Representation theory of SU(2) and SU(3) with applications to spin and quark models.

by

Shayan Farrokhi

to obtain the degree of Bachelor of Science in Applied Physics and Applied Mathematics
at the Delft University of Technology,

Student number:	4889827	
Project duration:	January 27, 2021 – 7 July 2021	
Thesis committee:	Dr. ir. W.G.M. Groenevelt,	TU Delft, supervisor
	Prof. Dr. Y. V. Nazarov,	TU Delft, supervisor
	Dr. A. R. Akhmerov,	TU Delft, committee member
	Dr. D. de Laat,	TU Delft, committee member

Abstract

In this thesis we provide an elementary introduction in finite dimensional representation theory of the Lie groups SU(2) and SU(3) for undergraduate students in physics and mathematics. We will also give two application of representation theory of these two groups in physics: the spin and quark models. We begin with first discussing representation theory for finite groups to create intuition for representations. We will explain notions such as intertwining maps and complete reducibility and we will mention some application of representation theory of finite groups in quantum mechanics. Hereafter, we begin with representation theory for the Lie groups and Lie algebras, especially of the groups SO(3) and SU(2), as these groups will play an important role in the description of spin. One of the main results is that SU(2) is the universal cover of SO(3). Furthermore, we give a description of spin by means of representation theory of SO(3) and its Lie algebra so(3). We will show that half integer representations of the Lie algebra so(3) cannot be exponentiated to representations of the Lie group SO(3), but it can be exponentiated to its unviversal cover SU(2). Moreover, we study the irreducible representations of SO(3) inside the Hilbert space $L^2(\mathbb{R}^3)$. We will argue that one of the simplest quantum Hilbert space of a particle $L^2(\mathbb{R}^3)$, can be modified to the tensor product $L^2(\mathbb{R}^3)\hat{\otimes}V$, where, V is a finite dimensional Hilbert space that incorporates the internal degrees of freedom: spin. V carries an irreducible projective representation of SO(3). We will also discuss the addition of angular momenuntum of two particles in quantum mechanics. For this, we show how the tensor product of irreducible representations V and W of so(3) decomposes into SO(3) invariant subspaces of $L^2(\mathbb{R}^3)$. Hereafter, we will turn to representation theory of the Lie group SU(3) for setting up the mathematical framework for analysing the quark model. We will proof that there is a one-to-one correspondence between the irreducible representations of sl(3;$\mathbb{C}$) and SU(3). We will also proof the theorem of the highest weight by which we can classify all the irreducible representations of SU(3) and sl(3;$\mathbb{C}$) by their highest weight. We will also introduce the notion of the Weyl group and show that the Weyl group is a symmetry of weights of the finite dimensional representation of sl(3;$\mathbb{C}$). Other properties of these representation, such as the dimension of the irreducible representations of sl(3;$\mathbb{C}$) will be provided. Lastly, the quark model is discussed by means of representition theory of SU(3). We will show how this model can be used to classify two type of particles which also interact by means of the strong force: baryons and mesons. We show that we can classify the lightest mesons and baryons in so-called multiplets by the irreducible representations of SU(3). However, we will also introduce a modification of the strong force which further refines this model.
A topic for further study would be how the symmetry group SU(3) describing Quantum Chromo Dynamics (QCD) can be used for the description of mesons and baryons. Also one could further study the whole symmetry group of the standard model: U(1)$\times$ SU(2)$\times$SU(3).

Contents

Introduction

Symmetries occur in the most unexpected places in nature: the mirror symmetry of a butterfly, the rotational symmetry of a flower to the translational symmetry of graphene. Moreover, in the 20th century symmetry became more and more important in physics. For example in the description of spin of particles, to the classification of elementary particles in the standard model, which describes the structure and interactions of these elementary particles. However, we would like a mathematical formulation to describe these symmetries as well.

An algebraic way to work with symmetries in mathematics is by means of Lie groups and Lie algebras. These mathematical structures were first introduced by the Norwegian mathematician Sophus Lie in the late nineteenth century and their symmetric nature made them particularly useful in the description of symmetry in physics. They can be studied by means of representation theory, that is, the elements of the Lie group or Lie algebra are continuously mapped to linear transformations of a vector space. Lie groups and Lie algebras have found their way in many applications in physics, such as the earlier mentioned classification of particles.

In the early 1960, it was a quite confusing time to be a particle physicist. There was an abundant list of elementary particles and more and more of them were being discovered. This time is usually described as the particle zoo era. Therefore, the quest of classifying these exotic particles began and it was not a trivial one. Ultimately, Murray Gell-Mann proposed a successful system based on representation theory of the symmetry Lie group $\mathrm{SU}(3)$ to classify a portion of the then known particles in so-called multiplets. With his model, called the Eightfold-Way because of a funny analogy with Buddhism, he even predicted the existence of a new particle which was discovered in 1964. For his efforts he received the Nobel price in physics in 1969.

The aim of this bachelor thesis is to provide an introduction, for undergraduate students in physics and mathematics, to finite dimensional representation theory of the symmetry groups $\mathrm{SU}(2)$ and $\mathrm{SU}(3)$. We will also provide two applications in physics: the spin and the quark models. This thesis is divided into 5 parts, where chapters 1, 2, 4 primarily give the reader a basic understanding in representation theory of $\mathrm{SU}(2)$ and $\mathrm{SU}(3)$ from a mathematical viewpoint, chapter 3 and 5 are dedicated to explaining the basics of the spin and quark models.

In chapter 1, the basic concepts of representations of finite groups and the axioms of quantum mechanics are given. Notions such as complete reducibility and character theory will be explained and applications of representation theory of finite groups will be mentioned. The purpose of this chapter is to introduce the basic concepts of representation theory for finite groups, which can be considered less abstract than representation theory for Lie groups and Lie algebras. Therefore, it is perfect for providing intuition concerning representation theory.

In chapter 2, we give an introduction to Lie groups and Lie algebras and the connection between them. We restrict ourselves to matrix Lie groups and focus on the groups $\mathrm{SO}(3)$ and $\mathrm{SU}(2)$. This focus is motivated by the fact that these groups play a central role in the description of spin, which will be treated in chapter 3.

In chapter 3 we will give our first physical application of representation theory: spin. We will define what spin is by means of representation theory of the Lie groups $\mathrm{SO}(3)$, $\mathrm{SU}(2)$ and their respective Lie algebras $\mathrm{so}(3)$ and $\mathrm{su}(2)$. In elementary quantum mechanics particles live inside the Hilbert space $\text{Ł}^2(\mathbb{R}^3)$. Therefore, our main goal is to obtain the irreducible representations of $\mathrm{SO}(3)$ inside $L^2(\mathbb{R}^3)$ and connect this to the notion of spin.

In chapter 4 we discuss representation theory for the Lie group SU(3). The purpose of this chapter is setting up the mathematical framework for analysing the quark model. Hereafter, in chapter 5, we give a short description of flavour symmetry of the lightest mesons and baryons by means of representation theory of SU(3). Lastly, in chapter 6, the limitations of these models will be discussed and some suggestions for further study will be given.

This report was written as bachelor thesis project for the degree of Applied Mathematics and Applied Physics at Delft University of Technology. The ideas introduced in this report come from several sources that can be found in the Bibliography. I claim in no way that the ideas introduced in this report are my own original work. I hope that you will enjoy reading this thesis.

1

Representions of Finite Groups

This chapter is dedicated to setting up the right framework of representation theory and quantum mechanics in which we will work. The purpose of this chapter is to give some insight in finite dimensional representations and character theory. To be precise, Schur's orthogonality relations are going to be proved, as well as the orthogonality relations of characters of groups with finite elements. Moreover, the notion that the irreducible representations are an orthonormal basis in the space of class functions is going to be proved as well. Several applications and examples will be provided throughout this chapter to make the several notions related to finite-dimensional representations more insightful. This will be convenient for creating intuition, since in later chapters we will consider representations of topological groups with countable cardinality which are considered more abstract. At the end of this chapter, the basic notions of quantum mechanics will be explained from a mathematical viewpoint. In this chapter I follow the works of [13], [21], [22] and [11]. Unless stated otherwise, G denotes a finite group.

1.1. Basic definitions

Definition 1.1. *A representation of a finite group G on a finite-dimensional complex vector space V is a homomorphism $\phi : G \to GL(V)$ of G to the group of automorphisms of V.*

Oftentimes we say that V itself is a representation of G when there is little doubt about the map ϕ. The degree of ϕ is the dimension of V. However, when studying such maps one would like to know whether these maps are "the same" or not.

Definition 1.2. *Two representations $\phi : G \to GL(V)$ and $\psi : G \to GL(W)$ are said to be equivalent if there exists an isomorphism $T : V \to W$ such that $\psi_g = T\phi_g T^{-1}$ for all $g \in G$. In this case, we write $\phi \sim \psi$.*

We can visualise the above definition in the diagram below. For T being a isomorphism this diagram has to commute.

$$\begin{array}{ccc} V & \xrightarrow{\quad T \quad} & W \\ {\scriptstyle \phi_g}\downarrow & & \downarrow{\scriptstyle \psi_g} \\ V & \xrightarrow[\quad T \quad]{} & W \end{array}$$

There are several more definitions which are going to be used frequently, one of them is the notion of an invariant subspace.

Definition 1.3. *Let $\phi : G \to GL(V)$ be a representation. A subspace $W \subset V$ is G-invariant if, for all $g \in G$ and $w \in W$, one has $\phi_g w \in W$.*

There are several ways to construct new representations out of existing ones, one could think of constructs such as the direct sum or the tensor product.

Definition 1.4. *Suppose that representations $\phi^{(1)} : G \to GL(V_1)$ and $\phi^{(2)} : G \to GL(V_2)$ are given. Then their (external) direct sum*

$$\phi^{(1)} \oplus \phi^{(2)} : G \to GL(V_1 \oplus V_2)$$

is given by

$$\left(\phi^{(1)} \oplus \phi^{(2)}\right)_g (v_1, v_2) = \left(\phi_g^{(1)}(v_1), \phi_g^{(2)}(v_2)\right)$$

In chapter 2 and chapter 3, we will see that when describing a system in quantum mechanics consisting of two particles, we need a mathematical tool called the "tensor product" (see Appendix B). Therefore, it is natural to introduce the notion of a tensor product representation.

Definition 1.5. *Let $\phi^{(1)} : G \to GL(V_1)$ and $\phi^{(2)} : G \to GL(V_2)$ be two finite dimensional linear representations of a group G. For $g \in G$ define an element ϕ_g of $GL(V_1 \otimes V_2)$ by the condition:*

$$\phi_g(v_1 \otimes v_2) = \phi_g^{(1)} v_1 \otimes \phi_g^{(2)} v_2$$

for $v_1 \in V_1$ and $v_2 \in V_2$. Hence, we write: $\phi_g = \phi_g^{(1)} \otimes \phi_g^{(2)}$. For the existence and uniqueness of this operator see Appendix B.

Note that we are working in a finite dimensional setting. Thus we can choose a basis (e_{i_1}) for V_1 and (e_{i_2}) for V_2 respectively. Then we note that we can express $\phi_g^{(1)} \in \mathrm{End}(V_1)$ with a matrix $(\phi^{(1)}_{i_1 j_1})$ with respect to the basis (e_{i_1}) and do the same for V_2 as well with matrix $(\phi^{(2)}_{i_2 j_2})$. Then we see that

$$\phi_g^{(1)}(e_{j_1}) = \sum_{i_1} \phi^{(1)}_{i_1 j_1}(g)\, e_{i_1} \text{ and } \phi_g^{(2)}(e_{j_2}) = \sum_{i_2} \phi^{(2)}_{i_2 j_2} e_{i_2},$$

which imply that

$$\phi_g(e_{j_1} \otimes e_{j_2}) = \sum_{i_1, i_2} \left(\phi^{(1)}_{i_1 j_1}(g) \cdot \phi^{(2)}_{i_2 j_2}(g)\right) e_{i_1} \otimes e_{i_2}.$$

Here, the matrix of ϕ_g is $\phi_g^{(1)} \otimes \phi_g^{(2)}$ is usually known as the Kronecker tensor product. See for example [18] for less formal treatment of the Kronecker tensor product.

Just as in algebra where one can express a number in $\mathbb{Z}$ as a factor of prime numbers, we are interested in whether a representation can be expressed with similar entities which are atomic in some sense. These entities are usually called irreducible representations.

Definition 1.6. *A non-zero representation $\phi : G \to GL(V)$ of a group G is said to be irreducible if the only G-invariant subspaces of V are $\{0\}$ and V.*

As we have remarked, just as a number in $\mathbb{Z}$ can be decomposed in prime factors, we are interested whether a representation of a finite group G given by $\rho : G \to GL(V)$ can be decomposed into irreducible representations. Hence, the following two definitions follow quite naturally.

Definition 1.7. ***(Complete reducibility)***
Let G be a group. A representation $\phi :\to GL(V)$ is said to be completely reducible if $V = V_1 \oplus \ldots \oplus V_n$ where the V_i are G-invariant subspaces and $\phi|_{V_i}$ is irreducible for all $i = 1, \ldots, n$.

Definition 1.8. *A non-zero representation ϕ of a group G is called decomposable if $V = V_1 \oplus V_2$ with V_1, V_2 non-zero G-invariant subspaces. Otherwise, V is called indecomposable.*

1.1.1. Complete Reducibility

In this section, I mostly follow section 3.2 of [22]. Reading the previous section, we may ask ourselves: "What are the conditions for a representation of finite group G to be completely reducible?". This is a quite interesting question, which will play an important role in analysing spin in chapter 3, but for Lie groups instead of finite groups. However, before we can answer this question, we need to introduce the notion of unitary representations.

Definition 1.9. ***(Unitary representation)*** *Let V be an inner product space. A representation $\phi : G \to GL(V)$ is said to be unitary if ϕ_g is unitary for all $g \in G$, i.e.,*

$$\left\langle \phi_g(v), \phi_g(w) \right\rangle = \langle v, w \rangle$$

for all $v, w \in W$. In other words, we may view ϕ as a map $\phi : G \to U(V)$.

As we can see from the definition of unitary representation, it preserves the inner product. This will play a huge role in physics (and later chapters), as one is often interested in unitary representations which preserve inner products of the spaces describing the physical systems, which are often Hilbert spaces.

Proposition 1.10. *Let $\phi : G \to GL(V)$ be a unitary representation of a group. Then ϕ is either irreducible or decomposable.*

Proof. Suppose ϕ is not irreducible. Then there is a non-zero G-invariant subspace W of V. Its orthogonal complement is also non zero and $V = W^\perp \oplus W$. Take an arbitrary elements from $w \in W^\perp$ and $v \in W$ and note that

$$\begin{aligned} \langle \phi_g(w), v \rangle &= \langle \phi_{g^{-1}} \phi_g(w), \phi_{g^{-1}}(v) \rangle \\ &= \langle w, \phi_{g^{-1}}(v) \rangle \\ &= 0 \end{aligned}$$

So $W^\perp$ is also G-invariant. Thus, ϕ is decomposable. □

What is interesting to note is that every finite dimensional representation of a finite group is in fact equivalent to a unitary one. Therefore, when we want to analyse an arbitrary representation of a finite group, we can make use of the fact that it is equivalent to a unitary one and use all the properties of a unitary representations.

Proposition 1.11. *Every representation of a finite group G is equivalent to a unitary representation.*

Proof. Let $\phi : G \to GL(V)$ be a representation where $\dim(V) = n$. We choose a basis B for V, and let $T : V \to \mathbb{C}^n$ be the isomorphism taking coordinates with respect to B. Then we can set $\rho_g = T \phi_g T^{-1}$ for $g \in G$, which yields a representation $\rho : G \to GL_n(\mathbb{C})$ equivalent to ϕ. Let $\langle \cdot, \cdot \rangle$ be the standard inner product on $\mathbb{C}^n$. We can define a new inner product $(\cdot, \cdot)$ on $\mathbb{C}^n$ in the following way:

$$(v, w) = \sum_{g \in G} \langle \rho_g v, \rho_g w \rangle$$

Such a summation is not a problem as G is finite. Later in this chapter we will introduce tricks for the case that G is infinite. This is indeed an inner product, as can easily be checked. For the unitary part we note the following. Let $h \in G$,

$$(\rho_g v, \rho_v w) = \sum_{g \in G} \left\langle \rho_g \rho_v v, \rho_g \rho_v v \right\rangle = \sum_{g \in G} \langle \rho_{gh} v, \rho_{gh} w \rangle.$$

Set $x = gh$. As g ranges over all G, we know that x ranges over all G as well. Hence

$$(\rho_h v, \rho_h w) = \sum_{x \in G} \langle \rho_x v, \rho_x w \rangle = (v, w).$$

This concludes the proof. □

Corollary 1.12. *Let $\phi : G \to GL(V)$ be a non-zero representation of a finite group. Then ϕ is either irreducible or decomposable.*

Proof. By the previous proposition 1.11 ϕ is equivalent to a unitary representation ρ. However, we know that every unitary representation is decomposable or irreducible by the previous two propositions unitaryrepresentation and 1.10. Hence, ϕ is equivalent to an irreducible or decomposable representation and thus irreducible or decomposable itself. □

We have now the tools for proving the theorem which is at the heart of this section.

Theorem 1.13. *Every representation of a finite group is completely reducible.*

Proof. Let $\phi : G \to GL(V)$ be a representation of a finite group G. The proof is based on induction on the degree of ϕ. Let $\dim(V) = 1$, then ϕ is irreducible since V has no non-trivial invariant subspaces. Now assume that this statement is true for $\dim(V) \leq n$. Let $\phi : G \to GL(V)$ be a representation with $\dim(V) = n+1$. If ϕ is irreducible, we are done. So assume it is not. We know that it decomposable by the previous corollary. Hence $V = V_1 \oplus V_2$ where $0 \neq V_1, V_2$ are G-invariant subspaces. We know that $\dim(V_1)$ and $\dim(V_2)$ are less than $\dim(V)$. Then we can apply the induction hypothesis and note that $\phi|_{V_1}$ and $\phi|_{V_2}$ are completely reducible. Hence, $V_1 = U_1 \oplus ... \oplus U_s$ and $V_2 = W_1 \oplus ... \oplus W_r$, where U_i, W_j are G-invariant subrepresentations and the subrepresentations $\phi|_{U_i}$ and ϕ_{W_j} are irreducible for all $i \leq i \leq s, 1 \leq j \leq r$. Then $V = U_1 \oplus ... \oplus U_s \oplus W_1 \oplus ... \oplus W_r$ and ϕ is thus completely reducible. □

This theorem tells us the following: when we want to study an arbitrary representation of a finite group G, it is enough to study the irreducible representations of the group, as the arbitrary one can be broken down into irreducible representations. Hence, when studying arbitrary representations one can focus on the irreducible ones.

1.2. Schur's Lemma and Its Applications

In this section, I mostly follow section 4.1 of [22]. We are interested in representations of a group G, but not all of them. We want a tool by which we can make distinction between two arbitrary representations of the same group. This tool is usually the intertwining map between two representations V and W of an arbitrary group G.

Definition 1.14. *Let $\phi : G \to GL(V)$ and $\rho : G \to GL(W)$ be representations. An intertwining map from ϕ to ρ is by definition a linear map $T : V \to W$ such that $T\phi_g = \rho_g T$ for all $g \in G$. The set of all intertwining maps from ϕ to ρ of representations of the group G is denoted $Hom_G(\phi, \rho)$.*

The next theorem will play a crucial role in later sections and derivations of the Schur's orthogonality relations and the orthogonality relations of characters in later sections.

Theorem 1.15. ***(Schur's Lemma)*** *Let ϕ, ρ be irreducible representations of G, and $T \in Hom_G(\phi, \rho)$. Then either T is invertible or $T = 0$. Consequently:*

- *If ϕ is not isomorphic to ρ, then $Hom_G(\phi, \rho) = 0$.*
- *If $\phi = \rho$, then $T = \lambda I$ with $\lambda \in \mathbb{C}$ (i.e., T is multiplication by a scalar).*

The proof of this theorem will be provided for the case of Lie groups in chapter 2, although the proof is in no way difficult at all. For now note that the number of ways one can construct intertwining maps between two irreducible representations is very limited. As promised, we will give an application of Schur's lemma, specifically to abelean groups.

Corollary 1.16. *Let G be an abelean group. Then any irreducible representation of G has degree one.*

Proof. Let $\phi : G \to GL(V)$ be an irreducible representation. Fix an arbitrary $h \in G$ and set $T = \phi_h$. Then we can see that

$$T\phi_g = \phi_h\phi_g = \phi_{hg} = \phi_{gh} = \phi_g\phi_h = \phi_g T.$$

Then $\phi_h = \lambda_h I$ by Schur's lemma for some $\lambda_h \in \mathbb{C}$. Then note that for any $v \in V$ and $k \in \mathbb{C}$ we have

$$\phi_h(kv) = \lambda_h Ikv = \lambda_h kv \in \mathbb{C}v.$$

Hence, $\mathbb{C}v$ is a G-invariant subspace, as h was arbitrary. Hence $V = \mathbb{C}v$ by irreducibility and thus $\dim(V) = 1$. □

At the beginning of this chapter we mentioned that we would provide some examples and applications. The previous corollary has some interesting applications in linear algebra.

Corollary 1.17. *Let G be a finite abelian group and $\phi : G \to GL_n(\mathbb{C})$ a representation. Then there is an invertible matrix T such that $T^{-1}\phi_g T$ is diagonal for all $g \in G$.*

Proof. As ϕ is a finite dimensional representation it is completely reducible and hence it is $\phi \cong \phi^{(1)} \oplus ... \oplus \phi^{(m)}$ where $\phi^{(1)}, ..., \phi^{(m)}$ are irreducible. We know that G is abelean, thus the degree of each $\phi^{(i)}$ is equal to 1. Thus $\phi_g^{(i)} \in \mathbb{C}^*$ for all $g \in G$. If $T : \mathbb{C}^n \to \mathbb{C}^n$ is the intertwining map of ϕ with $\phi^{(1)} \oplus ... \oplus \phi^{(n)}$, then

$$T^{-1}\phi_g T = \begin{bmatrix} \phi_g^{(1)} & & \\ & \ddots & \\ & & \phi_g^{(n)} \end{bmatrix}$$

is diagonal for all $g \in G$

□

Corollary 1.18. *Let $A \in GL_m(\mathbb{C})$ be a matrix of finite order. Then A is diagonalizable. Moreover, if $A^n = I$, then the eigenvalues of A are nth-roots of unity.*

Proof. Assume that $A^n = I$. Define $\phi : \mathbb{Z}/n\mathbb{Z} \to GL_m(\mathbb{C})$ by $\phi([k]) = A^k$. With some elementary group theory, this is easily seen to give a well-defined representation just by looking at the orders of the elements which are mapped to one another. Hence, there exists an $T \in GL_m(\mathbb{C})$ such that $T^{-1}AT$ is diagonal by the previous corollary.

Suppose

$$T^{-1}AT^1 = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{bmatrix} = D$$

Then $D^n = (T^{-1}AT)^n = I$. Hence, we have

$$\begin{bmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_m^n \end{bmatrix} = D^n = I$$

and thus $\lambda_i^n = 1$ for all $i \in \{1, ..., m\}$. □

1.2.1. Orthogonality Relations

The purpose of this section is to give some orthogonality relations which we will use quite frequently from this point, especially for the treatment of character theory in the next section. Note that we still work with groups that are bounded in cardinality, thus we do not have to think of problems such as convergence when summing over all the elements of the representation of the group. The orthogonality relations we are proving in this section are called the Schur's orthogonality relations for groups of finitely many elements. From this point on we will mostly follow sections 2.2-2.5 of [21].

Lemma 1.19. *Suppose we have a group G which is finite and two irreducible representations $\phi^{(1)}: G \to GL(V_1)$ and $\phi^{(2)}: G \to GL(V_1)$. Let h be a linear mapping of V_1 into V_2, and put:*

$$h^0 = \frac{1}{|G|} \sum_{g \in G} (\phi_g^{(2)})^{-1} h \phi_g^{(1)}.$$

Then:

1. *If $\phi^{(1)}$ and $\phi^{(2)}$ are not isomorphic, then $h^0 = 0$.*
2. *If $V_1 = V_2$ and $\phi^{(1)} \cong \phi^{(2)}$, h_0 is a scalar multiple of the identity of ratio $\frac{1}{n}Tr(h)$, with $n = \dim(V_1)$.*

Proof. Note that $\phi_g^{(1)} h^0 = h^0 \phi_g^{(2)}$:

$$\left(\phi_g^{(2)}\right)^{-1} h_0 \phi_g^{(1)} = \frac{1}{|G|} \sum_{t \in G} \left(\phi_g^{(2)}\right)^{-1} \left(\phi_t^{(2)}\right)^{-1} h \phi_t^1 \phi_g^{(1)} = \frac{1}{|G|} \sum_{t \in G} \left(\phi_{tg}^{(2)}\right)^{-1} h \phi_{ts}^{(1)} = h^0.$$

Note that by Schur's lemma we have that $h^0 = 0$ in case (1). In case (2) we see that h^0 is equal to a scalar multiple of the identity. In the latter case, we can see that

$$\mathrm{Tr}\left(h^0\right) = \frac{1}{|G|} \sum_{g \in G} \mathrm{Tr}\left(\left(\phi_g^{(1)}\right)^{-1} h \phi_g^{(1)}\right) = \mathrm{Tr}(h).$$

Let λ be the scalar multiple of the identity. Since $\mathrm{Tr}(\lambda I) = n\lambda$, we obtain that $\lambda = \frac{\mathrm{Tr}(h)}{n}$. □

We said that all representations are of finite dimension. Hence, we can write $\phi^{(1)}$ and $\phi^{(2)}$ in matrix form:

$$\phi_g^{(1)} = \left(\phi_{i_1 j_1}^{(1)}(g)\right), \qquad \phi_g^{(1)} = \left(\phi_{i_2 j_2}^{(2)}(g)\right).$$

We can also say that h and h^0 are defined by matrices $(x_{i_2 i_1})$ and $(x_{i_2 i_1}^0)$ respectively. We have

$$x_{i_2 i_1}^0 = \frac{1}{|G|} \sum_{g, j_1, j_2} \phi_{i_2 j_2}^{(2)}(g^{-1}) x_{j_2 j_1} \phi_{i_1 j_1}^{(1)}(g).$$

Note that the right hand side is a linear form with respect to $x_{j_2 j_1}$. For case (1) of lemma 1.19, this linear form vanishes for all systems of values $x_{j_2 j_1}$. Hence, its coefficients are zero. Therefore, the following corollary is a direct consequence of this.

Corollary 1.20. *For the case that $\phi^{(1)}$ and $\phi^{(2)}$ are as defined before, we have*

$$\frac{1}{|G|} \sum_{g \in G} \phi_{i_2 j_2}^{(2)}(g^{-1}) \phi_{j_1 i_1}^{(1)}(g) = 0,$$

which holds for arbitrary i_1, i_2, j_1, j_2

Note that in the second case we have that $h^0 = \lambda I$, hence $\left(x_{i_2 i_1}^0\right) = \lambda \delta_{i_2 i_1}$, with $\lambda = \frac{1}{n}\mathrm{Tr}(h)$. The latter expression is $\lambda = \frac{1}{n} \sum_{j_1, j_2} \delta_{j_2 j_1} x_{j_2 j_1}$. Hence, we obtain the equality:

$$\frac{1}{|G|} \sum_{g, j_1, j_2} \phi_{i_2 j_2}^{(2)}(g^{-1}) x_{j_2 j_1} \phi_{j_1 i_1}^{(1)}(g) = \frac{1}{n} \sum_{j_1, j_2} \delta_{i_2 i_1} \delta_{j_2 j_1} x_{j_2 j_1}.$$

When we now equate the coefficients of $x_{j_2 j_1}$ we obtain the following result.

Corollary 1.21. *In case (2) of lemma 1.19, if $\phi^{(1)} \cong \phi^{(2)}$, we have:*

$$\frac{1}{|G|} \sum_{g \in G} \phi_{i_2 j_2}^{(2)}(g^{-1}) \phi_{j_1 i_1}^{(1)}(g) = \frac{1}{n} \delta_{i_2 i_1} \delta_{j_2 j_1} = \begin{cases} \frac{1}{n} & \text{if } i_1 = i_2 \text{ and } j_1 = j_2 \\ 0 & \text{else.} \end{cases}$$

A few interesting remarks can be made. First of all if ϕ and ψ are arbitrary complex valued functions $G \to \mathbb{C}$, set

$$\langle \phi, \psi \rangle = \frac{1}{|G|} \sum_{t \in G} \phi(t^{-1}) \psi(t) = \frac{1}{|G|} \sum_{t \in G} \phi(t) \psi(t^{-1}).$$

We have $\langle \phi, \psi \rangle = \langle \psi, \phi \rangle$. Furthermore, from the definition it is clear that $\langle \psi, \phi \rangle$ is linear in both ψ and ϕ. The corollaries 2 and 3 become respectively

$$\left\langle \phi^{(2)}_{i_2 j_2}, \phi^{(1)}_{j_1 i_1} \right\rangle = 0 \text{ and } \left\langle \phi^{(2)}_{i_2 j_2}, \phi^{(1)}_{j_1 i_1} \right\rangle = \frac{1}{n} \delta_{i_2 i_1} \delta_{j_2 j_1}.$$

Note that all finite dimensional representations are equivalent to a unitary representation by proposition 1.11. Hence we can realize that the matrices $(\phi_{ij}(g))$ are unitary. Then $\phi_{ij}(g^{-1}) = \phi_{ji}(g)^*$ and corollaries 1.20 and 1.21 are the orthogonality relations for (ϕ, ψ), as defined in section 1.3.1, which one usually see in literature such as [13], [21], [22].

1.3. Character Theory

Character theory is a beautiful theory to classify finite dimensional representations and has many, many applications in physics. What we will do in the next section is giving the basics of character theory and mention the applications of character theory in physics. We begin with the definition of character.

Definition 1.22. *If $\phi : G \to GL(V)$ is a linear representation of a finite group G, its character χ_ϕ is the complex-valued function on the group defined by*

$$\chi_\phi(g) = Tr(\phi_g),$$

the trace of ϕ_g on V.

Note that $\dim(V) = \chi_\phi(1)$. One of the identities of a character which makes it so powerful is that the character is the same for all conjugacy classes. Hence, $\chi_\phi(hgh^{-1}) = \chi_\phi(g)$. We demonstrate this as follows:

$$\chi_\phi(\phi_{hgh^{-1}}) = \mathrm{Tr}(\phi_h \phi_g \phi_{h^{-1}}) = \mathrm{Tr}(\phi_{h^{-1}} \phi_h \phi_g) = \mathrm{Tr}(\phi_g) = \chi_\phi(g).$$

More interestingly, we can derive some properties of the characters of a direct sum and tensor product of representations.

Proposition 1.23. *Let $\phi^{(1)} : G \to GL(V_1)$ and $\phi^{(2)} : G \to GL(V_2)$ be two linear representations of G, and let χ_1 and χ_2 be their characters. Then:*

1. *The character χ of the direct sum representation $V_1 \oplus V_2$ is equal to $\chi_1 + \chi_2$.*
2. *The character χ of the tensor product representation $V_1 \otimes V_2$ is equal to $\chi_1 \cdot \chi_2$*

Proof. Let $\phi^{(1)}_g$ and $\phi^{(2)}_g$ be given in their matrix notation R^1_g and R^2_g. Then the representation is given by

$$R_g = \begin{pmatrix} R^1_g & 0 \\ 0 & R^2_g \end{pmatrix}.$$

We can then conclude that $\mathrm{Tr}(R_g) = \mathrm{Tr}(R^{(1)}_g) + \mathrm{Tr}(R^{(2)}_g)$. Thus $\chi(g) = \chi_1(g) + \chi_2(g)$.

For the second part, note that

$$\chi_1(g) = \sum_{i_1} \phi^{(1)}_{i_1 i_1}(g), \qquad \chi_2(g) = \sum_{i_2} \phi^{(2)}_{i_2 i_2}(g).$$

With the matrix convention we have introduced in definition 1.5 it is easy to see that

$$\chi(g) = \sum_{i_1 i_2} \phi^{(1)}_{i_1 i_1}(g) \phi^{(2)}_{i_2 i_2}(g) = \chi_1(g) \chi_2(g).$$

□

1.3.1. Orthogonality Relations for Characters

We are going to derive some orthogonality relations of character with respect to their representations. We see that characters give us quite a lot of information whether two irreducible representations are isomorphic or not.

When ψ, χ are two complex valued functions such that $\psi, \chi : G \to \mathbb{C}$, then we can define an inner product

$$(\psi|\chi) = \frac{1}{|G|} \sum_{g \in G} \psi(g) \chi(g)^* .$$

This map is clearly linear in ϕ, but semi-linear (linear up to a field automorphism of $\mathbb{C}$) in χ. Also note that $(\psi|\psi) > 0$ for all $\psi \neq 0$ and that $(\psi|\psi) = 0$ implies that $\psi = 0$ since $|\psi(g)|^2 \geq 0$ for all $g \in G$.

We now switch to χ being a character. One can see that $\chi_\phi(g^{-1}) = \chi_\phi(g)^*$ in the following way. We note that, again, we are working with representations of finite order and by corollary 1.18, that the eigenvalues of the matrix representation ϕ_g are the n-th roots of unity. Hence, we have that

$$\chi_\phi(g)^* = \mathrm{Tr}(\phi_g)^* = \sum_i \lambda^* = \sum_i \lambda^{-1} = \mathrm{Tr}\left(\phi_g^{-1}\right) = \mathrm{Tr}\left(\phi_{g^{-1}}\right) = \chi_\phi(g^{-1}).$$

From this, we can see that $\chi_\phi(g)^* = \chi_\phi(g^{-1})$. Hence, we have

$$(\phi|\chi) = \frac{1}{|G|} \sum_{g \in G} \phi(g) \chi(g)^* = \frac{1}{|G|} \sum_{g \in G} \phi(g) \chi(g^{-1}) = \langle \phi, \chi \rangle.$$

Therefore, we have $(\phi|\chi) = \langle \phi, \chi \rangle$ for all $\phi : G \to \mathbb{C}$. This holds as long as χ is a character. The next theorem is the first orthogonality relation for characters.

Theorem 1.24.

- *If χ is the character of an irreducible representation, we have $(\chi|\chi) = 1$, i.e. χ has norm 1.*
- *If χ and χ' are the characters of two nonisomorphic irreducible representations, we have $(\chi|\chi') = 0$. (i.e. χ and χ' are orthogonal).*

Proof. Let ϕ be an irreducible representation with character χ. Thus for all $g \in G$ we have $\chi(g) = \sum_i \phi(g)_{ii}$. Thus,

$$(\chi|\chi) = \langle \chi, \chi \rangle = \sum_{g \in G} \chi(g) \chi(g^{-1}) = \frac{1}{|G|} \sum_{i,j,g} \phi_{ii}(g) \phi_{jj}(g^{-1}) = \sum_{i,j} \langle \phi_{ii}, \phi_{jj} \rangle.$$

We can use corollary 1.21 to note that $\langle \phi_{ii}, \phi_{jj} \rangle = \frac{\delta_{ij}}{n}$. Here, n is the degree of ϕ. Hence, we can see that

$$(\chi|\chi) = \frac{\left(\sum_{i,j} \delta_{ij}\right)}{n} = \frac{n}{n} = 1.$$

Note that point 2 is proved along the same lines, just use corollary 1.20. □

In theorem 1.13 we proved that an arbitrary representation breaks down into irreducible ones. However, we did not specify how many times these irreducible relations occur in the decomposition.

Theorem 1.25. *Let V be a linear representation of G, with character χ_ϕ, and suppose that V decomposes into a direct sum of irreducible representations:*

$$V = W_1 \oplus \ldots \oplus W_k.$$

Then, if W is an irreducible representation with character χ, the number of W_i isomorphic to W is equal to the scalar product $(\chi_\phi|\chi) = \langle \chi_\phi|\chi \rangle$.

Proof. Let χ_i be the character of W_i. By proposition 1.23 we have

$$\chi_\phi = \chi_1 + ... + \chi_k.$$

Thus $(\chi_\phi|\chi) = (\chi_1|\chi) + ... + (\chi_k|\chi)$. The result follows by the previous theorem. □

Just as one can express a number in $\mathbb{Z}$ in prime numbers uniquely up to order of the prime numbers, we can state a similar result for the decomposition of an arbitrary representation.

Corollary 1.26. *The number of W_i isomorphic to W does not depend on the chosen decomposition.*

Proof. Note that $(\chi|\chi_i)$ does not depend on the decomposition of the previous theorem. □

Moreover, we can also check whether two representations are isomorphic or not.

Corollary 1.27. *Two representations with the same character are isomorphic.*

Proof. This trivially follows from corollary 1.26. Note that each irreducible representation occurs the same number of times. □

Lastly, with the aid of characters we can check whether a representation V is irreducible or not. Before we proof this we need some additional observations first. By the previous two theorems we can conclude that the study of representations can be reduced to that of their characters. If $\chi_1, ..., \chi_h$ are the distinct irreducible characters of G and $W_1, .., W_h$ denote the corresponding representations, then each representation V is isomorphic to

$$V = m_1 W_1 \oplus ... \oplus m_h W_h \text{ with } m_i \text{ integers } \geq 0.$$

The character of V is equal to $m_1\chi_1 + ... + m_h\chi_h$ and the orthogonality relation among the χ_i implies $(\phi|\phi) = \sum_{i=1}^h m_i^2$.

Theorem 1.28. *If χ is the character of a representation V, $(\chi|\chi)$ is a positive integer and $(\chi|\chi) = 1$ if and only if V is irreducible.*

Proof. From the previous we can conclude that $\sum_i m_i^2$ is equal to 1 if and only if one of the m_i is equal to one and the others are equal to 0. That is, V is equal to one of the W_i. □

To conclude, as said before, the study of representations can be reduced to that of their characters, which makes character theory so powerfull.

1.4. Decomposition of the Regular Representation

Before we plunge in the analysis of the number of irreducible representations of a group G, we first need to introduce the concept of regular representation. Let $|G|$ be the order of the group G and let V be a finite dimensional vector space of dimension $|G|$ with a basis $(e_t)_{t\in G}$ indexed by the elements t of G. Let $\rho_s \in \mathrm{End}(V)$ with $s \in G$ be the representation of G such that e_t is mapped to e_{st}. This representation is oftentimes called the regular representation of G. Its degree is $|G|$ and note that $e_s = \rho_s(e_1)$. Hence the images of e_1 from a basis of V. We first have to make a claim concerning the character of the regular representation.

Proposition 1.29. *The character r_G of the regular representation is given by the formulas:[21]*

$$r_G(1) = |G|, \qquad r_G(s) = 0 \; if \; s \neq 1.$$

Corollary 1.30. *Every irreducible representation W_i is contained in the regular representation with multiplicity equal to its degree n_i.*

Proof. By theorem 1.25 this number is equal to $\langle r_G, \chi_i \rangle$ and this number is

$$\langle r_G, \chi_i \rangle = \frac{1}{|G|} \sum_{s\in G} r_G\left(s^{-1}\right) \chi_i(s) = \frac{1}{|G|} |G| \cdot \chi_i(1) = \chi_i(1) = n_i.$$

□

Corollary 1.31.

1. *The degrees n_i of irreducible representations of G satisfy the relation $\sum_{i=1}^{h} n_i^2 = |G|$.*

2. *If $s \in G$ is different from 1, we have $\sum_{i=1}^{h} n_i \chi_i(s) = 0$.*

Proof. This immediately follows from the previous corollary. □

Note that the above results can be used to determine the irreducible representations of a group G. For example, when one has found non-isomorphic irreducible representations of degrees $n_1, ..., n_k$ such that $n_1^2 + ... + n_k^2 = |G|$, then we have found all.

1.5. Number of Irreducible Representations

With the notions we have introduced thus far we are now in the position to make claims about the number of irreducible representation of a group G. For this end, we will make use of the class functions of the group G.

Proposition 1.32. *Let f be a class function on G, and let $\rho : G \to GL(V)$ be a linear representation of G. Let ρ_f be the linear mapping of V into itself defined by:*

$$\rho_f = \sum_{t \in G} f(t) \rho_t.$$

If V is irreducible of degree n and character χ, then ρ_f is a scalar multiple of the identity of ratio λ given by:

$$\lambda = \frac{1}{n} \sum_{t \in G} f(t)\chi(t) = \frac{|G|}{n}(f|\chi^*).$$

Proof. Note that $\rho_s^{-1} \rho_f \rho_s$ is given by

$$\rho_s^{-1} \rho_f \rho_s = \sum_{t \in G} f(t) \rho_s^{-1} \rho_t \rho_s = \sum_{t \in G} f(t) \rho_{s^{-1}ts}.$$

We can make the substitution $x = s^{-1}ts$, this will give

$$\rho_s^{-1} \rho_f \rho_s = \sum_{x \in G} f(sxs^{-1}) \rho_x = \sum_{x \in G} f(x) \rho_x = \rho_f.$$

Hence, this result implies that $\rho_f \rho_s = \rho_s \rho_f$. By Schur's lemma, this shows that ρ_f is a scalar multiple of the identity, denoted by λI. We know that the trace of the scalar multiple of the identity is given by λn. The trace of ρ_f is given by $\sum_{t \in G} f(t) \mathrm{Tr}(\rho_t) = \sum_{t \in G} f(t)\chi(t)$. From this point we can easily verify the relation given in the proposition. Note that

$$\lambda = \frac{1}{n} \sum_{t \in G} f(t)\chi(t) = \left(f|\chi^*\right).$$

□

Now we have come to one of the most interesting observations of this chapter

Theorem 1.33. *The characters $\chi_1, ..., \chi_h$ form an orthonormal basis of H. Here H is the space of class functions on the group G.*

Proof. It is already clear from theorem 1.24 that the χ_i form an orthonormal set in H. It remains to proof that the χ_i from a basis for H. Note that it is enough to proof that every element orthogonal to the χ_i^* is zero. We set f to be such an element. For each representation ρ of G, we set $\rho_f = \sum_{t \in G} f(t)\rho_t$. Since f is orthogonal to the χ_i^*, we know by the previous proposition 1.32 that ρ_f is zero as long as ρ is irreducible. Since every finite group representation can be decomposed as the direct sum of irreducible representations, we can conclude that ρ_f is always zero, even though ρ might not be irreducible.

We are going to apply this to the regular representation R. We take the basis vector e_1 under ρ_f to obtain

$$\rho_f e_1 = \sum_{t \in G} f(t) \rho_t e_1 = \sum_{t \in G} f(t) e_t.$$

Note that ρ_f is always zero, hence, $\rho e_1 = 0$ and therefore $f(t) = 0$ for all $t \in G$. Therefore, $f = 0$ and this concludes the proof. □

From the previous theorem it quite naturally follows that the number of irreducible representations of a group G is equal to the number of conjugacy classes of G, which is actually quite a remarkable result.

Theorem 1.34. *The number of irreducible representations of G up to isomorphism is equal to the number of classes of G.*

Proof. Let $C_1, ..., C_k$ be the distinct classes of G. To say that a function f on G is a class function is equivalent to the restriction for f being constant on each of the C_i with $i \in \{1, ..., k\}$. Hence, the space of class functions is equal to the number of distinct classes of G, which is k. The previous theorem told us that the number of irreducible representations up to isomporphism is equal to the number of conjugacy classes k. Hence, the result follows. □

With this last result, we end this analysis of representation theory of groups with finite elements. We will introduce some more involved definitions in later chapters with the aid of the definitions and results obtained in these sections.

1.6. Quantum Mechanics Preliminaries

As said in the beginning of this chapter, we have to introduce the basic notions of quantum mechanics before proceeding further. This section is dedicated to introducing the axioms of quantum mechanics that are of interest for us in later chapters. Note that these axioms are not axioms in the mathematical sense regarding rules from which all other results are derived. Rather, the axioms can be regarded as the main principles of how quantum mechanics works. Since in this report we will only work with time independent case, there is no need to also include the time dependent axioms of quantum mechanics. We closely follow in this section 3.6 of [11].

In quantum mechanics, physical quantities such as position, momentum and energy are represented by operators on a Hilbert space $\mathbf{H}$.

Definition 1.35. *A Hilbert space is a vector space $\mathbf{H}$ over $\mathbb{R}$ or $\mathbb{C}$, equipped with an inner product $\langle \cdot, \cdot \rangle$ such that $\mathbf{H}$ is complete in the norm $\| \cdot \| : V \to \mathbb{R}$ defined by*

$$\|\psi\| = \sqrt{\langle \psi, \psi \rangle}.$$

We begin with the first axiom of quantum mechanics.
Axiom 1: The state of the system is represented by a unit vector ψ in an appropriate Hilbert space $\mathbf{H}$. If ψ_1 and ψ_2 are two unit vectors in $\mathbf{H}$ with $\psi_2 = c\psi_1$ for some constant $c \in \mathbb{C}$, then ψ_1 and ψ_2 represent the same physical state.

For example, the state of a quantum system, like a particle in a box, is the complex valued function

$$\psi \in L^2\left(\mathbb{R}^3\right),$$

on the Hilbert space $L^2\left(\mathbb{R}^3\right)$ For the combination of quantum systems, such as cascade of multiple particles, we need some auxiliary measure theoretical results which will be discussed in chapter 3.

Axiom 2: To each real-valued function f on the classical phase space there is associated a self-adjoint operator $\hat{f}$ on the quantum Hilbert space.

In physics literature, a function f on a classical phase space is called a classical observable. This means that it is some physical quantity which can be observed by taking a measurement of the system, such as position or momentum. The corresponding operator $\hat{f}$ is then called a quantum observable. We will now move on to the third axiom of quantum mechanics.

Axiom 3: If a quantum system is in a state described by a unit vector $\psi \in \mathbf{H}$, the probability distribution for the measurement of some observable f satisfies

$$E(f^m) = \langle \psi, (\hat{f})^m \psi \rangle .$$

In particular, the expectation value for a measurement of f is given by $\langle \psi, \hat{f}\psi \rangle$.

Hence, we have maintained the point of view that in a quantum system, what one is measuring is the classical observable f. In the quantum case, f has no longer a definite value, but only probabilities, which are encoded by the quantum observable $\hat{f}$ and the vector $\psi \in \mathbf{H}$. Since $\hat{f}$ is self adjoint and the operator is thus symmetric, we can say that $E(f)$ is real valued. However, if $\psi \in \mathbf{H}$ is an eigenvector of $\hat{f}$, the outcome of the measurement of the observable f is not totally random.

Proposition 1.36. ***(Eigenvectors)*** *If a quantum system is in a state described by a unit vector $\psi \in \mathbf{H}$ and fore some quantum observable $\hat{f}$ we have $\hat{f}\psi = \lambda\psi$ for some $\lambda \in \mathbb{R}$, then*

$$E(f^m) = \langle \psi, (\hat{f})^m \psi \rangle = \lambda^m$$

for all positive integers m. The unique probability measure consistent with this condition is the one in which f has the definite value λ, with probability one.

What is interesting in quantum mechanics is when the state of the system is a linear combination of eigenvectors for $\hat{f}$, because the measurement of f will no longer be deterministic.

Example 1.37. *Suppose $\hat{f}$ has an orthonormal basis $\{e_j\}$ of eigenvectors with distinct real eigenvalues λ_j. Suppose also that ψ is a unit vector in $\mathbf{H}$ with the expansion*

$$\psi = \sum_{j=1}^{\infty} a_j e_j .$$

Then for a measurement in the state ψ of the observable f, the observed value of f will always be one of the numbers λ_j. Furthermore, the probability of observing the value λ_j is given by

$$Prob\{f = \lambda_j\} = |a_j|^2 .$$

Last but not least, we introduce the fourth axiom of time independent quantum mechanics.
Axiom 4: Suppose a quantum system is initially is a state ψ and that a measurement of an observable f is performed. If the result of the measurement is the number $\lambda \in \mathbb{R}$, then immediately after the measurement, the system will be in a state ψ' that satisfies

$$\hat{f}\psi' = \lambda\psi' .$$

The passage from ψ to ψ' is called the collapse of the wave function. Here $\hat{f}$ is the self-adjoint operator associated with f by Axiom 2.

This axiom says the following. Assume that $\hat{f}$ has an orthonormal basis of eigenvectors $\{e_j\}$ with distinct eigenvalues λ_j. When we observe the value λ_j in the measurement of $\hat{f}$, then $\psi' = e_j$. Hence we can say that the wavefunction "collapses" to the state e_j immediately after the measurement.

1.7. Representation Theory and Quantum Mechanics

We introduced quite some abstract theory in this chapter. However, this is not just all theory without applications, on the contrary! We are first going to introduce the general setting of representation theory in quantum mechanics and in the second part we will mention some applications. In this section we follow chapter 5 of [5].

Suppose we have a Hamiltonian $\mathcal{H}$ describing our physical system and the group of symmetry operations ϕ_g, thus being a representation of a group G, such that $[\mathcal{H},\phi_g] = 0$. This finite group consisting of the ϕ_g satisfying the last condition are called the group of Schrödinger's equation. This is a group representing for example rotations, translations or permutations of the particles in the system. These set of operations do not change the Hamiltonian or its eigenvalues. For example, if ψ_n is an eigenfunction of the Hamiltonian $\mathcal{H}$ and $\mathcal{H}$ and ϕ_g commute, then

$$\phi_g\mathcal{H}\psi_n = \phi_g E_n\psi_n = \mathcal{H}(\phi_g\psi_n) = E_n(\phi_g\psi_n).$$

Hence, $\phi_g\psi_n$ is just as good an eigenfunction of $\mathcal{H}$ as ψ_n itself and corresponds to the same eigenvalue E_n. This also works for applying multiple symmetry operators ϕ_g on the eigenfunction ψ_n.

We are now going to make the matrix representation of ϕ_g for $g \in G$ more precise. We are first going to introduce the main claim: when E_n for $n \in \mathbb{N}$ is a k-fold degenerate level, we can find eigenfunctions $\psi_{n1},...,\psi_{nk}$ which are solutions to the Schrödinger's equation. Then the matrix representation of ϕ_g is given by

$$\phi_g\psi_{n\alpha} = \sum_{j=1}^{k}\psi_{nj}D^{(n)}\left(g\right)_{ja}$$

where $D^{(n)}\left(g\right)_{ja}$ is an irreducible matrix which defines the linear combination, n labels the energy index and α denotes the degeneracy index. We will show that the matrices $D^{(n)}\left(g\right)$ form an k dimensional irreducible representation of the group of Schrödinger's equation. Let $g,h \in G$ such that ϕ_g,ϕ_h commute with the Hamiltonian. Then we can write

$$\begin{aligned}\phi_{gh}\psi_{n\alpha} = \phi_g\phi_h\psi_{n\alpha} &= \phi_g\sum_j\psi_{nj}D^{(n)}\left(g\right)_{j\alpha}\\ &= \sum_{jk}\psi_{jk}D^{(n)}\left(g\right)_{kj}D^{(n)}\left(g\right)_{j\alpha}\\ &= \sum_k\psi_{nk}\left[D^{(n)}\left(g\right)D^{(n)}\left(g\right)\right]_{k\alpha}\end{aligned}$$

By definition, we know that

$$\phi_{gh}\psi_{n\alpha} = \sum_k\psi_{nk}D^{(n)}(gh)_{k\alpha}$$

by which we can conclude that

$$D^{(n)}(gh) = D^{(n)}\left(g\right)D^{(n)}(h)$$

and therefore we can conclude that the matrices $D^{(n)}\left(g\right)$ for $g \in G$ from a representation. As a last remark, we say that the dimension of the irreducible representation is equal to the degeneracy of the eigenvalue of E_n. The representation of $D^{(n)}\left(g\right)$ corresponding to $\phi\left(g\right)$ is an irreducible representation if all the ψ_{nk} correspond to a single eigenvalue of E_n. Otherwise, it would be possible to find two or more subsets of $\{\psi_{nk}\}$ for which their linear combinations would transform among them self. However, this suggests that the eigenvalue could be different for these two sets, which is not the case. Therefore, the transformation matrices for the symmetry operations form indeed an irreducible representation of the group Schrödinger's equation.

1.7.1. Applications of Group Theory to Quantum Mechanics

There are several examples of applications of group theory in physics, we will mention them but not discuss them in detail. For example, the irreducible representations of the symmetry group of Schrödinger's equation label the states and specify their degeneracies. We can also use representation theory for finding the correct linear combinations of wavefunctions to diagonalise the Hamiltonian. Lastly, we mention that group theory is also usefull in following the changes in degeneracy as the symmetry of the system is lowered. For example, when we have the Hamiltonian

$$\mathscr{H} = \mathscr{H}_0 + \mathscr{H}_1,$$

where $\mathscr{H}_0$ has high symmetry corresponding to an arbitrary group G and $\mathscr{H}_1$ corresponds to a pertubation corresponding to group of lower order G'. Usually G' is a subgroup of G [5] and the key is to find which symmetry operations of the higher symmetry group G are contained in G'. Note that the irreducible representations of G' label the states of the lower symmetry exactly. When adding a pertubation and thus lowering the symmetry of the system, we want to know what happens to the degeneracy of the states in the unperturbed system. To put it in terms of representation theory; the higher symmetry group forms reducible representations for the lower symmetry groups.

The theory introduced in this chapter is a good stepping stone for further analysing these applications which we will not pursue in this report. Was everything of this chapter written in vain? Not exactly, as mentioned several times, we will see that most of the concepts we introduce in later chapters are oftentimes generalisations of concepts introduced in this chapter.

2

Lie Groups and Lie Algebras

In the first chapter we introduced some concepts such as representations of groups with finite elements, Schur's orthogonality relations and lastly character theory. In this chapter we are going to further develop some of these notions to groups with infinite amount of elements: Lie groups. Lie groups are smooth manifolds which are endowed with a group structure such that operations as taking the inverse and group multiplication are smooth, thus being topological groups as well. The Lie algebra of a Lie group is the tangent space of a Lie group with a natural bracket operation. The reason we are going to introduce these structures is because they are of great importance in quantum mechanics.

Like we said in the last paragraph of the first chapter: symmetry such as rotational symmetry, translational symmetry or permutation of the particles in a system are of great importance in physics. Oftentimes, these groups are the previously mentioned Lie groups. In this chapter we are going to introduce several matrix Lie groups and how they relate to one another. This is of great importance for chapter 3, since they describe a physical property of particles in quantum mechanics: spin.

Connected to the notion of a Lie group is a Lie algebra, the tangent space of a Lie group at the identity. We introduce this notion because Lie algebras are linear spaces and thus easy to work with. Again, these play an important role in the analysis of spin in chapter 3. We will also discuss the connection between these two notions in depth in this chapter.

Furthermore, in this chapter the notions of complete reducibility is introduced. We will calculate the fundamental group for some matrix Lie groups in this chapter, as this will shed light in the analysis of spin in chapter 3 from a slightly different perspective. In this chapter we mostly follow the works of chapter 16 of [11] and parts of section 3.5 and 3.6 of [10].

2.1. Lie Groups

We begin by noting that we will work in the space of $M_n(\mathbb{C})$, which is the space of $n \times n$ matrices with complex entries. We identify this space with the space $\mathbb{C}^{n^2}$ equiped with the usual topology, e.g. the open balls. This can be seen by identifying each entry in $A \in M_n(\mathbb{C})$ by a coördinate in $\mathbb{C}^{n^2}$. The notion of standard topology implies that the open sets are open balls.

Definition 2.1. *A subgroup G of $GL(n;\mathbb{C})$ is closed if for each sequence A_m in G that converges to a matrix A, either A is again in G or A is not invertible. A matrix Lie group is a closed subgroup of some $GL(n,\mathbb{C})$.*

Here $\mathrm{GL}(n;\mathbb{C})$ is the general linear group consisting of all invertible matrices with complex entries. In a similar way, one can define $\mathrm{GL}(n,\mathbb{R})$ as the general linear group consisting of all invertible $n \times n$ matrices with real entries. It is easy to see that these are groups under the operation of matrix multiplication. Other groups which are of interest are the groups $\mathrm{SL}(n,\mathbb{C}), \mathrm{SO}(n,\mathbb{R})$ and

SU$(n,\mathbb{C})$, the special linear, special orthogonal and the special unitary group respectively. As a quick reminder, we give the most basic definitions of these groups.

The special linear group is the group consisting of all the complex matrices with determinant 1. The special unitary group consists of all the $n \times n$ matrices $U \in M_n(\mathbb{C})$ such that $UU^* = U^*U = I$ and the determinant of U is 1. Lastly, the special orthogonal group consists of all $n \times n$ matrices $R \in M_n(\mathbb{R})$ such that $RR^{tr} = R^{tr}R = I$ and the determinant of R is equal to 1. We assume the theory posed in [7], [24] on orthogonal, unitary matrices to be known.

It is of interest to know whether matrix Lie Groups are homomorphic to one another and therefore it is natural to introduce the following notion.

Definition 2.2. *If G_1 and G_2 are matrix Lie groups, then a Lie group homomorphism of G_1 and G_2 is a continuous group homomorphism of G_1 into G_2. A Lie group homomorphism is called a Lie group isomorphism if it is one-to-one and onto with continuous inverse. Two matrix Lie groups are called isomorphic if there exists a Lie group isomorphism between them.*

This definition is quite natural, since it a convenient way to identify matrix Lie groups that are the same up to homomorphism. We are working in a setting which is closely related to topology. Therefore, we are interersted in properties of spaces which are conserved under continuous mappings, such as connectedness and simply connectedness.

Definition 2.3. *A matrix Lie group G is connected if for all $A,B \in G$ there is a continuous path $A:[0,1] \to M_n(\mathbb{C})$ such that $A(0) = A$ and $A(1) = B$ and such that $A(t)$ lies in G for all t. A matrix Lie group G is simply connected if it is connected and every continuous loop in G can be shrunk continuously to a point in G. A matrix Lie group G is compact if it is compact as a subset of $M_n(\mathbb{C}) \cong \mathbb{C}^{n^2}$.*

In this definition connectedness is actually path connectedness. One may look in Appendix A for a more in depth look at what connectedness and simply connectedness actually mean in an algebraic topological sense. To give some intuition of simply connectedness and connectedness, we are going to proof the simply connectedness of SU(2) and the path connectedness of SO(3).

Example 2.4. *The group SU(2) is simply connected.*

Proof. We begin with claiming $\mathrm{SU}(2) = \left\{ \begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix} | \alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 = 1 \right\}$.
It is easy to see that a matrix on the r.h.s. is indeed unitary and has determinant equal to 1. For the converse, take an element A of the set on the r.h.s., then the first column $(\alpha,\beta) \in \mathbb{C}^2$ is a unit vector. The second column of A is a multiple of $(-\overline{\beta},\overline{\alpha})$ as this is orthogonal to (α,β). However, the only multiple that produces a matrix of determinant 1 is $(-\overline{\beta},\overline{\alpha})$. Therefore one can see that SU(2) is topologically the unit sphere $\mathbb{S}^3$ inside $\mathbb{C}^2 \cong \mathbb{R}^4$ and, therefore, is simply connected. □

Example 2.5. *SO(n) is path-connected, for $n \geq 1$*

Proof. Suppose we have an orthonormal basis $\{a_1,...,a_n\}$ of $\mathbb{R}^n$. We would like to know whether we can find a continuous path $\gamma:[0,1] \to \mathrm{SO}(n)$ such that $(\gamma(0)a_1,...,\gamma(0)a_n = (a_1,..,a_n)$ and $(\gamma(1)a_1,...,\gamma(1)a_n) = (e_1,...,e_n)$. Here $e_n = (0,..,1,...,0)$, with a 1 on the n-th entry.

We begin by looking for a path $\gamma_1:[0,1] \to \mathrm{SO}(n)$ which takes a_1 to e_1. Hence we look for a path such that $\gamma_1(0)a_1 = a_1$ and $\gamma_1(1)a_1 = e_1$. We begin by introducing a new basis vector u such that u is orthogonal to a_1 and u is in the span of a_1 and e_1. A method to obtain this vector u is to apply Gram-Schmidt to a_1 and e_1 when they are linearly independent. Otherwise, pick a arbitrary basis vector that is orthogonal to a_1. Complete the basis arbitrarily. Note that $e_1 \in \mathrm{span}(a_1,\mathrm{u})$. Also note that a_1 and e_1 are normalised, hence we can define a rotation which does not act as the identity on the complement of $\mathrm{span}(a_1,\mathrm{u})$,

$$e_1 = \begin{pmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & I_{n-2} \end{pmatrix} a_1.$$

Here $\phi \in [0, 2\pi]$. Note that we can set

$$\gamma(t) = \begin{pmatrix} \cos(t\phi) & \sin(t\phi) & 0 \\ -\sin(t\phi) & \cos(t\phi) & 0 \\ 0 & 0 & I_{n-2} \end{pmatrix},$$

which takes a_1 to e_1 and $\gamma_1(t)$ is in SO(n) for all $t \in [0,1]$ trivially. Now we can define a map $\gamma_2(t), t \in [0,1]$ that takes $\gamma_1(1)a_1$ to a_2 in the same manner as before. However, we would like that this rotation $\gamma_2(t)$ leaves e_1 at rest.

This will be the case as we will see. Note that e_1 and e_2 are trivially orthogonal. Also e_1 and $\gamma_1(1)a_2$ are orthogonal. This follows because γ_1 maps one orthogonal basis $\{a_1, ..., a_n\}$ to another one $(e_1, \gamma_1(1)a_1, ..., \gamma_1(1)a_n)$. Note that e_1 is then in the complement of the subspace of the rotation which takes $\gamma_1(1)a_1$ to e_2. Hence, we can apply this process $n-1$ times to obtain a composed path $\gamma = \gamma_1 * \gamma_2 * .... * \gamma_{n-1}$ such that $(\gamma(1)a_1, ..., \gamma(1)a_n) = (e_1, ..., e_n)$ and $\gamma(0) = I_n$. □

See Appendix A for more explanation of the $*$ operation.

2.2. Lie Algebras

The matrix Lie groups are quite convenient structures from a mathematical viewpoint, since it allows us to make use of the results of group theory which has already been studied quite extensively. However, as said before many times already, it often more useful to talk about the notion of Lie algebra instead of a matrix Lie group, since these are linear spaces and hence, easier to work with.

Definition 2.6. *A Lie algebra over a field* $\mathbb{F}$ *is a vector space* $\mathfrak{g}$ *over* $\mathbb{F}$*, together with a bracket map* $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ *having the following properties:*

1. $[\cdot,\cdot]$ *is bilinear*
2. $[Y, X] = -[X, Y]$ *for all* $X, Y \in \mathfrak{g}$
3. *For all* $X, Y, Z \in \mathfrak{g}$ *we have the Jacobi identity*

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

Example 2.7. *We will now introduce some standard Lie algebras for the sake of giving examples. In section 2.4 it will be clear what the connection of these Lie algebras are with their respective matrix Lie groups.*

The Lie algebra $gl(n,\mathbb{C})$ *is the complex vector space* $M_n(\mathbb{C})$ *with the Lie bracket*

$$[X, Y] = XY - YX.$$

This Lie bracket is clearly anti-symmetric and bilinear. Also note that

$$\begin{aligned}[X,[Y,Z]] &+ [Y,[Z,X]] + [Z,[X,Y]] \\ &= X(YZ - ZY) - (YZ - ZY)X + Y(ZX - XZ) \\ &- (ZX - XZ)Y + Z(XY - YX) - (XY - YX)Z = 0.\end{aligned}$$

Hence, the Jacobi identity is satisfied, as $M_n(\mathbb{C})$ *is associative. The set of unit matrices* E_{ij} *with* $i, j \in \{1, ..., n\}$ *is a basis for* $gl(n,\mathbb{C})$ *and thus* $gl(n,\mathbb{C})$ *has dimension* n^2.

Example 2.8. *Another example of a Lie algebra is the following.*

$$sl(n,\mathbb{C}) = \left\{X \in gl(n,\mathbb{C}) \mid \mathrm{Tr}(X) = 0\right\}.$$

Again, in section 2.4 it will become clear that this is the Lie algebra is actually associated with the matrix Lie group $SL(n,\mathbb{C})$. *For now we note that this is a Lie subalgebra of* $gl(n,\mathbb{C})$. *This means that* $sl(n,\mathbb{C}) \subset gl(n,\mathbb{C})$ *and that* $[X,Y] \in sl(n,\mathbb{C})$ *for all* $X,Y \in sl(n,\mathbb{C})$. *Indeed, for* $X,Y \in sl(n,\mathbb{C})$ *we have*

$$\mathrm{Tr}([X,Y]) = \mathrm{Tr}(XY - YX) = \mathrm{Tr}(XY) - \mathrm{Tr}(YX) = 0.$$

We can pick as a basis the same set of $E_{i,j}$ *with* $i,j \in \{1,...,n\}$ *as we previously used for* $gl(n,\mathbb{C})$. *We only have to realise that the* E_{ii} *for* $i \in \{1,...,n\}$ *have* $\mathrm{Tr}(E_{ii}) = 1$ *and hence, are no elements of* $sl(n,\mathbb{C})$. *There are* n *such matrices. Nevertheless, we can choose* $n-1$ *different matrices with a one on the first diagonal entry and choose a* -1 *on one another diagonal entry. Then we can see that we have* $n^2 - n + (n-1) = n^2 - 1$ *matrices in our basis of* $sl(n,\mathbb{C})$.

In this report $sl(2,\mathbb{C})$ *and* $sl(3,\mathbb{C})$ *will be used extensively. Particularly, the Lie algebra* $sl(2,\mathbb{C})$ *has a basis*

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

The Lie brackets of these basis elements are

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

With this last observation, we end our example.

However, just as in group theory, we are interested in whether Lie algebras are the same up to homomorphism. Therefore, similarly to the case of matrix Lie groups, we can define what a Lie algebra homomorphism is.

Definition 2.9. *If* $\mathfrak{g}_1$ *and* $\mathfrak{g}_2$ *are Lie algebras, a map* $\phi : \mathfrak{g}_1 \to \mathfrak{g}_2$ *is called a Lie algebra homomorphism if* ϕ *is linear and* ϕ *satisfies*

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

for all $X,Y \in \mathfrak{g}_1$. *A Lie algebra homomorphism is called a Lie algebra isomorphism if it is one-to-one and onto.*

Definition 2.10. *An ideal in* $\mathfrak{g}$ *is a subalgebra* $\mathfrak{h}$ *of* $\mathfrak{g}$ *with the stronger property that* $[X,Y] \in \mathfrak{h}$ *for all* X *in* $\mathfrak{g}$ *and* Y *in* $\mathfrak{h}$.

Note that an ideal is to a Lie algebra homomorphism what a normal group is to a group homomorphism. We can also define a direct sum in this setting.

Definition 2.11. *The direct sum of Lie algebras* $\mathfrak{g}_1$ *and* $\mathfrak{g}_2$, *denoted by* $\mathfrak{g}_1 \oplus \mathfrak{g}_2$, *is the direct sum of* $\mathfrak{g}_1$ *and* $\mathfrak{g}_2$ *as a vector space, equipped with the bracket given by*

$$[(X_1,Y_1),(X_2,Y_2)] = ([X_1,X_2],[Y_1,Y_2])$$

for all $X_1, X_2 \in \}_1$ *and* $Y_1, Y_2 \in \}_2$

The next part of this paragraph is dedicated to some notions which will play a significant role in the analysis of spin and the Baryon and Meson model which we discuss in more detail in chapters 3, 4 and 5. The first we are going to discuss are the simple Lie groups.

Definition 2.12. *A Lie algebra* $\mathfrak{g}$ *is called irreducible if the only ideals in* $\mathfrak{g}$ *are* $\{0\}$ *and* $\{\mathfrak{g}\}$. *A lie algebra* $\mathfrak{g}$ *is called simple if it is irreducible and* $\dim(\mathfrak{g}) \geq 2$.

Definition 2.13. *Let* $\mathfrak{g}$ *be a finite dimensional real or complex Lie algebra and let* $X_1,...,X_N$ *be a basis for* $\mathfrak{g}$ *(as a vector space). Then the unique constants* C_{jkl} *such that*

$$[X_j, X_k] = \sum_{l=1}^{N} c_{jkl} X_l$$

are called the structure constants of $\mathfrak{g}$ *(with respect to the chosen basis).*

2.3. Matrix Exponentials

When we are talking in the language of matrix Lie algebras, we are often considering matrix exponentials. For completeness sake, we go over some elementary notions of matrix exponentials and provide proofs for most of the theorems. We begin with the definition of a matrix exponential. For this end take $X \in M_n(\mathbb{C})$, then the matrix exponential is defined in the following way,

$$e^X = \sum_{m=0}^{\infty} \frac{X^m}{m!},$$

However, we would like to have a norm.

Definition 2.14. *For any* $X \in M_n(\mathbb{C})$*, we have*

$$\|X\| = \left(\sum_{j,k=1}^{n} |X_{ij}|^2 \right)^{1/2}$$

This is called the Hilbert Schmidt norm. Suppose $\|X\| \leq R$, then since $\|XY\| \leq \|X\|\|Y\|$ it follow that $\|X\|^n \leq R^n$. Since $\sum_{n=1}^{\infty} \frac{R^n}{n!}$ converges, the series for e^X converges on the set $\{X \in M_n(\mathbb{C}) | \|X\| \leq R\}$ by the Weierstrass M-test [9]. Similarly to the matrix exponential, one can define the logarithm as follows.

Definition 2.15. *For any* $A \in M_n(\mathbb{C})$*, define* $\log(A)$ *by*

$$\log(A) = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A-I)^m}{m}$$

whenever $\|A - I\| < 1$.

The matrix exponential has some interesting properties which we will exploit quite extensively.

Theorem 2.16. *The matrix exponential has the following properties for all* $X, Y \in M_n(\mathbb{C})$

1. $e^0 = I$
2. $e^{X^{tr}} = \left(e^X\right)^{tr}$ *and* $e^{X^*} = \left(e^X\right)^*$
3. *If A is an invertible* $n \times n$ *matrix, then* $e^{AXA^{-1}} = Ae^X A^{-1}$
4. *det*$\left(e^X\right) = e^{\mathrm{Tr}(X)}$
5. *If* $XY = YX$*, then* $e^{X+Y} = e^X e^Y$
6. e^X *is invertible and* $\left(e^X\right)^{-1} = e^{-X}$
7. *Even if* $XY \neq YX$*, we have* $e^{X+Y} = \lim_{m \to \infty} \left(e^{X/m} e^{Y/m}\right)^m$

We will not go over these proofs in full detail, as most of them are considered elementary. However, the proof of point 7 is a good exercise and since this identity is of great importance in one of the most fundamental theorems of this chapter, we are going to proof this. Actually, point 7 is a specific case of the Trotter product formula. However, before we are proceeding we have to introduce an auxiliary lemma.

Lemma 2.17. *There exists a constant c such that for all* $n \times n$ *matrices B with* $\|B\| < \frac{1}{2}$*, we have*

$$\|\log(I + B) - B\| \leq c\|B\|^2$$

Proof.

$$\log(I+B) - B = \sum_{m=2}^{\infty} (-1)^{m+1} \frac{B^m}{m} = B^2 \sum_{m=2}^{\infty} (-1)^{m+1} \frac{B^{m-2}}{m}.$$

Hence we can easily see that

$$\|\log(I+B)\| \leq \|B\|^2 \sum_{m=2}^{\infty} \frac{\left(\frac{1}{2}\right)^{m-2}}{m}.$$

□

We are now in the position to proof the last part of theorem 2.16.

Proof. For point 7 we take $X, Y \in M_n(\mathbb{C})$ arbitrary. Then we multiply the power series of $e^{X/m}$ and $e^{Y/m}$ to obtain

$$e^{X/m} e^{Y/m} = I + \frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right).$$

Note that $e^{\frac{X}{m}} e^{\frac{Y}{m}} \to I$ as $m \to \infty$. Hence, as we choose sufficiently large m we have that $e^{\frac{X}{m}} e^{\frac{Y}{m}}$ is in the domain of convergence of the logarithm as defined in definition 2.15 Hence we obtain

$$\begin{aligned}\log\left(e^{\frac{X}{m}} e^{\frac{Y}{m}}\right) &= \log\left(I + \frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right)\right) \\ &= \frac{X}{m} + \frac{Y}{m} + O\left(\left\|\frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right)\right\|^2\right) \\ &= \frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right).\end{aligned}$$

From this point all the heavy lifting has already been done, so

$$e^{\frac{X}{m}} e^{\frac{Y}{m}} = e^{\frac{X}{m} + \frac{Y}{m} + O\left(\frac{1}{m^2}\right)},$$

and hence,

$$\lim_{m\to\infty} \left(e^{\frac{X}{m}} e^{\frac{Y}{m}}\right)^m = \lim_{m\to\infty} \left(e^{X+Y+O\left(\frac{1}{m}\right)}\right) = e^{X+Y}.$$

□

Another tool which we will use quite extensively is that of a the derivative of a matrix exponential. For completeness sake, we included the proof as well.

Proposition 2.18. *Let $X \in M_n(\mathbb{C})$, then term by term differentiation yields $\frac{d}{dt} e^{tX}\Big|_{t=0} = X$.*

Proof. Note that therm by term differentiation yield the following result:

$$\frac{d}{dt} e^{tX} = \frac{d}{dt} \sum_{m=0}^{\infty} \frac{(tX)^m}{m!} = \sum_{m=1}^{\infty} \frac{X^m t^{m-1}}{(m-1)!} = \sum_{m=0}^{\infty} \frac{X^{m+1} t^m}{m!}.$$

Hence we can see that $\frac{d}{dt} e^{tX}|_{t=0} = X$ □

To conclude this section, we will introduce the notion of a one-parameter subgroup, as this will be of great use in proving the relationship between matrix Lie groups and matrix Lie algebras in the next section.

Definition 2.19. *A one-parameter subgroup of $GL(n;\mathbb{C})$ is a continuous homomorphism of $\mathbb{R}$ into $GL(n;\mathbb{C})$, that is,*

1. *A continuous map $A : \mathbb{R} \to GL(n;\mathbb{R})$,*

2. $A(0) = I$,

3. $A(s+t) = A(s)A(t)$ *for all* $s, t \in \mathbb{R}$.

Theorem 2.20. *If* $A(\cdot)$ *is a one-parameter subgroup of* $GL(n;\mathbb{C})$*, there exists a unique* $n \times n$ *complex matrix* X *such that*

$$A(t) = e^{tX}.$$

We follow the proof given in the lecture notes of [9].

Proof. We first need to show that A is differentiable. Take $s \in \mathbb{R}$ arbitrary, then

$$\int_0^h A(s+t)dt = A(s)\int_0^h A(t)dt.$$

Since $A(0) = I$ and A is a continuous homomorphism by definition, we know that the integral on the right hand side is invertible for small enough h. Furthermore,

$$\int_0^h A(s+t)dt = \int_s^{s+h} A(\mathrm{u})du,$$

Hence, we can see that

$$A(s) = \int_s^{s+h} A(\mathrm{u})du\left(\int_0^h A(t)dt\right)^{-1},$$

which is differentiable in s. Suppose that $A'(0) = X$, then

$$A'(t) = \lim_{h\to 0}\frac{A(t+h)-A(t)}{h} = \lim_{h\to 0}\frac{A(t)A(h)-A(t)A(0)}{h} = A'(0)A(t),$$

so A is a solution of the initial value problem

$$y'(t) = Xy(t), \quad y(0) = I.$$

Hence, we can easily see that the unique solution is $y(t) = \exp(tX)$. □

2.4. Lie Algebra of Matrix Lie Groups

We have defined what matrix Lie groups are and we have seen some elementary examples of matrix Lie groups such as SO(3). We are going to further our studies of Lie algebras of such matrix Lie groups. The goal of this section is to obtain a basis for so(3), the Lie algebra which will be used in the analysis of spin.

Definition 2.21. *If* $G \subset GL(n;\mathbb{C})$ *is a matrix Lie group, then the Lie algebra of* $\mathfrak{g}$ *of* G *is defined as follows:*

$$\mathfrak{g} = \{X \in M_n(\mathbb{C}) | e^{tX} \in G \text{ for all } t \in \mathbb{R}\}.$$

There are some properties which are worthwhile to note considering matrix Lie algebras,

Lemma 2.22. *For any matrix Lie group* G*, the Lie algebra* $\mathfrak{g}$ *of* G *has the following properties.*

1. *The zero matrix* 0 *belongs to* $\mathfrak{g}$

2. *For all* X *in* $\mathfrak{g}$*,* tX *belongs to* $\mathfrak{g}$ *for all real numbers* t*.*

3. *For all* X *and* Y *in* $\mathfrak{g}$*,* $X+Y$ *belongs to* $\mathfrak{g}$*.*

4. *For all* $A \in G$ *and* $X \in \mathfrak{g}$ *we have* $AXA^{-1} \in \mathfrak{g}$*.*

5. *For all* X *and* Y *in* $\mathfrak{g}$*, the commutator* $[X,Y] := XY - YX$ *belongs to* $\mathfrak{g}$*.*

Proof. Point 1 and 2 are trivial. Point 3 follows from applying the Lie product formula under the assumption that G is closed. Point 4 follows from point 3 of the properties of the matrix exponential. Point 5 is somewhat harder to proof. Note that by applying the chain rule for matrix exponentials we obtain

$$[X,Y] = \frac{d}{dt} e^{tX} Y e^{-tX}\Big|_{t=0},$$

where point 4 is used to show that the expression that is differentiated is indeed an element of $\mathfrak{g}$. Note that $\mathfrak{g}$ is a closed subspace of $M_n(\mathbb{C})$ and hence

$$[X,Y] = \lim_{h\to\infty} \frac{e^{hX} Y e^{-hX} - Y}{h}$$

belongs to $\mathfrak{g}$. □

It is not immediately clear that this is indeed a Lie algebra. So we will check all the conditions in definition 2.6. When we equip $\mathfrak{g}$ with the bracket $[X,Y] = XY - YX$ for all $X, Y \in \mathfrak{g}$, then blinearity follows trivially from point three and point 5 of lemma 2.22. When $X, Y \in \mathfrak{g}$ then $[X,Y] = -[Y,X]$, which can be checked by direct calculations. The Jacobi identity also follows by direct calculation similar to the one in example 2.7. Note that from lemma 2.22 it is clear that $\mathfrak{g}$ is a real Lie algebra.

In order to compute the basis of the Lie algebra of so(3) we are going to introduce some specific examples to further develop our intuition about basic matrix Lie algebras.

Proposition 2.23. *Let $gl(n,\mathbb{C})$, $u(n)$, $su(n)$, $so(n)$ and $sl(n;\mathbb{C})$denote the Lie algebras of the matrix Lie groups $GL(n;\mathbb{C})$, $U(n)$, $SU(n)$, $SO(n)$ and $SL(n;\mathbb{C})$. Then the following holds:*

$$\begin{aligned}
&gl(n;\mathbb{C}) = M_n(\mathbb{C})\\
&sl(n,\mathbb{C}) = \{X \in gl(n) \mid \mathrm{Tr}(X) = 0\}\\
&u(n) = \{X \in M_n(\mathbb{C}) \mid X^* = -X\}\\
&su(n) = \{X \in u(n) \mid \mathrm{Tr}(X) = 0\}\\
&so(n) = \{X \in u(n) \mid X^{tr} = -X\}.
\end{aligned}$$

Proof. Let us consider $\mathrm{su}(n,\mathbb{C})$. Assume that we have $X \in M_n(\mathbb{C})$ such that $\mathrm{Tr}(X) = 0$, then

$$\det\left(e^{tX}\right) = e^{t\,\mathrm{Tr}(X)} = e^0 = \mathbb{I}.$$

This implies that $e^{tX} \in \mathrm{SU}(n;\mathbb{C})$. Conversely, suppose that $X \in \mathrm{su}(n;\mathbb{C})$, then by inspecting the above calculation we have to obtain $e^{t\,\mathrm{Tr}(X)} = \mathbb{I}$ for all $t \in \mathbb{R}$, which is possible if and only if $\mathrm{Tr}(X) = 0$. The proof for $\mathrm{sl}(n,\mathbb{C})$ is similar and thus omitted.

We continue by inspecting $\mathrm{u}(n)$. For this end take an arbitrary $X \in M_n(\mathbb{C})$ such that $X^* = -X$. Then by the properties of the matrix exponential we have

$$\left(e^{tX}\right)^* = e^{tX^*} = e^{-tX} = \left(e^{tX}\right)^{-1}.$$

Hence, e^{tX} is a unitary matrix. Conversely, if e^{tX} is unitary for all $t \in \mathbb{R}$, then $\left(e^{tX}\right)^* = \left(e^{tX}\right)^{-1} = e^{-tX}$. When we differentiate $e^{tX^*} = e^{-tX}$ on both sides at $t = 0$, we obtain $X^* = -X$. Thus, $\mathrm{u}(n) = \{X \in M_n(\mathbb{C}) \mid X^* = -X\}$. □

Something interesting can be concluded from these proofs. That is, we can represent every element of so(3) in the following way:

$$\mathrm{so}(3) = \left\{ \begin{pmatrix} 0 & a & b\\ -a & 0 & c\\ -b & -c & 0 \end{pmatrix} \middle|\ a,b,c \in \mathbb{R} \right\}$$

We obtain the following set of generators of so(3):

$$F_1 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}; \quad F_2 := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}; \quad F_3 := \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

It is not hard to check that $[F_1, F_2] = F_3$, $[F_2, F_3] = F_1$ and $[F_3, F_1] = F_2$. With the usual relations of the bracket map the other permutations are easily obtained.

2.5. Complexification of Lie Algebras

In this section we discuss the basics of the complexification of a matric Lie algebra which will play an important role in the analysis of the root-weight system of $\mathrm{SU}(3)$.

Definition 2.24. *If V is a finite dimensional real vector space, then the complexification of V, denoted by $V_{\mathbb{C}}$, is the space of formal linear commbinations*

$$v_1 \oplus i v_2,$$

with $v_1, v_2 \in V$. This space becomes a complex vector space if we define

$$i(v_1 + i v_2) = -v_1 + i v_2$$

When we have a Lie algebra $\mathfrak{g}$ equipped with a bracket $[\cdot,\cdot]$, we there is a unique extension such that its complexification $\mathfrak{g}_{\mathbb{C}}$ is a Lie algebra as well.

Proposition 2.25. *Let $\mathfrak{g}$ be a finite-dimensional real Lie algebra and $\mathfrak{g}_{\mathbb{C}}$ its complexification as a vector space. 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. The complex Lie algebra $\mathfrak{g}_{\mathbb{C}}$ is called the complexification of the real Lie algebra $\mathfrak{g}$.*

Proof. The uniqueness is not hard to see. Take X_1, X_2, Y_1 and Y_2 as elements of $\mathfrak{g}$, then

$$[X_1 + iX_2, Y_1 + iY_2] = ([X_1, Y_1] - [X_2, Y_2]) + i([X_1, Y_2] + [X_2, Y_1]).$$

When one would define two extensions which agree on the real subset of the vector space, the brackets must agree on the complex vector space as well. To show existence, we need to check that this map is bilinear over $\mathbb{C}$, skew-symmetric and satisfies the Jacobi identity. Note that this map is an extension of the bracket over $\mathfrak{g}$, so it is at least bilinear over $\mathbb{R}$. Before we proof that this map is complex linear in both arguments, we proof that this map is skew-symmetric.

$$\begin{aligned} [X_1 + iX_2, Y_1 + iY_2] &= ([X_1, Y_1] - [X_2, Y_2]) + i([X_1, Y_2] + [X_2, Y_1]) \\ &= -([Y_1, X_1] - [Y_2, X_2]) - i([Y_2, X_1] + [Y_1, X_2]) \\ &= -[Y_1 + iY_2, X_1 + iX_2]. \end{aligned}$$

Hence, for complex bilinearity we only need to show that this map is indeed complex linear in one factor.

$$\begin{aligned} [i(X_1 + iX_2), Y_1 + iY_2] &= [-X_2 + iX_1, Y_1 + iY_2] \\ &= -([X_2, Y_1] - [X_1, Y_2]) + i([X_1, Y_1] - [X_2, Y_2]) \\ &= i(([X_1, Y_1] - [X_2, Y_2]) + i([X_2, Y_1] + [X_1, Y_2])) \\ &= i[X_1 + iX_2, Y_1 + iY_2]. \end{aligned}$$

Hence, bilinearity holds. Lastly, the Jacobi idenity has to be checked. The Jacobi idenity holds if $X, Y, Z \in \mathfrak{g}$.

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

Note that the Jacobi identity will still hold if $X \in \mathfrak{g}_{\mathbb{C}}$ and $Y, Z \in \mathfrak{g}$ as we showed that it is complex linear in X with Y, Z fixed. For the same reasoning the Jacobi identity still holds when Y, Z are also in $\mathfrak{g}_{\mathbb{C}}$. □

Proposition 2.26. *Suppose that $\mathfrak{g} \subset M_n(\mathbb{C})$ is a real Lie algebra and that for all nonzero X in $\mathfrak{g}$, the element iX is not in $\mathfrak{g}$. Then the complexification $\mathfrak{g}_{\mathbb{C}}$ of $\mathfrak{g}$ in definition 2.24 is isomorphic to the set of matrices in $M_n(\mathbb{C})$ that can be expressed in the form $X + iY$ with X and Y in $\mathfrak{g}$*

Proof. The proof is quite standard. We can consider the map $f : \mathfrak{g}_{\mathbb{C}} \to M_n(\mathbb{C})$ such that $f(X+iY) = X+iY$, sending the formal linear combination $X+iY$ to the linear combination $X+iY$ of matrices. Take $X+iY, A+iB \in \mathfrak{g}_{\mathbb{C}}$ arbitrary. Then note that

$$
\begin{aligned}
&f((X+iY)(A+iB)) = f(XA - YB + i(YA+XB)) = XA - YB + i(YA+XB) = \\
&(X+iY)(A+iB) = f(X+iY)f(A+iB),
\end{aligned}
$$

which indeed gives an homomorphism. Note that $iX \notin \mathfrak{g}$ for every nonzero $X \in \mathfrak{g}$. Hence, this map $f : \mathfrak{g}_{\mathbb{C}} \to M_n(\mathbb{C})$ is injective as well. Note that this means that there is an isomorphism between $\mathfrak{g}_{\mathbb{C}}$ and $\mathfrak{g} + i\mathfrak{g}$. □

From this proposition we can conclude that $\mathrm{su}(n)_{\mathbb{C}} \cong \mathrm{sl}(n;\mathbb{C})$. This can be seen in the following way. Note that $\mathrm{su}(n)$ is a subalgebra of $\mathrm{u}(n)$ and therefore we have for all $X \in \mathrm{su}(n)$ that $X^* = -X$. However, also note that $(iX)^* = iX$ and hence, not both X and iX can be an element of $\mathrm{su}(n)$, unless X is zero. Now note that $\mathrm{sl}(n,\mathbb{C})$ consists of all the $n \times n$ traceless matrices X. We know that every $X \in \mathrm{sl}(n,\mathbb{C})$ can be expressed as $X = X_1 + iX_2$ where $X_1 = \frac{X-X^*}{2}$ and $X_2 = \frac{X+X^*}{2}$, which are both in $\mathrm{su}(n)$. Hence, $\mathrm{su}(n)_{\mathbb{C}} \cong \mathrm{sl}(n;\mathbb{C})$. To recap, from the first proposition 2.25 we could conclude that $\mathrm{su}(n)_{\mathbb{C}}$ is a complex vector space and from the second proposition 2.26 it is clear that this vector space is isomorphic to $\mathrm{sl}(n,\mathbb{C})$.

2.6. Connection between Lie Groups and Lie Algebras

In the previous sections we have seen several examples of Lie groups and Lie algebras $\mathrm{so}(3), \mathrm{su}(2), \mathrm{SO}(3)$ and $\mathrm{SU}(2)$. However, we never explicitly stated the connection between them. In this section, we will proof that $\mathrm{SU}(2)$ is a double cover of $\mathrm{SO}(3)$. This will play a great role in the analysis of spin. We begin our study with stating the relationship between matrix Lie groups and the corresponding Lie algebras.

Theorem 2.27. *Suppose G_1 and G_2 are matrix Lie groups with Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$, respectively, and suppose that $\Phi : G_1 \to G_2$ is a Lie group homomorphism, Then there exists a unique linear map $\phi : \mathfrak{g}_1 \to \mathfrak{g}_2$ such that $\Phi(e^{tX}) = e^{t\phi(X)}$ for all $t \in \mathbb{R}$ and $X \in \mathfrak{g}$. This linear map has the following additional properties:*

1. *$\phi([X,Y]) = [\phi(X), \phi(Y)]$ for all $X, Y \in \mathfrak{g}$*
2. *$\phi(AXA^{-1}) = \Phi(A)\phi(X)\Phi(A)^{-1}$ for all $A \in G$ and $X \in \mathfrak{g}$*
3. *$\phi(X)$ may be computed as*

$$
\phi(X) = \frac{d}{dt}\Phi\left(e^{tX}\right)\bigg|_{t=0}
$$

Proof. Note that Φ is a continuous group homomorphism from G_1 to G_2. Note as well that $\Phi(e^{tX})$ is a one parameter subgroup in G_2. Hence, by theorem 2.20 we have that there is a unique matrix Z such that

$$
\Phi(e^{tX}) = e^{tZ}
$$

for all $t \in \mathbb{R}$. We define $\phi(X) = Z$ and check that it has the desired properties.

Let $t = 1$, then it follows that $\Phi(e^X) = e^{\phi(x)}$ for all $X \in \mathfrak{g}$. For the following point, if $\Phi(e^{tX}) = e^{tZ}$ for all t, then it naturally follows that $\Phi(e^{ts(X)}) = e^{ts(Z)}$, therefore $\phi(sX) = s\phi(X)$. Note that Φ is continuous, we can therefore use the Lie product formula in the following way:

$$\begin{aligned} e^{t\phi(X+Y)} &= \Phi\Big(\lim_{m\to\infty}\big(e^{tX/m}e^{tY/m}\big)^m\Big) = \lim_{m\to\infty}(\Phi(e^{tX/m}\Phi(e^{tY/m}))^m \\ &= \lim_{m\to\infty}\big(e^{t\phi(X)}e^{t\phi(Y)/m}\big)^m = e^{t(\phi(X)+\phi(Y))} \end{aligned}$$

We can now use proposition 2.18 to both sides to obtain the equality: $\phi(X+Y) = \phi(X) + \phi(Y)$.

We have constructed a real-linear map which satisfies $\Phi(e^X) = e^{t\phi(X)}$. We still have to proof that this map is indeed unique. For this end take another real linear map ϕ' with all the previous properties. Then one would note that $e^{t\phi(X)} = \phi(e^{tX}) = e^{t\phi'(X)}$ for all $t \in \mathbb{R}$. Again, we can apply proposition 2.18 to both sides to obtain the equality $\phi(X) = \phi'(X)$.

We still have to proof the properties 1,2,3 and these all follow more or less from proposition 2.18 For point 1, take an arbitrary $A \in G$. We then have

$$\begin{aligned} e^{t\phi(AXA^{-1})} &= e^{t\phi(tAXA^{-1})} = \Phi(e^{tAXA^{-1}}) \\ &= \Phi(A)\Phi(e^{tX})\Phi(A)^{-1} = \Phi(A)e^{t\phi(X)}\Phi(A)^{-1}. \end{aligned}$$

One can differentiate both sides at $t = 0$ to obtain the equality in point 1.

For point 2 we have to use the following identity which we already have used in the proof of lemma 2.22.

$$\phi([X,Y]) = \phi\left(\frac{d}{dt}e^{tX}Ye^{-tX}\bigg|_{t=0}\right) = \frac{d}{dt}\phi\left(e^{tX}Ye^{-tX}\right)\bigg|_{t=0}.$$

Note that we have used that the derivative commutes with the linear transformation.

$$\begin{aligned} \phi([X,Y]) &= \frac{d}{dt}\phi\big(e^{tX}Ye^{-tX}\big)\bigg|_{t=0} = \frac{d}{dt}\Phi(e^{tX})\phi(Y)\Phi(e^{-tX})\bigg|_{t=0} \\ &= \frac{d}{dt}e^{t\phi(X)}\phi(Y)e^{-t\phi(X)}\bigg|_{t=0} = [\phi(X),\phi(Y)]. \end{aligned}$$

This established point 2. To finish this proof, note that for point 3 we can write $\phi(e^{tX}) = e^{\phi(tX)} = e^{t\phi(X)}$. Again using proposition 2.18 we can compute $\phi(X)$ as is shown in point 3. □

We have now constructed a relationship between Lie groups and Lie algebras. This is of great use, since we may study for example matrix Lie groups and use these results to deduce the relationship between Lie algebras and vice-versa. For example, when we have two matrix Lie groups which are isomorphic, then we are able to deduce whether the related matrix Lie algebras are isomorphic or not.

Theorem 2.28. *Suppose that G_1 and G_2 are matrix Lie groups with Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$, respectively. If G_1 is isomorphic to G_2, then $\mathfrak{g}_1$ is isomorphic to $\mathfrak{g}_2$.*

Proof. Suppose $\Phi : G_1 \to G_2$ is a Lie group isomorphism from G_1 to G_2 with corresponding Lie algebra homomorphism $\phi : \mathfrak{g}_1 \to \mathfrak{g}_2$. Then note that $\Phi^{-1} : G_2 \to G_2$ is an Lie group isomorphism as well with corresponding Lie algebra homomorphism $\psi : \mathfrak{g}_2 \to \mathfrak{g}_1$. Let $X \in \mathfrak{g}_1$ arbitrary, then $\Phi^{-1}\big(\Phi\big(e^{tX}\big)\big) = e^{t\psi(\phi(X))}$ which results in $e^{tX} = e^{t\psi(\phi(X))}$ for all $t \in \mathbb{R}$. Hence $X = \psi\big(\phi(X)\big)$. A similar argument can be used to show that for any $X \in \mathfrak{g}_2$ we obtain $X = \phi\big(\psi(X)\big)$. Hence, this shows that $\phi \circ \psi = \mathrm{Id}_{\mathfrak{g}_2}$ and $\psi \circ \phi = \mathrm{Id}_{\mathfrak{g}_1}$. Hence $\phi : \mathfrak{g}_1 \to \mathfrak{g}_2$ is a bijection and hence Lie algebra isomorphism. □

This is a result which one would expect to be true. However, the converse may not be always true. Nevertheless, there are some cases in which it is. The following theorem is an example of such a case.

Theorem 2.29. *Suppose that G_1 and G_2 are matrix Lie groups with Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$, respectively, and suppose that $\phi : \mathfrak{g}_1 \to \mathfrak{g}_2$ is a Lie algebra homomorphism. If G_1 is connected and simply connected, then there exists a unique Lie group homomorphism $\Phi : G_1 \to G_2$ such that Φ and ϕ are related as in theorem 2.27.*

The proof can be found in [10]. However, when the requirements of this theorem are not satisfied, this theorem might not always hold. We will encounter this with for example SU(2) and SO(3). In fact, SU(2) is a double cover of SO(3), meaning there is a two to one continuous mapping from SU(2) onto SO(3).

Definition 2.30. *Suppose G is a connected matrix Lie group with Lie algebra $\mathfrak{g}$. A universal cover of G is an ordered pair $\left(\overline{G}, \Phi\right)$ consisting of a simply connected matrix Lie group $\overline{G}$ and a Lie group homomorphism $\Phi : \overline{G} \to G$ such that the associated Lie algebra homomorphism $\phi : \overline{\mathfrak{g}} \to \mathfrak{g}$ is an isomorphism of the Lie algebra $\overline{\mathfrak{g}}$ of $\overline{G}$ with $\mathfrak{g}$. The map Φ is called the covering map for $\overline{G}$.*

When we want SU(2) to be a double cover of SO(3), we first need to proof that so(3) and su(2) are isomorphic to one another.

Proposition 2.31. *The Lie algebra su(2) and so(3) are isomorphic, but the Lie groups SU(2) and SO(3) are not isomorphic.*

Proof. We know from proposition 2.23 that $\mathrm{su}(2) = \left\{ \begin{pmatrix} ia & b+ic \\ -b+ic & -ia \end{pmatrix} \middle| \, a,b,c \in \mathbb{R} \right\}$. Then it trivially follows that one can choose a basis:

$$E_1 = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}; \quad E_2 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}; \quad E_3 = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix};$$

The commutator relations $[E_1, E_2] = E_3$ satisfy the same cyclic permutation as the $F_j, j = 1,2,3$ for the basis of so(3), as can be checked from direct calculations. This is quite interesting, as one can construct an isomoprhism $\phi : \mathrm{su}(2) \to \mathrm{so}(3)$ mapping the basis of su(2) directly to the basis of so(3).

There is, however, no isomorphism between SU(2) and SO(3), as it can be easily seen that SU(2) has non-trivial center containing at least $\{I, -I\}$, whereas the center of SO(3) is trivial.

To conclude this proof, we will show that the center of SO(3) is indeed trivial. We will do this by contradiction. Assume that $R_1 \neq I$ is in the center of SO(3). We know that every $A \in \mathrm{SO}(3)$ has an eigenvalue 1, as one can see in Appendix A. Let $A \neq I \in \mathrm{SO}(3)$ arbitrary. We know that there exists an $v \in \mathbb{C}^3$ such that $R_1 v = v$. Since R_1 is in the center we may assume $AR_1 v = R_1 Av$, which results in $(R_1 - I)Av = 0$. Since R_1 and A commute and v is an eigenvector of R_1, we know that Av is again an eigenvector of R_1 and hence non-zero. Therefore, we may safely conclude that $R_1 = I$, which is a contradiction. □

Before proving that $(\mathrm{SU}(2), \Phi)$ is indeed a cover of SO(3), three auxiliary theorems and lemmas will aid us.

Theorem 2.32. *Let G be a matrix Lie group with Lie algebra $\mathfrak{g}$. Then there exists a neighborhood U of 0 in $M_n(\mathbb{C})$ and a neighborhood V of I in $M_n(\mathbb{C})$ such that the matrix exponential maps U diffeomorphically onto V and such that for all $X \in U$, we have that X belongs to $\mathfrak{g}$ if and only if e^X belongs to G.*

The proof of this theorem can be found in [10], but we will not dive into the proof, since we need a lot of auxiliary tools to proof this. This theorem essentially tells us that our Lie algebra $\mathfrak{g}$ of a matrix Lie group G is large enough to tell us what happens in a neighbourhood of the identity in G. We will continue with two fun lemmas.

Lemma 2.33. *A discrete normal subgroup of a connected group is central.*

Proof. Let G be a connected group and N a discrete normal subgroup of G. Let $n \in N$ be arbitrary. Define the mapping $f: G \to N$ by $g \mapsto gng^{-1}$. We know that this function is continuous and therefore the image of f is connected as connectedness is a topological property. Note that for $g = e$ the identity element $f(e) = n$, therefore n is contained in the image of f. Also note that N is a discrete normal subgroup and hence the only connected sets are singletons, so $f(g) = n$ for all $g \in G$. Therefore $gng^{-1} = n$ and hence $gn = ng$ for all $g \in G$. □

Lemma 2.34. *The center of $SU(2)$ is no larger than $\{I, -I\}$*

Proof. One can take $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{SU}(2)$ and $B = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \in \mathrm{SU}(2)$. We can conclude that for an element in $\begin{pmatrix} \alpha & -\overline{\beta} \\ \beta & \overline{\alpha} \end{pmatrix} \in \mathrm{SU}(2)$ to commute with A, B, then α should be real and β should be purely real and imaginary. This leads to $\alpha = \pm 1$ and $\beta = 0$. Hence the only elements in the center of $\mathrm{SU}(2)$ are $\pm I$. □

Finally, we are now in the position to proof the main theorem of this section.

Theorem 2.35. *Let $\Phi : SU(2) \to SO(3)$ be the unique Lie group homomorphism for which the associated Lie algebra homomorphism ϕ satisfies $\phi(E_j) = F_j, j = 1,2,3$. Then $\ker(\Phi) = \{I, -I\}$ and $(SU(2), \Phi)$ is a universal cover of $SO(3)$.*

Proof. As we already know, E_1 is diagonal. From trivial calculations, it follows that $e^{2\pi E_1} = -I$ in $\mathrm{SU}(2)$. We also have

$$e^{aF_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(a) & -\sin(a) \\ 0 & \sin(a) & \cos(a) \end{pmatrix};$$

for all $a \in \mathbb{R}$. Also

$$\Phi(-I) = \Phi\left(e^{2\pi E_1}\right) = e^{2\pi\phi(E_1)} = e^{2\pi F_1} = I.$$

As one can see, $-I$ is an element of the kernel of Φ.

From theorem 2.32 we can conclude that as ϕ is injective, Φ is injective near a neighbourhood of I. Take $A, B \in \mathrm{SU}(2)$ and $A \neq B$ close enough to I, then by theorem 2.32 we have $A = e^X$ and $B = e^Y$ for $X, Y \in \mathrm{su}(2)$ with $X \neq Y$.

We know that ϕ is injective, hence $\phi(X), \phi(Y)$ are distinct elements in $\mathrm{so}(3)$. Theorem 2.32 can be applied once again to note that $\Phi(A) = e^{\phi(X)}$ and $\Phi(B) = e^{\phi(Y)}$ are distinct as well. Therefore, we can see that the $\ker(\Phi)$ is a discrete normal subgroup. We now from lemma 2.33 that the discrete normal subgroup of a connected group is central.

From lemma 2.34 we know that the center of $\mathrm{SU}(2)$ is no larger than $\{I, -I\}$. From this it follows that the kernel of Φ cannot be larger than $\{\pm I\}$. Lastly, we have to proof that Φ maps onto $\mathrm{SO}(3)$. This can be done by noting that each element of $\mathrm{SO}(3)$ can be expressed as $R = e^{\phi(X)}$ with $X \in \mathrm{so}(3)$.

However, this claim is far from trivial. Note the following:

$$e^{bF_2} = \begin{pmatrix} \cos(b) & 0 & -\sin(b) \\ 0 & 1 & 0 \\ \sin(b) & 0 & \cos(b) \end{pmatrix},$$

and

$$e^{cF_3} = \begin{pmatrix} \cos(c) & -\sin(c) & 0 \\ \sin(c) & \cos(c) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Hence, we see that e^{aF_1}, e^{bF_2} and e^{cF_3} are rotations around the x, y and z axis respectively. We denote these subgroups as R_x, R_y and R_z respectively. Here $a, b, c \in \mathbb{R}$. Take an arbitrary $R \in \mathrm{SO}(3)$, then note that $R\mathbf{x}$, with $\mathbf{x} \in \mathbb{R}^3$ parallel with the z-axis, can be rotated in the (x,z) plane by an $Q \in R_z$. Then note that $QR\mathbf{x}$ can be rotated to $\mathbf{z}$ by an $W \in R_y$, hence $WQR\mathbf{z} = \mathbf{z}$. Hence, this shows that $T = WQR \in R_z$ and that $R = Q^{-1}W^{-1}T$ with $Q^{-1}, T \in R_z$ and $W^{-1} \in R_y$. Hence, we can see that $R = e^{\phi F_3}e^{\theta F_2}e^{\psi F_3}$ with $(\phi,\theta,\psi) \in \mathbb{R}^3$. Hence, we see that Φ is a surjective map. □

2.7. Finite Dimensional Representations of Lie Groups and Lie Algebras

Now we have introduced all the necessary tools to introduce the notion of a representation. In this section, we will give some basic definitions concerning finite dimensional representations in general. We will also provide some tools to work with representations and identify them, such as Schur's Lemma. The representations we will consider are those involving matrix Lie groups, which has a topological structure as discussed before. Untill section 2.7.2 we mostly follow section 16.7 of [11].

Definition 2.36. *Let $G \subset GL(n;\mathbb{C})$ be a matrix Lie group. A finite dimensional representation of G is a continuous homomorphism of G into $GL(V)$, the group of all invertible linear transformations of a finite dimensional vector space V.*

As we have noted, the matrix Lie groups inherited the topological structure of $M_n(\mathbb{C})$. Therefore, it is only natural to require for the homomorphism to be continuous. The vector spaces we consider are over the field $\mathbb{C}$. We further our studies with the notion of Lie algebras.

Definition 2.37. *A finite-dimensional representation of a Lie algebra $\mathfrak{g}$ is a Lie algebra homomorphism of $\mathfrak{g}$ into $gl(V)$, the space of all linear transformations of V. Here $gl(V)$ is considered as a Lie algebra with bracket given by $[X,Y] = XY - YX$.*

As we have introduced the important definitions of Lie group and Lie algebra homomorphisms, we can note the following. Given a representation $\Pi : G \to \mathrm{GL}(V)$ of a matrix Lie group G, we can pick a basis for V and note that we can identify $\mathrm{GL}(V)$ with $\mathrm{GL}(n;\mathbb{C})$ and $\mathrm{gl}(V)$ with $\mathrm{gl}(n;\mathbb{C})$. Using theorem 2.27 we obtain a representation $\pi : \mathfrak{g} \to \mathrm{gl}(V)$ such that

$$\Pi\left(e^X\right) = e^{\pi(X)}$$

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

As an intermezzo we are going to give a little intro in how representations of $\mathrm{SO}(n)$ can be used in quantum mechanics. An important example of a representation in quantum mechanics arises from the time-independent Schrödinger equation in $\mathbb{R}^n$, namely $\mathscr{H}\psi = E\psi$, for a fixed $E \in \mathbb{R}$ and $\mathscr{H}$ is the Hamiltonian which describes our physical system. This is what we already have seen at the end of the first chapter.

Suppose that $\mathscr{H}$ is invariant under rotations. This means that $\mathscr{H}$ commutes with the relevant representations of the rotation group and also with the associated Lie algebra operators. This implies that the eigenspaces for $\mathscr{H}$ are invariant under rotations. Even when a solution ψ of the Schrödinger equation is not a radial function, thus rotationally invariant, the action of $\mathrm{SO}(n)$ on this solution will give another solution of the Schrödinger equation in the case that $\mathscr{H}$ is rotationally invariant.

Even when the quantum Hilbert space describing our physical system is infinite dimensional, the solution spaces are typically finite dimensional [11]. Therefore, if we understand what all finite dimensional representations of $\mathrm{SO}(n)$ look like, we already made some great steps in understanding what the solutions to $\mathscr{H}\psi = E\psi$ look like in the rotational invariant case. This is one of the examples of the power of representation theory in quantum mechanics. This ends our little intermezzo.

Before proving the following theorem, we need an additional proposition.

Proposition 2.38. *If a matrix Lie group G is connected, then for all $A \in G$ there exists a finite sequence $X_1, X_2, ..., X_N$ of elements of $\mathfrak{g}$ such that*

$$A = e^{X_1} e^{X_2} ... e^{X_N}.$$

The proof can be found in [11].

Just as with groups, it is oftentimes useful to know which representations are isomorphic to one another.

Theorem 2.39. *Suppose G is a connected matrix Lie group with Lie algebra $\mathfrak{g}$. Suppose that $\Pi : G \to GL(V)$ is a finite-dimensional representation of G and $\pi : \mathfrak{g} \to gl(V)$ is the associated Lie algebra representation. Then a subspace W of V is invariant under the action of G if and only if it is invariant under the action of $\mathfrak{g}$. In particular, Π is irreducible if and only if π is irreducible. Furthermore, two representations of G are isomorphic if and only if the associated Lie algebra representations are isomorphic.*

Proof. Suppose that $W \subset V$ is an invariant subspace of W under the action of $\pi(X)$ for all $X \in \mathfrak{g}$. Note that W is invariant under $\pi(X)^m$ for all $m \in \mathbb{N}$. We assumed that V is a finite dimensional subspace, and therefore any subspace of it is a closed subset. Hence, W is invariant under

$$\Pi(e^X) = e^{\pi(X)} = \sum_{m=0}^{\infty} \frac{\pi(X)^m}{m!}.$$

Since G is connected, we know that every element of G is a product of exponentials of elements of $\mathfrak{g}$, as the previous proposition 2.38 dictates. Therefore, we see that W is invariant under $\Pi(A)$ for all $A \in G$.

Now we are going to proof the converse implication. Suppose that W is invariant under $\Pi(A)$ for all $A \in G$. Then, as W is closed, it is invariant under

$$\pi(X) = \lim_{h \to 0} \frac{\Pi\left(e^{hX}\right) - I}{h}$$

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

For the last part of the proof, suppose that Π_1 and Π_2 are two representation of G, acting on two vector spaces V_1 and V_2. Also suppose that $\Phi : V_1 \to V_2$ is an invertible linear map. Then by following similar steps as above we can see that $\Phi\Pi(A) = \Pi(A)\Phi$ for all $A \in G$ if and only if $\Phi\pi(X) = \pi(X)\Phi$ for all $X \in \mathfrak{g}$. Hence, Φ is an isomorphism of group representations if and only if it is an isomorphism of Lie algebra representations. □

As we have foreshadowed in the first chapter, we are going to proof Schur's lemma.

Theorem 2.40. *If V_1 and V_2 are two irreducible representations of a group or Lie algebra, then the following hold.*

1. *If $\Phi : V_1 \to V_2$ is an intertwining map, then either $\Phi = 0$ or Φ is an isomorphism.*
2. *If $\Phi : V_1 \to V_2$ and $\Psi : V_1 \to V_2$ are nonzero intertwining maps, then there exists a nonzero constant $c \in \mathbb{C}$ such that $\Phi = c\Psi$. In particular, if Φ is an intertwining map of V_1 to itself then $\Phi = cI$.*

Proof. Ker(Φ) is an invariant subspace of V_1. Note that V_1 is assumed to be irreducible, hence $\ker(\Phi) = \{0\}$, thus injective, or $\ker(\Phi) = V_1$, thus $\Phi = 0$. Similarly, note that Im(Φ) is an invariant subspace of V_2. Therefore, we note that either $\text{Im}(\Phi) = \{0\}$ or $\text{Im}(\Phi) = V_2$. In the latter case we know that Φ has to be injective as well, hence bijective. Point 1 has now been proven.

For point 2 note the following. Since Φ and Ψ are non zero we know that they must be isomorphism by point 1. It is enough to proof that $\Gamma = \Phi^{-1}\Psi : V_1 \to V_1$ is a multiple of the identity. To prove this we first note the following: $\mathbb{C}$ is an algebraïcally closed field. Therefore, Γ must have at least one eigenvalue $\lambda \in \mathbb{C}$. We denote W the λ eigenspace of Γ. Then we know this subspace is invariant under the action of the Lie group or Lie algebra.

The last part follows quite easily. Take an arbitrary $w \in W$, then note that $\Gamma w = \lambda w$. We can then conclude that $\Gamma(\Pi(A)w) = \lambda\Pi(A)w = \Pi(A)\Gamma w$. Since w is an eigenvector of Γ, the invariant subspace W is nonzero and hence $V_1 = W$. This yields that $\Gamma = \lambda I$. □

The first part of Schur's Lemma holds for representations over an arbitrary field while the second part only holds for representations over algebraically closed fields.[11]

2.7.1. Unitary Representations

In quantum mechanics, we are not interested in all vector spaces, but mainly in Hilbert spaces. The expectation values in quantum mechanics are defined in terms of inner products. This construct is used for making predictions concerning observables in quantum mechanics, as is specified in chapter 1. Hence, it is only natural to study the actions of a group that preserves the inner product and the linear structure. As noted before, in this section the finite-dimensional case will be highlighted.

Definition 2.41. *Suppose V is a finite dimensional Hilbert space over $\mathbb{C}$. $U(V)$ is the group of invertible linear transformations of V that preserve the inner product. A (finite dimensional) unitary representation of a matrix Lie group G is a continuous homomorphism of $\Pi : G \to U(V)$ for some finite-dimensional Hilbert space V.*

As we have introduced the notion of unitary representation, one might wonder when an arbitrary representation of a Lie group is unitary with respect to an inner product defined on a Hilbert space.

Theorem 2.42. *Let $\Pi : G \to GL(V)$ be a finite dimensional representation of a connected matrix Lie group G and let π be the associated representation of the Lie algebra $\mathfrak{g}$ of G. Let $\langle\cdot,\cdot\rangle$ be an inner product on V. Then Π is unitary with respect to $\langle\cdot,\cdot\rangle$ if and only if $\pi(X)$ is skew self-adjoint with respect to $\langle\cdot,\cdot\rangle$ for all $X \in \mathfrak{g}$, that is, if and only if*

$$\pi(X)^* = -\pi(X)$$

for all $X \in \mathfrak{g}$

Proof. The proof in both directions is quite easy. First assume that $\Pi(A)$ is unitary for all $A \in G$. Then for all $X \in \mathfrak{g}$ and $t \in \mathbb{R}$ we have

$$\Pi\left(e^{tX}\right)^* = \Pi\left(e^{tX}\right)^{-1} = \Pi\left(e^{-tX}\right) = e^{-t\pi(X)}.$$

However, we also have

$$\Pi\left(e^{tX}\right)^* = \left(e^{t\pi(X)}\right)^* = e^{t\pi(X)^*}.$$

Combining these two observations we arrive at

$$e^{t\pi(X)^*} = e^{-t\pi(X)}$$

for all $t \in \mathbb{R}$. We can apply lemma 2.18 to both sides to obtain $\pi(X)^* = -\pi(X)$.

Now assume that $\pi(X)^* = -\pi(X)$ for all $X \in \mathfrak{g}$, then

$$\Pi\left(e^X\right)^* = e^{\pi(X)^*} = e^{-\pi(X)} = \Pi\left(e^{-X}\right) = \Pi\left(e^X\right)^{-1}.$$

Hence, $\Pi(e^X)$ is unitary. Since G is connected we can use theorem 2.38. This theorem tells us that all $A \in G$ are expressible as a product of exponentials, thus $\Pi(A)$ is unitary. □

2.7.2. Projective Unitary Representation

Note that by axiom 1 in section 1.6 we can see that two unit vectors ψ_1, ψ_2 in a Hilbert space **H** such that $\psi_1 = c\psi_2$ with $c \in \mathbb{C}$ describe the same physical state. Therefore an operator of the form $e^{i\theta}I$ with $\theta \in \mathbb{R}$ acts as the identity operator on the level of physical states. Suppose that V is a finite dimensional vector space over $\mathbb{C}$, then it is only natural to consider representations not only to $\mathrm{U}(V)$ but also to $\mathrm{U}(V)/\{e^{i\theta}\}$ for $\theta \in \mathbb{R}$.

Definition 2.43. *Suppose that V is a finite-dimensional Hilbert space over $\mathbb{C}$. Then the projective unitary group over V, denoted $PU(V)$, is the quotient group*

$$PU(V) = U(V)/\{e^{i\theta}\},$$

where $\{e^{i\theta}\}$ denotes the group of matrices of the form $e^{i\theta}I$, $\theta \in R$.

Note that $\{e^{i\theta}\}$ is easily seen to be a normal subgroup of a matrix Lie group G. Note that $\mathrm{U}(V)$ is a matrix Lie group by picking a orthonormal basis for V. However, note that at first sight it is not clear that the quotient of a matrix Lie group with a normal subgroup is a matrix Lie group again.

Proposition 2.44. *If V is a finite-dimensional Hilbert space over $\mathbb{C}$, then $PU(V)$ is isomorphic to a matrix Lie group. Let $Q : U(V) \to PU(V)$ be the quotient homomorphism and let $q : u(V) \to pu(V)$ be the associated Lie algebra homomorphism. Then q maps $u(V)$ onto $pu(V)$ and the kernel of q is the space of matrices of the form iaI with $a \in \mathbb{R}$. Thus $pu(V)$ is isomorphic to $u(V)/\{ia\}$.*

Proof. If $\dim(V) = N$, then $\mathrm{gl}(V)$ has dimension N^2. We can set for $U \in \mathrm{U}(V)$

$$C_U : \mathrm{gl}(V) \to \mathrm{gl}(V)$$

by

$$C_U(X) = UXU^{-1}.$$

Note that $(C_U)^{-1} = C_{U^{-1}}$ and that $C_{UV} = C_U C_V$. Hence, the map $U \to C_U$ is a homomorphism from $\mathrm{U}(V)$ to $\mathrm{GL}(\mathrm{gl}(V))$. The homomorphism is continuous. If U is a multiple of the idenity, then C_U is the identity operator on $\mathrm{gl}(V)$ and conversely, if C_U is the identity operator, then $UX = XU$ for all $X \in \mathrm{gl}(V)$. Hence, U is a multiple of the the identity. Hence, the kernel of C consists of the scalar multiples of the identity that are in $\mathrm{U}(V)$. Hence $\ker(C) = \{e^{i\theta}\}$.

Note that C is thus a homomorphism from $\mathrm{U}(V)$ to $\mathrm{GL}(\mathrm{gl}(V)) \cong \mathrm{GL}(N^2, \mathbb{C})$ with kernel $\{e^{i\theta}\}$. It follows that the image of $\mathrm{U}(V)$ under the homomorphism is isomorphic to $\mathrm{U}(V)/\{e^{i\theta}\}$. Since, $\mathrm{U}(V)$ is compact, we know that the image of $\mathrm{U}(V)$ under C is compact as well as C is a continuous mapping. Therefore, a matrix Lie group is isomorphic to $\mathrm{PU}(V)$.

Let c be the Lie algebra homomorphism associated with C. We can calculate that

$$\begin{aligned} c_X(Y) &= \left.\frac{d}{dt} e^{tX} Y e^{-tX}\right|_{t=0} \\ &= XY - YX \\ &= [X, Y]. \end{aligned}$$

Note that $c_X = 0$ if and only if X is a scalar multiple of the identity. Therefore, the kernel of c is simply $\{iaI\} \subset \mathrm{u}(V)$.

Note that the image of $\mathrm{U}(V)$ under C is isomorphic to $\mathrm{U}(V)/\{e^{i\theta}\}$. Therefore C maps $\mathrm{U}(V)$ onto $\mathrm{PU}(V)$ [10]. Note that c then maps $\mathrm{u}(V)$ onto $\mathrm{pu}(V)$. Hence $\mathrm{pu}(V) \cong \mathrm{u}(V)/\{ia\}$. □

With the observation that $\mathrm{PU}(V)$ is indeed isomorphic to a matrix Lie group, we can define what a projective unitary representation is.

Definition 2.45. *A finite-dimensional projective unitary representation of a matrix Lie group G is a continuous homomorphism Π of G into $PU(V)$, where V is a finite dimensional Hilbert space over $\mathbb{C}$. A subspace W of V is said to be invariant under Π if for each $A \in G$, W is invariant under U for every $U \in U(V)$ such that $[U] = \Pi(A)$. Here, $[U]$ denotes all elements $X \in U(V)$ such that $X = e^{i\theta}U$ for some $\theta \in \mathbb{R}$. A projective unitary representation (Π, V) is irreducible if the only invariant subspaces are $\{0\}$ and V.*

The next proposition tells us that every finite-dimensional projective representation can be decomposed into a representations $\sigma : \mathfrak{g} \to \mathrm{u}(V)$ and $q : \mathrm{u}(V) \to \mathrm{pu}(V)$ at the Lie algebra level.

Proposition 2.46. *Let $\Pi : G \to PU(V)$ be a finite-dimensional projective unitary representation of a matrix Lie group G, and let $\pi : \mathfrak{g} \to pu(V)$ be the associated Lie algebra homomorphism. Then there exists a Lie algebra homomorphism $\sigma : \mathfrak{g} \to u(V)$ such that $\pi(X) = q(\sigma(X))$ for all $X \in \mathfrak{g}$. It is possible to choose σ so that $Tr(\sigma(X)) = 0$ for all $X \in \mathfrak{g}$, and σ is unique if we require this condition.*

Proof. Note that $\mathrm{pu}(V) \cong \mathrm{u}(V)/\{iaI\}$. Hence, for each $X \in \mathfrak{g}$, $\pi(X)$ is a whole family of operators that differ by adding iaI. Note that when $Y \in \mathrm{u}(n)$ is any representative of $\pi(X)$, then since $Y^* = -Y$, the trace will be pure imaginary. Hence, there is a unique constant $c = \frac{-\mathrm{Tr}(X)}{\dim(V)}$ such that the trace of $Y + cI$ is zero. Then set $\sigma(X) = Y + cI$. Since $\pi : \mathrm{u}(V) \to \mathrm{pu}(V)$ is a Lie algebra homomorphism, we note

$$\sigma([X,Y]) = [\sigma(X), \sigma(Y)] + iaI,$$

for some $a \in \mathbb{R}$. Since $\mathrm{Tr}(\sigma([X,Y])) = 0$ by construction, we see that $a = 0$. Hence, we see that such σ exists and it is unique if we require that $\sigma(X)$ has trace zero. □

The final theorem of this chapter is the following. It says that in the finite dimensional case, there is a one-to-one correspondence between the irreducible projective unitary representations of G and irreducible, determinant-one unitary representations of its universal cover $\tilde{G}$.

Theorem 2.47. *Suppose G is a matrix Lie group and $\tilde{G}$ is a universal cover of G, with covering map Φ. Then the following hold.*

1. *Let $\Pi : G \to PU(V)$ be a finite-dimensional projective unitary representation of G. Then there exists an ordinary unitary representation $\Sigma : \tilde{G} \to U(V)$ of $\tilde{G}$ such that $\Pi \circ \Phi = Q \circ \Sigma$. Here, $Q : U(V) \to PU(V)$ is the quotient homomorphism and q its associated Lie algebra homomorphism. Any such Σ is irreducible if and only if Π is irreducible. It is possible to choose Σ so that $\det(\Sigma(A)) = 1$ for all $A \in \tilde{G}$, and Σ is unique if we require this condition.*

2. *Let Σ be a finite-dimensional irreducible representation of $\tilde{G}$. Then the kernel of the associated projective unitary representation $Q \circ \Sigma$ contains the kernel of the covering map Φ. Thus $Q \circ \Sigma$ factors through G and gives rise to a projective unitary representation of G.*

Proof. Let $\mathfrak{g}$ be the Lie algebra of G. The previous proposition 2.46 tells us that there is an representation $\sigma : \mathfrak{g} \to \mathrm{u}(V)$ such that $q \circ \sigma = \phi$. We then define $\tilde{\sigma} : \tilde{\mathfrak{g}} \to \mathrm{u}(V)$ of the Lie algebra $\tilde{\mathfrak{g}}$ of $\tilde{G}$ by setting $\tilde{\sigma}(X) = \sigma(\phi(X))$, for all $X \in \tilde{\mathfrak{g}}$. Note that $\tilde{G}$ is simply connected, and hence we have a unique representation $\Sigma : \tilde{G} \to \mathrm{U}(V)$ such that $\Sigma(e^X) = e^{\tilde{\sigma}(X)}$ for all $X \in \mathfrak{g}$. Note that

$$q \circ \tilde{\sigma} = q \circ \sigma \circ \phi = \pi \circ \phi,$$

we may conclude that $Q \circ \Sigma = \Pi \circ \Phi$. Furthermore, if Σ maps into $\mathrm{SU}(V)$, then $\sigma = \tilde{\sigma} \circ \phi^{-1}$ maps into $\mathrm{su}(n)$. This condition uniquely determines σ and thus also $\tilde{\sigma}$ and Σ.
Note that $\ker(\Phi)$ is a discrete normal subgroup of $\tilde{G}$, thus central. Thus, for all $A \in \ker(\Phi)$, we obtain

$$\Sigma(A)\Sigma(B) = \Sigma(AB) = \Sigma(BA) = \Sigma(B)\Sigma(A)$$

for all $B \in \tilde{G}$. Hence, Σ is an intertwining map from V to itself. However, we know that V is irreducible and hence by Schur's Lemma 2.40 we know that $\Sigma = cI$ with $|c| = 1$ since $\Sigma(A) \in \mathrm{U}(V)$ for all $A \in \ker(\Phi)$. Hence, A is in the kernel of the projective representation $Q \circ \Sigma$. □

2.7.3. The Adjoint Representation

Another important operation on elements of the Lie algebra are the adjoint map which will play an important role in finding the roots of of a Lie group such as SU(3) as we will see in chapter 4. They can be defined for both the Lie algebra as the Lie group.

Definition 2.48. *Let G be matrix Lie group, with Lie algebra $\mathfrak{g}$. Then for each $A \in G$, define a linear map $Ad_A : \mathfrak{g} \to \mathfrak{g}$ by the formula*

$$Ad_A(X) = AXA^{-1}.$$

Note that by lemma 2.22 point 4 we have that $\mathrm{Ad}(X) \in \mathfrak{g}$ for all $X \in \mathfrak{g}$, hence this map is well-defined. Also note that the map $A \to \mathrm{Ad}_A$ is a homomorphism of G into $\mathrm{GL}(\mathfrak{g})$, the group of all linear transformations of $\mathfrak{g}$.

Furthermore, note that for this map, the following holds by means of direct calculation $\mathrm{Ad}_A([X,Y]) = [\mathrm{Ad}_A(X), \mathrm{Ad}_A(Y)]$. Also note that we are working in finite dimensional real vector space $\mathfrak{g}$ of some dimension k. Therefore $\mathrm{GL}(\mathfrak{g}) \cong \mathrm{GL}(k;\mathbb{R})$ for some $k \in \mathbb{N}$. Then we can easily see that the map $\mathrm{Ad}_A : G \to \mathrm{GL}(\mathfrak{g})$ is continuous.

Definition 2.49. *Let $\mathfrak{g}$ be a Lie algebra with Lie bracket $[\cdot,\cdot]$. We define the adjoint map $ad : \mathfrak{g} \to gl(\mathfrak{g})$ to be*

$$ad(X)Z = [X,Z], \ X, Z \in \mathfrak{g}$$

We still need to verify that this is indeed a representation. For this, note that $\mathrm{ad}(X)$ is indeed a linear map since the Lie bracket is linear in its second component. Next, note that we still need to check whether

$$\mathrm{ad}([X,Y]) = [\mathrm{ad}(X), \mathrm{ad}(Y)].$$

Observe that

$$\mathrm{ad}([X,Y])Z = [[X,Y],Z] = -[Z,[X,Y]],$$

and

$$\begin{aligned}[\mathrm{ad}(X), \mathrm{ad}(Y)]Z &= \mathrm{ad}(X)\mathrm{ad}(Y)Z - \mathrm{ad}(Y)\mathrm{ad}(X)Z \\ &= [X,[Y,Z]] - [Y,[X,Z]] \\ &= [X,[Y,Z]] + [Y,[Z,X]].\end{aligned}$$

Now $\mathrm{ad}([X,Y]) = [\mathrm{ad}(X), \mathrm{ad}(Y)]$ follows from the Jacobi identity.

There is a relation between these two maps, as one might already have guessed.

Proposition 2.50. *Let G be a matrix Lie group, let $\mathfrak{g}$ be its Lie algebra, and let $Ad : G \to GL(\mathfrak{g})$ be defined as definition 2.48. Let $ad : \mathfrak{g} \to gl(\mathfrak{g})$ be the associated Lie algebra map. Then for all $X, Y \in \mathfrak{g}$*

$$ad_X(Y) = [X,Y].$$

Proof. By point 3 of theorem 2.27 ad can be calculated as

$$\mathrm{ad}_X = \frac{d}{dt}\mathrm{Ad}_{e^{tX}}\Big|_{t=0}.$$

Hence,

$$\mathrm{ad}_X(Y) = \frac{d}{dt} e^{tX} Y e^{-tX}\Big|_{t=0} = [X,Y],$$

as claimed. □

2.7.4. Constructing New Representations

So far, we have only seen some elementary Lie group and Lie algebra representations, however, in our analysis of angular momentum and spin we need some new representations which arise from the old ones. This is analogous to what we have seen in chapter 1.

Definition 2.51. *Suppose* (Π_1, V_1) *and* (Π_2, V_2) *are representations of a matrix Lie group* G. *The direct sum of these two representations is the representation* $\Pi_1 \oplus \Pi_2 : G \to GL(V_1 \oplus V_2)$ *given by*

$$(\Pi_1 \oplus \Pi_2)(A) = \Pi_1(A) \oplus \Pi_2(A).$$

The tensor product of Π_1 *and* Π_2 *is the representation* $\Pi_1 \otimes \Pi_2 : G \to GL(V_1 \otimes V_2)$ *given by*

$$(\Pi_1 \otimes \Pi_2)(A) = \Pi_1(A) \otimes \Pi_2(A).$$

The dual of Π_1 *is the representation* $\Pi_1^{tr} : G \to GL(V^*)$ *given by*

$$\Pi_1^{tr}(A) = \Pi_1(A^{-1})^{tr} = (\Pi_1(A)^{tr})^{-1}.$$

Similarly, the tensor product of Lie algebra representations and its dual can be defined by

$$\begin{aligned}
(\pi_1 \oplus \pi_2)(X) &= \pi_1(X) \oplus \pi_2(X) \\
(\pi_1 \otimes \pi_2)(X) &= \pi_1(X) \otimes I + I \otimes \pi_2(X) \\
\pi_1^{tr}(X) &= -\pi_1(X)^{tr}
\end{aligned}$$

Note the difference in the formulas for the tensor product. This is easy to motivate, because

$$\frac{d}{dt}\left(\Pi_1(e^{tX}) \otimes \Pi_2(e^{tX})\right)\Big|_{t=0} = \pi_1(X) \otimes I + I \otimes \pi_2(X)$$

As a final remark, suppose that we have two spaces V and W with inner product $\langle \cdot, \cdot \rangle_V$ and $\langle \cdot, \cdot \rangle_W$, then we can define an inner product on the space $V \otimes W$ as follows.

Theorem 2.52. *Assume that* V *and* W *are Hilbert spaces, then* $V \otimes W$ *is a Hilbert space with inner product*

$$\langle v \otimes w, v' \otimes w' \rangle_{V \otimes W} = \langle v, v' \rangle_V \langle w, w' \rangle_W.$$

If Π_1 *and* Π_2 *are unitary representations of* V *and* W *of a Lie group* G, *then* $\Pi_1 \otimes \Pi_2$ *is again a unitary representation.[9]*

2.8. Complete Reducibility

In chapter 1 section 1.2.1 we introduced some orthogonality relations for groups with finite number of elements. Now we consider Lie groups with an infinite number of elements. Therefore, the "avaraging" trick we used for constructing inner products will not work as well with since we have to consider an infinite number of elements to sum over. Therefore, we need to introduce a certain measure. The existence of this measure will not be proven in this report, since it will not add much insight in representations theory and is just measure theoretic of nature. The approach for proving that every compact Lie group is completely reducible is the following.

Theorem 2.53. *Let* G *be a compact Lie group. There exists a unique Borel measure* μ *on* G *such that*

$$\int_G f(hg)\,d\mu(g) = \int_G f(g)\,d\mu(g) = \int_G f(gh)\,d\mu(g),$$

for all continuous functions $f : G \to \mathbb{C}$ *and all* $h \in G$. *This unique measure* μ *is called the Haar measure on* G.

We are now in the position to extend the inner product of $C(G) = \{f : G \to \mathbb{C}\}|f\text{continuous}\}$ to:

$$\langle f_1, f_2\rangle_G = \int_G f_1(g)\overline{f_2(g)}d\mu(g).$$

Then we can extend Schur's orthogonality relations in the following way.

Lemma 2.54. *Let G be a matrix Lie group, V a finite dimensional Hilbert space, and $\Pi : G \to GL(V)$ a unitary representation. Then Π is completely reducible. Similarly, if $\mathfrak{g}$ is a real Lie algebra and π is a finite-dimensional skew-adjoint representation of $\mathfrak{g}$, then π is completely reducible.*

Proof. Let $W \subset V$ be an invariant subspace. Let $W^\perp$ be the orthogonal complement of W in V. $W^\perp$ is an invariant subspace. Take $v \in W^\perp$ and $w \in W^\perp$ arbitrary then

$$\langle \Pi(g)v, w\rangle = \langle v, \Pi(g)^* w\rangle = \langle v, \Pi(g^{-1})w\rangle = 0,$$

since $\Pi(g^{-1})w \in W$. Hence, $\Pi(g)W^\perp \subset W^\perp$. The steps for the for π are the same, except $\pi(g)^* = -\pi(g)$.

We will now prove the result by induction on $n = \dim(V)$. If $n = 1$, Π is irreducible. Suppose that any representation on a Hilbert space of dimension less than n is completely reducible. Let $W \neq 0$ be an invariant subspace of V with smallest dimension, then $\Pi|_W$ is irreducible. Since $\dim(W^\perp) < n$ and $\Pi|_{W^\perp}$ is unitary, we can conclude that $\Pi_{W^\perp}$ is the direct sum of irreducible representations. Therefore, $\Pi = \Pi_W \oplus \Pi_{W^\perp}$ is also reducible. □

Lemma 2.55. *Let G be a compact Lie group, and let Π be a representation of G on a finite dimensional vector space V. Then there exists an inner product on V such that Π is a unitary representation.*

Proof. Let $\langle\cdot,\cdot\rangle$ be an arbitrary inner product on V. Define a new inner product on V by

$$(v, w)_G = \int_G \langle \Pi(g)v, \Pi(g)w\rangle d\mu(g) \text{ with } v, w \in V.$$

Then, for $h \in G$ and $v, w \in V$ arbitrary we have,

$$\begin{aligned}(\Pi(h)v, \Pi(h)w)_G &= \int_G \langle \Pi(h)\Pi(g)v, \Pi(h)\Pi(g)w\rangle d\mu(g)\\ &= \int_G \langle \Pi(hg)v, \Pi(hg)w\rangle d\mu(g)\\ &= \int_G \langle \Pi(g)v, \Pi(g)w\rangle d\mu(g)\\ &= (v, w)_G.\end{aligned}$$

Note that we made use of the invariance of the Haar measure. □

With this observation, we can make the statement which is at the heart of this section.

Theorem 2.56. *Let G be a compact Lie group. Every finite dimensional representation of G is completely reducible.*

Proof. Let Π be a finite dimensional rperesentation of G. We already have seen in lemma 2.54 that every finite dimensional unitary representation is completely reducible. By the previous lemma 2.55, π can be reduced into a unitary representation and Π is completely reducible. □

Theorem 2.57. *The groups $SO(n)$ and $SU(n)$ are compact.*

Proof. Note that we are working with the Hilbert-Schmidt norm. The group SU(n), for example, is defined by setting $(U^*U)_{jk} = \delta_{jk}$ for each j and k and by setting $\det(U) = 1$. These groups are also subsets of $M_n(\mathbb{C})$. Also note that each of these groups has the property that each column of any matrix in the group is a unit vector. Thus, each group is a bounded and closed subset of $M_n(\mathbb{C})$. The case for SO(n) goes similar. □

With the previous results we can conclude that every finite dimensional representation of SO(n) or SU(n) is completely reducible, which is a remarkable result indeed!

3

Angular Momentum and Spin

Thus far we have seen quite some representation theory of groups which have a finite number of elements, in chapter 1, and Lie groups and algebras, which are treated in chapter 2 of this report. We have now introduced all the necessary tools to give and explain an application of representation theory in quantum physics: spin representations. In this chapter we will closely follow chapter 17 of [11] with the exception of section 17.3 of [11].

We will first explain the notion of angular momentum in classical mechanics and then make the connection with quantum mechanics. In classical mechanics, angular momentum is the generator of rotations [11]. This may sound daunting, so we will explain the intuition behind this statement. Suppose we have a particle living in $\mathbb{R}^n$ for which the Hamiltonian is given by $\mathcal{H}(\mathbf{x},\mathbf{p}) = \frac{1}{2m}\sum_{j=1}^{n} p_j^2 + V(\mathbf{x})$. Here p_j denotes the momentum in the j-th coördinate, $V(\mathbf{x})$ denotes the its potential and m denotes its mass. From classical mechanics [23] we know that the Hamilton equations are as follows:

$$\frac{dx_j}{dt} = \frac{\partial \mathcal{H}}{\partial p_j}$$
$$\frac{dp_j}{dt} = -\frac{\partial \mathcal{H}}{\partial x_j}$$
$$.$$

Lets take the same particle and restrict the space to $\mathbb{R}^2$. When we solve the Hamilton equations for the angular momentum operator $J(\mathbf{x},\mathbf{p}) = x_1 p_2 - x_2 p_1$ we see that its solution is of the form [11]

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix}$$
$$\begin{pmatrix} p_1(t) \\ p_2(t) \end{pmatrix} = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} p_1(0) \\ p_2(0) \end{pmatrix}.$$

Here $t \in [0, 2\pi)$. This can be directly verified by plugging $J(\mathbf{x},\mathbf{p})$ into the Hamilton equations and solving the differential equations. What is interesting to note is that the solution consists of rotations of $\mathbf{x}$ and $\mathbf{p}$. Hence, we say that angular momentum is the Hamiltonian generator of rotations.

When a system has rotational symmetry angular momentum is a conserved quantity [23]. In quantum mechanics, angular momentum is the (infinitesimal) generator of a one parameter group of unitary group of rotation operators [11]. This may sound daunting and confusing, so we will give an example involving generators and explain the connection to angular momentum in more detail in the next section.

The first abstract notion is that of a generator regarding Lie groups. We will explain this by means of SO(3) and so(3), which we will use quite extensively in this chapter as angular momentum and so(3) are linked to each other. In chapter 2 we saw that the basis of so(3) could be exponentiated surjectively to a neighborhood of the identity of SO(3) by theorem 2.32 and theorem 2.35. We also showed that SO(3) is connected, see example 2.5. Therefore, we know that these elements generate the whole group SO(3). In that case we say that the the basis so(3) are infinitesimal generators of SO(3) [14]. This is just a simple example and in no way accounting for the complexity regarding generators of Lie groups.

The quantum angular momentum is also conserved in systems with rotational symmetry[11], meaning that if the Hamiltonian $\mathscr{H}$ is invariant under rotations, then $\mathscr{H}$ commutes with the angular momentum operators. It will be shown that the components of the angular momentum operator in $\mathbb{R}^3$ satisfy the same commutation relations as the commutation relations of the Lie algebra so(3) of the rotation group SO(3). Hence, if $\mathscr{H}$ commutes with each component of the angular momentum operator, each eigenspace of $\mathscr{H}$ is invariant under the angular momentum operator, as we already saw in section 1.7 with the group of Schrödinger's equation. Therefore, the eigenspace is a representation of the Lie algebra so(3).

Because of this, it is of interest to study the irreducible representations of so(3), as these will tell us a lot about the structure of the solution spaces to the equation $\mathscr{H}\psi = \lambda\psi$ where ψ is an eigenvector of the Hamiltonian and $\mathscr{H}$ is rotationally invariant. It is important and necessary for us to first understand these important concepts before diving into the theory of spin, as these two are closely related.

The aim of this chapter, as mentioned many times before, is to explain what spin is by means of representation theory. We do this by first analysing the connection between angular momentum and Lie algebras, then turning to the irreducible representations of SO(3) and so(3). Since in elementary quantum mechanics the Hilbert particles live in the space $L^2(\mathbb{R}^3)$, we then turn to representations of SO(3) inside $L^2(\mathbb{R}^3)$. Having identified the irreducible representations of SO(3) inside $L^2(\mathbb{R}^3)$, we turn to the analysis of spin and the addition of angular momentum and spin in quantum mechanics.

3.1. Angular Momentum Operators

We begin with introducing the concept of angular momentum in quantum mechanics, which will be given as

$$\mathbf{J} = \mathbf{X} \times \mathbf{P}$$

Here $\mathbf{X}$ and $\mathbf{P}$ are the position and the momentum operators. To give an example in 3 dimensions, one may consider $J_1 = X_2P_3 - X_3P_2 = x_2\frac{\partial}{\partial x_3} - x_3\frac{\partial}{\partial x_2}$. Note that each component of the angular momentum involves products of distinct components of the operators $\mathbf{X}$ or $\mathbf{P}$, which commute. So in the expression for J_3 it does not matter whether we write X_2P_3 or P_3X_2.

As we know from elementary calculations[8], one can write

$$\frac{1}{i\hbar}[J_1, J_2] = J_3; \quad \frac{1}{i\hbar}[J_2, J_3] = J_1; \quad \frac{1}{i\hbar}[J_3, J_1] = J_2;$$

However, when one writes J_3 explicitly, one would obtain [11]

$$\begin{aligned}(J_3\psi)(\mathbf{x}) &= -i\hbar\left(x_1\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_1}\right)\psi(\mathbf{x}) \\ &= -i\hbar\frac{d}{d\theta}\psi(R_\theta\mathbf{x})\bigg|_{\theta=0}.\end{aligned}$$

Here R_θ is the counterclockwise rotation by angle θ in the (x_1, x_2) plane given by the matrix

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

We can construct similar expressions for J_1 and J_2. What is important to remember from this section is the following.: we now see that $J_i, i = 1,2,3$ are closely related to rotations.

To conlcude this seciton, we can also look at the natural action of the rotation group SO(3) on $L^2(\mathbb{R}^3)$.

Definition 3.1. *For each $R \in SO(3)$, define $\Pi(R) : L^2(\mathbb{R}^3) \to L^2(\mathbb{R}^3)$ by*

$$(\Pi(R)\psi)(\mathbf{x}) = \psi\left(R^{-1}\mathbf{x}\right)$$

.

This map is continuous and unitary, see [10].

3.2. The Irreducible Representations of so(3)

As said in the introduction, it is of importance to understand the finite-dimensional irreducible representations of the Lie algebra so(3) in order to understand spin in later sections. The aim of this section is to obtain such finite-dimensional irreducible representations of so(3). We take all representations over the field of complex numbers with dimension of at least 1 and we take the standard basis of so(3): F_1, F_2 and F_3 used extensively in chapter 2.

Theorem 3.2. *Let $\pi : so(3) \to gl(V)$ be a finite-dimensional irreducible representation of so(3). Define operators L^+, L^- and L^3 on V by*

$$\begin{aligned} L^+ &= i\pi(F_1) - \pi(F_2) \\ L^- &= i\pi(F_1) + \pi(F_2) \\ L_3 &= i\pi(F_3). \end{aligned}$$

Let $l = \frac{1}{2}(\dim(V) - 1)$, so that $\dim(V) = 2l+1$. Then there exists a basis $v_0, v_1, ..., v_{2l}$ of V sucht that

$$\begin{aligned} L_3 v_j &= (l-j)v_j \\ L^- v_j &= \begin{cases} v_{j+1} & \text{if } j < 2l \\ 0 & \text{if } j = 2l \end{cases} \\ L^+ v_j &= \begin{cases} j(2l+1-j)v_{j-1} & \text{if } j > 0 \\ 0 & \text{if } j = 0 \end{cases}. \end{aligned}$$

Beware of the signs of L^- and L^+ operators. Note that the dimension of V completely determines the structure of the irreducible representation of so(3). Interestingly, as $\dim(V)$ is a natural number we have that $l = 0, \frac{1}{2}, 1, \frac{3}{2}, ...$ As one can see, the number l completely determines the structure of the irreducible representation.

Proof. The proof of this property is actually quite long, but not that hard. We are going to exploit the fact that we are working in an algebraïcally closed field $\mathbb{C}$ and that since π is a Lie algebra homomorphism it satisfies the same commutator relations as the $F_j, j \in \{1,2,3\}$. From the last observation we can easily deduce that

$$\begin{aligned} [L_3, L^+] &= L^+ \\ [L_3, L^-] &= -L^- \\ [L^+, L^-] &= 2L_3. \end{aligned}$$

As we are working over an algebraïcally closed field $\mathbb{C}$, the operator L_3 has at least one eigenvector v with eigenvalue λ. Note that we can use this fact to obtain

$$L_3 L^+ v = (L^+ L_3 + L^+) v = L^+(\lambda v) + L^+ v = (\lambda + 1) v.$$

We observe that either $L^+ v = 0$ or $L^+ v$ is an eigenvector for L_3 with eigenvalue $\lambda + 1$. The operator L^+ is oftentimes called the raising operator for this particular reason.

We know that the operator L_3 has finitely many eigenvectors, hence there exists a $k \in \mathbb{N}$ such that $(L^+)^k v \neq 0$ but $(L^+)^{k+1} = 0$. By applying the same steps k times we can see that $(L^+)^k v$ is an eigenvector of L_3 with eigenvalue $\lambda + k$.

Now set $v_0 = (L^+)^k v$ and $\mu = \lambda + k$. Then note that $L_3 v_0 = \mu v_0$. We forget about the original vectors v and eigenvalue λ. Define v_j as follows.

$$v_j = (L^-)^j v_0, \quad j = 0, 1, 2, \ldots$$

Using the same logic, we can easily see that

$$L_3 v_j = (\mu - j) v_j.$$

Next, we claim for $j \geq 1$ that

$$L^+ v_j = j(2\mu + 1 - j) v_{j-1}, \quad j = 1, 2, 3, \ldots.,$$

which follows by induction. L_3 has only finitely many eigenvectors, as before, thus v_j must eventually be zero. Therefore, there is an $N \in \mathbb{N}$ such that $v_N \neq 0$ but $v_{N+1} = 0$. Applying the relation of $L^+ v_j$ with $j = N + 1$ gives

$$0 = L^+ v_{N+1} = (N+1)(2\mu - N) v_N.$$

Note that $v_N \neq 0$ and that $N + 1 > 0$. Hence $(2\mu - N) = 0$. Therefore $\mu = \frac{N}{2}$. We set $l = \frac{N}{2} = \mu$, which is also stated in the theorem. We conclude this proof with the observation that the dimension of V is indeed $2l + 1$.

Note that v_j are eigenvectors of L_3 with distinct eigenvalues, hence linearly independent. Also observe that the span of the v_j's is independent under L^+, L^- and L_3 and hence under all of so(3). V is irreducible and thus the span of the v_j must be all of V. Hence, the v_j's are a basis of V. The dimension of V is therefore equal to the number of $v_j's$ which is $2l + 1$. □

We are now in the position to properly define what spin actually is in a mathematical sense.

Definition 3.3. *If (π, V) is an irreducible finite dimensional representation of so(3), then the spin of (π, V) is the largest eigenvalue of the operator $L_3 := i\pi(F_3)$. Equivalently, l is the unique number such that* $\dim(V) = 2l + 1$.

The question that arises is whether for any $l = 0, \frac{1}{2}, 1, \ldots$ there exists an irreducible representation of so(3) of dimension $2l + 1$ and more importantly, when one would find such representation, one would also know whether there exists multiple of such representations which are not isomorphic to one another. These questions will be answered by the following theorem.

Theorem 3.4. *For any $l = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ there exists an irreducible representation of so(3) of dimension $2l + 1$, and any two irreducible representations of so(3) of dimension $2l + 1$ are isomorphic.*

Proof. We begin by just defining a space V of dimension $2l + 1$ with basis $v_0, \ldots, v_{2l}$. We define the action of so(3) on this space just as we have defined in theorem 3.2. We can easily check that L^+, L^-, L_3 have the right commutation relations in order to let $\pi : \mathrm{so}(3) \to \mathrm{gl}(V)$ indeed be a representation. We do this for the first relation $[L_3, L^+] = L^+$ for $v_j, j > 0$ and the other ones are similar.

$$\begin{aligned}[L_3,L^+]v_j &= (L_3L^+ - L^+L_3)v_j = L_3 j(2l+1-j)v_{j-1} - L^+(l-j)v_j \\ &= j(2l+1-j)(l-j+1)v_{j-1} - j(2l+1-j)(l-j)v_{j-1} \\ &= j(2l+1-j)v_{j-1} = L^+v_j.\end{aligned}$$

We still have to show that V is indeed irreducible. For this end suppose that W is an invariant subspace of V and that $W \neq \{0\}$. Our goals is to show that $W = V$. So take an arbitrary $w \in W$ such that $w \neq 0$ and note that we can express w as $w = \sum_{j=0}^{2l} a_j v_j$ with $a_j \in \mathbb{C}$. Take j_0 to be the largest index such that a_j is nonzero. Therefore, as the relations in theorem 3.2 suggest $(L^+)^{j_0}w$ will be a non zero multiple of v_0. This means that v_0 belongs to W as we set W to be invariant. We can reverse this trick as with the operator L^- as defined in theorem 3.2 to show that for all $j \in \{0,...,2l\}$, v_j are elements of W. Then it naturally follows that $W = V$ as the v_j's are a basis for V as well. When we have an arbitrary irreducible representation of so(3) of dimension $2l+1$, then it has a basis as defined in theorem 3.2. We can construct an isomorphism between any two irreducible representations of dimension $2l+1$ just by mapping one basis into the other. □

Note that in the proof of theorem 3.2 irreducibility was only used to show that the span of $v_0, v_1, ..., v_{2l}$ is equal to V. Therefore, the following result is quite logical and this will be used later in this chapter for the addition of angular momentum.

Proposition 3.5. *Let (π, V) be any finite-dimensional representation of so(3), not necessarily irreducible. Suppose v_0 is a nonzero element of V such that $L^+v_0 = 0$ and $L_3v_0 = \lambda v_0$ for some $\lambda \in \mathbb{C}$. Then λ is equal to a non-negative integer or half-integer l. Furthermore, the vectors $v_0, v_1, ..., v_{2l}$ defined by*

$$v_j = (L^-)^j v_0, \ j = 0,1,...,2l,$$

span an irreducible invariant subspace of V of dimension $2l+1$, and L^+, L^-, and L_3 act on these vectors according to the formulas in theorem 3.2.

Another interesting result is the following, which will also be used in the analysis of angular momentum in the last section of this chapter.

Proposition 3.6. *Let $\pi : so(3) \to gl(V)$ be an irreducible representation of so(3). Then there exists an inner product on V, unique up to multiplication by a constant, such that $\pi(X)$ is skew self-adjoint for all $X \in so(3)$.*

Proof. Recall how the operators L^+, L^- and L_3 are defined in definition 3.2. When we impose that $\pi(X), X \in \text{so}(3)$ is skew-self-adjoint, we can see that L_3 is self adjoint trivially. It is not hard to see that L^+ and L^- are adjoints of one another. Since v_j's are eigenvectors for L_3 with distinct eigenvalues, if L_3 is self adjoint, then the v_j must be orthogonal. This is quite a standard proof, but for completeness sake we will include it. So assume that v_j and v_l are eigenvectors of L_3 with $j \neq l$. We know that L_3 is self adjoint and using the relations in theorem 3.2 we obtain $(l-j)\langle v_j, v_l\rangle = 0$. Since $j \neq l$ we know that v_j and v_l are orthogonal.

Conversely, if we have any inner product for which the v_j are orthogonal, we have that L^3 will be self-adjoint. This can be easily verified.

It remains to look at the consequences of $(L^+)^* = L^-$. Assuming this, we obtain

$$\langle v_j, v_j\rangle = \langle L^- v_{j-1}, L^- v_{j-1}\rangle = \langle v_{j-1}, L^+L^- v_{j-1}\rangle.$$

Note that $L^+L^- = L^-L^+ + 2L_3$, $L_3v_{j-1} = (l-j+1)v_{j-1}$ and $L^+v_{j-1} = (j-1)(2l-j+2)v_{j-2}$ and thus,

$$\begin{aligned}\langle v_j, v_j\rangle &= \langle v_{j-1}, L^+L^- v_{j-1}\rangle \\ &= (j-1)(2l-j+2)\langle v_{j-1}, L^- v_{j-2}\rangle + 2(l-j+1)\langle v_{j-1}, v_{j-1}\rangle.\end{aligned}$$

With the relation $L^- v_{j-2} = v_{j-1}$ in the back of our heads , we can write

$$\langle v_j, v_j \rangle = j(2l - j + 1)\langle v_{j-1}, v_{j-1} \rangle.$$

So we see that if the v_j's are orthogonal, then L^+ and L^- are adjoints if and only if the normalization condition in the above equation holds for $j = 1, ..., 2l$. Since all the terms $j(2l - j + 1)$ are positive for each $j = 1, ..., 2l$, there is no obstruction to normalizing these v_j's such that the these conditions hold. Hence, the inner product with the desired properties exists. Since the normalization of the v_j's is the only choice in defining the inner product, the inner product is unique up to multiplication by a constant. □

This proposition can be interpreted as follows: there exists an inner product on the space V such that the representation π of so(3) is skew-adjoint, see theorem 2.42, which is exactly what wants in quantum mechanics.

3.3. The Irreducible Representations of SO(3)

We will now turn to the representations of the Lie group SO(3). Note that SO(3) is connected as we have shown in example 2.5. Therefore, proposition 2.39 tells us that a Lie group representation SO(3) is irreducible if and only if the associated Lie algebra representation is irreducible. Moreover, it tells us as well that two representations of the Lie group SO(3) are isomorphic if and only if the associated Lie algebra representations are isomorphic. Hence, to classify the irreducible representations of the Lie group SO(3) it is enough to consider which irreducible representations of the Lie algebra so(3) come from a representation of the Lie group SO(3). Before we are going to proof the main result, we need an additional lemma first.

Lemma 3.7. *Suppose that G, H and K are matrix Lie groups and $\Phi : H \to K$ and $\Psi : G \to H$ are Lie group homomorphisms. Let $\Lambda : G \to K$ be the composition of Φ and Ψ and let ϕ, ψ, and λ be the Lie algebra maps associated to Φ, Ψ and Λ, respectively. Then we have*

$$\lambda = \phi \circ \psi.$$

Proof. For any $X \in \mathfrak{g}$,

$$\Lambda(e^{tX}) = \Phi(\Psi(e^{tX})) = \Phi(e^{t\psi(X)}) = e^{t\phi(\psi(X))}.$$

Hence, $\lambda(X) = \phi(\psi(X))$. □

Proposition 3.8. *Let $\pi_l : so(3) \to gl(V)$ be an irreducible representation of $so(3)$ with spin $l := \frac{1}{2}(\dim(V) - 1)$. If l is an integer (i.e., if the dimension of V is odd), there exists a representation $\Pi_l : SO(3) \to GL(V)$ such that Π_l and π_l are related as in theorem 2.27. If l is a half-integer (i.e. if the dimension of V is even) then no such representation Π_l exists.*

Proof. We begin by assuming that l is half a integer. Then we know that L_3 is diagonal in the basis $\{v_j\}$ as we have defined earlier, with eigenvalues being half integers. Therefore,

$$e^{2\pi\pi_l(F_3)} = e^{2\pi i L_3} = -I.$$

Note that for $F_3 \in$ so(3) gives $e^{2\pi F_3} = I$. Thus if there would exist a Π_l for π_l as in theorem 2.27. We would have

$$\Pi_l(I) = \Pi_l\left(e^{2\pi F_3}\right) = e^{2\pi\pi_l(F_3)} = -I$$

and this is a clear contradiction.

We will continue with premise that l is an integer. Note that we have already seen that so(3) and su(2) are isomorphic to each other in proposition 2.31. We set the isomorphism ϕ as described in proposition 2.31, which maps the basis $\{E_1, E_2, E_3\}$ of su(2) to the basis $\{F_1, F_2, F_3\}$ of so(3).

Therefore, we can define a representation of su(2) just by defining $\pi_l'(X) = \pi_l(\phi(X))$. As SU(2) is simply connected, proved in example 2.4, and theorem 2.29 tells us that there is a representation Π_l of SU(2) related to π' as in theorem 2.27. Then we can safely say

$$\Pi_l'(-I) = \Pi_l'\left(e^{2\pi E_1}\right) = e^{2\pi\pi_l'(E_1)} = e^{2\pi\pi_l(F_1)} = e^{2\pi i L_3} = I.$$

The last equality follows from the fact that the eigenvalues of L_3 are integers.

By theorem 2.35 there is a surjective homomorphism Φ from SU(2) onto SO(3) for which the associated Lie algebra homomorphism is ϕ and $\ker\Phi = \{I, -I\}$. We now make the observation that the kernel of Π_l' contains $\{I, -I\}$ and hence that the map Π_l' factors through SO(3), giving a representation Π_l of SO(3). This map $\Pi_l : \mathrm{SO}(3) \to \mathrm{GL}(V)$ satisfies $\Pi_l' = \Pi_l \circ \Phi$. From theorem 3.7 we know that the associated Lie algebra representation σ_l of so(3) satisfies $\pi_l' = \sigma_l \circ \phi$, so that $\sigma_l = \pi_l' \circ \phi^{-1} = \pi_l$. Hence, Π_l is the representation of SO(3) which we want. □

3.4. Representations Inside $L^2(S^2)$

At this stage, the physics literature will usually introduce spherical harmonics and derive the properties we are going to derive in a quite dreadful way. However, we can also consider the spherical harmonics to be the restrictions to the unit sphere of certain polynomials on $\mathbb{R}^3$[11]. Note that we have a natural unitary representation $\Pi : \mathrm{SO}(3) \to L^2(\mathbb{R}^3)$ given by $\Pi(R)\psi(x) = \psi(R^{-1}x)$ which we will make use of. Suppose we have $\psi \in L^2(\mathbb{R}^3)$ and consider the natural action of the rotation group SO(3) on this function as defined in definition 3.1. We can express ψ in terms of spherical coordinates and note that SO(3) only acts on the angle variables of ψ. Hence, it is insightfull to consider the action of SO(3) on $L^2(S^2)$, where S^2 is the unit sphere inside $\mathbb{R}^3$, given by

$$(\Pi(R)\psi)(\mathbf{x}) = \psi\left(R^{-1}\mathbf{x}\right), \quad \mathbf{x} \in S^2.$$

There exists a measure on S^2, which is invariant under SO(3). We will use this measure for computing the norm for $L^2(S^2)$. For completeness sake, we state the measure in this report. For the proof of this proposition, see [9].

Proposition 3.9. *There exists a measure ν defined by*

$$\int_{S^2} f(\mathbf{x})d\nu(\mathbf{x}) := \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} f(\theta,\psi)\sin(\theta)d\theta d\psi, \quad f \in C(S^2).$$

The measure ν is a $SO(3)$-invariant measure on S^2. Here ψ and θ can be considered polar coordinates which describe the positions on the surface S^2.

The notion of harmonic polynomials on $\mathbb{R}^3$ will be crucial in this section. These are the polynomials for which hold that $\Delta p = 0$ with Δ the Laplacian.

Definition 3.10. *Let l be a non-negative integer. Define a subspace V_l of $L^2(S^2)$ by setting V_l equal to the space of restrictions to S^2 of harmonic polynomials on $\mathbb{R}^3$ that are homogeneous of degree l. Then V_l is called the space of spherical harmonics of degree l.*

Suppose that we have that p is a homogeneous polynomial on $\mathbb{R}^3$ of some degree l. Then the restriction of p to S^2 is zero if and only if p is zero. Since p is homogeneuous of degree l, we may write

$$p(\mathbf{x}) = |\mathbf{x}|^l p\left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right),$$

for all $\|\mathbf{x}\| \neq 0$. Which is zero when p is zero on S^2. Hence, by continuity, also for $\mathbf{x} = 0$. Therefore, we can think of elements of V_l both as functions on S^2 or as $\mathbb{R}^3$. In the next theorem some interesting properties of these spaces V_l are stated.

Theorem 3.11. *The spaces V_l have the following properties.*

- *Each V_l has dimension $2l+1$.*
- *Each V_l is invariant under the action of the rotation group and irreducible under this action.*
- *For $l \neq m$, the spaces V_l and V_m are orthogonal in $L^2(S^2)$.*
- *The Hilbert space $L^2(S^2)$ decomposes as the orthogonal direct sum of the V_l's, as l ranges over the non-negative integers.*

This theorem will turn out to be crucial in analysing the irreducible representations of SO(3) inside $L^2(\mathbb{R}^3)$. We begin by proving the first point of this theorem. We need to introduce an additional lemma in order to properly define an inner product on the space of polynomials. This inner product then has some nice properties which we will use when we want to determine the dimension of the spaces V_l.

Lemma 3.12. *Suppose that μ is a smooth, strictly positive density on $\mathbb{C}^n$ and that F and G are sufficiently nice (but not necessarily holomorphic) functions on $\mathbb{C}^n$. Then*

$$\int_{\mathbb{C}^n} \overline{F(\boldsymbol{z})} \frac{\partial G}{\partial z_j} \mu(\boldsymbol{z}) d\boldsymbol{z} = \\ -\int_{\mathbb{C}^n} \overline{\frac{\partial F}{\partial \bar{z}_j}} G(\boldsymbol{z}) \mu(\boldsymbol{z}) d\boldsymbol{z} - \int_{\mathbb{C}^n} \overline{\frac{\partial \log \mu}{\partial \bar{z}_j}} \overline{F(\boldsymbol{z})} G(\boldsymbol{z}) d\boldsymbol{z}.$$

where $d\boldsymbol{z}$ denotes the $2n$-dimensional Lebesgue measure on $\mathbb{C}^n \cong \mathbb{R}^{2n}$.

We will not proof this result here, since this is more complex analysis than representation theory. The proof can be found in [11]. We will now show that there exists an inner product on the space of polynomials $\mathscr{P}_l$, the space of homogeneous polynomials of degree l, which we will use in determining the dimension of the space $\mathscr{P}_l$. We will us the dimension of $\mathscr{P}_l$ to determine the dimension of the space V_l.

Theorem 3.13. *Let $\mathscr{P}$ denote the space of polynomials on $\mathbb{R}^3$ with complex coefficients. There exists an inner product $\langle \cdot,\cdot \rangle$ on $\mathscr{P}$ with the property that*

$$\langle p, \Delta q \rangle_P = \langle x^2 p, q \rangle_P$$

where,

$$x^2 = x_1^2 + x_2^2 + x_3^2.$$

Proof. Note that every polynomial on $\mathbb{R}^3$ has a holomorphic extension to $\mathbb{C}^3$, denoted by $p_{\mathbb{C}}$. Then we can define an inner product in the following way.

$$\langle p, q \rangle_{\mathscr{P}} = \int_{\mathbb{C}^3} \overline{p_{\mathbb{C}}(\mathbf{z})} q_{\mathbb{C}}(\mathbf{z}) \frac{e^{-|z|^2}}{\pi^3} d^6 z.$$

By lemma 3.12, and noting that $\frac{\partial p_{\mathbb{C}}}{\partial \bar{z}_j}(\mathbf{z}) = 0$, we see that

$$\int_{\mathbb{C}^3} \overline{p_{\mathbb{C}}(\mathbf{z})} \frac{\partial q_{\mathbb{C}}}{\partial z_j}(\mathbf{z}) \frac{e^{-|z|^2}}{\pi^3} d^6 z = \int_{\mathbb{C}^3} \overline{z_j p_{\mathbb{C}}(\mathbf{z})} q_{\mathbb{C}}(\mathbf{z}) \frac{e^{-|z|^2}}{\pi^3} d^6 z$$

for all $p, q \in \mathscr{P}$ and all $j = 1,2,3$. Therefore we can conclude that

$$\left\langle p, \frac{\partial q}{\partial x_j} \right\rangle_{\mathscr{P}} = \langle x_j p, q \rangle_{\mathscr{P}}.$$

From this relation and the definition of the laplacian one would obtain the desired relation of the lemma. □

Before we determine the dimension of the space V_l we have to make an important observation. The number of triplets $(l_1, l_2, l_3) \in \mathbb{N}^3$ such that $l = l_1 + l_2 + l_3$ is equal to $\frac{(l+1)(l+2)}{2}$. Also note that $x_1^{l_1} x_2^{l_2} x_3^{l_3}$ form a basis of of the space of polynomials on $\mathbb{R}^3$ homogeneous of degree l. Therefore we can conclude that the dimension of $\mathscr{P}_l$ is equal to $\frac{(l+1)(l+2)}{2}$.

Corollary 3.14. *If $\mathscr{P}_l$ denotes the space of polynomials on $\mathbb{R}^3$ that are homogeneous of degree l, then the laplacian Δ maps $\mathscr{P}_l$ onto $\mathscr{P}_{l-2}$ for all $l \geq 2$. Hence, for all $l \geq 2$ we have*

$$\begin{aligned}\dim(V_l) &= \dim(P_l) - \dim(P_{l-2}) \\ &= \frac{(l+2)(l+1)}{2} - \frac{l(l-1)}{2} \\ &= 2l + 1.\end{aligned}$$

Proof. Equip $\mathscr{P}_l$ and $\mathscr{P}_{l-2}$ with the inner product of the previous theorem 3.13. Note that for maps of one finite dimensional inner product space into the other, the orthogonal complement of the image is the kernel of the adjoint. By theorem 3.13 we know that the adjoint of $\Delta : \mathscr{P}_l \to \mathscr{P}_{l-2}$ is $x^2 : \mathscr{P}_{l-2} \to \mathscr{P}_l$. Note that the map $x^2 : \mathscr{P}_{l-2} \to \mathscr{P}_l$ is injective, as x^2 is only zero for $(x_1, x_2, x_3) = (0, 0, 0)$. Hence, the orthogonal complement of the image of Δ is $\{0\}$. Since $\mathscr{P}_l$ and $\mathscr{P}_{l-2}$ are finite dimensional, we know that Δ maps $\mathscr{P}_l$ onto $\mathscr{P}_{l-2}$. □

We are now going to show that V_l is irreducible under the action Π of SO(3). We are going to do this with the aid of the associated Lie algebra so(3) and the result that a Lie group representation of SO(3) is irreducibel if and only if the associated Lie algebra representation is irreducibel as well, see theorem 2.39. As we have showed earlier, the restriction to the sphere is injective on homogeneous polynomials. Therefore we can consider the elements of V_j as polynomials on $\mathbb{R}^3$. The Lie algebra action of π associated to Π is then given in terms of the angular momentum operators.

Lemma 3.15. *Let $L_3 = i\pi(F_3) = \frac{1}{\hbar}\hat{J}_3$ and let $L^+ = i\pi(F_1) - \pi(F_2) = \frac{1}{\hbar}(\hat{J}_1 + i\hat{J}_2)$. For any non-negative integer l, the polynomial $p(x_1, x_2, x_3) := (x_1 + ix_2)^l$ belongs to V_l and satisfies*

$$L_3 p = lp$$

and

$$L^+ p = 0.$$

Proof. We will begin with the claim that the polynomial p is harmonic. We will do this by means of direct calculation. For $l = 0, 1$ this is trivial, so assume $l \geq 2$.

$$\begin{aligned}\Delta p &= \Delta(x_1 + ix_2)^l \\ &= \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}\right)(x_1 + ix_2)^l \\ &= l(l-1)(x_1 + ix_2)^{l-2} - (x_1 + ix_2)^{l-2} \\ &= 0.\end{aligned}$$

Note that applying L_3 to p gives

$$\begin{aligned}&-i\left(x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1}\right)(x_1 + ix_2)^l \\ &= -i\left(x_1 l(x_1 + ix_2)^{l-1}(i) - x_2 l(x_1 + ix_2)^{l-1}\right) \\ &= l(x_1 + ix_2)^l = lp.\end{aligned}$$

We can do the same with $L^+ := i\pi(F_1) - \pi(F_2)$ to obtain

$$
\begin{aligned}
&-i\left(x_2\frac{\partial}{\partial x_3} - x_3\frac{\partial}{\partial x_2}\right)p + \left(x_3\frac{\partial}{\partial x_1} - x_1\frac{\partial}{\partial x_3}\right)p \\
&= -i(-x_3 l(x_1+ix_2)^{l-1}(i)) + x_3 l(x_1+ix_2)^{l-1} \\
&= 0,
\end{aligned}
$$

as we have claimed. □

We are now in the position to proof the second point of theorem 3.11. Namely, that each V_l is invariant under the action of the rotation group and irreducible under this action.

Corollary 3.16. *The space V_l is irreducible under the action of $SO(3)$.*

Proof. By proposition 3.5, for which the condition are satisfied by the previous lemma 3.15, when we apply L^- repeatedly tot the polynomial p, then we obtain a chain of eigenvectors of length $2l+1$. Note that we have already shown that the dimension on V_l is equal to $2l+1$. Also note that the chain of eigenvectors also span an irreducible subspace of dimension $2l+1$. This means that the elements of the chain must span V_l, hence V_l is irreducible. □

We will now proof the third point of theorem 3.18. For a quick reminder, this point says that for $l \neq m$, the spaces V_l and V_m are orthogonal in $L^2(S^2)$.

Proof. Note that each V_l is an irreducible representation of SO(3) and when $l \neq m$ we know that V_l, V_m are not isomorphic since they have different dimension. Note that the orthogonal projection operator of $P : L^2(\mathbb{R}^3) \to V_l$ forms an intertwining map between V_l and V_m. However, we can see by Schur's lemma and the fact that V_l and V_m are not isomorphic, that $P = 0$. Hence, we can conclude that V_l and V_m are orthogonal inside $L^2(S^2)$ for $m \neq l$, which proves point 3. □

At this point we showed points 1 to 3. The last point to be proven is the following: the Hilbert space $L^2(S^2)$ decomposed as the orthogonal direct sum of the V_l's, as l ranges over the non-negative integers. The strategy to proof this result is the following. We will show that the restrictions to S^2 of polynomials on $\mathbb{R}^3$ form a dense subspace of $L^2(S^2)$. When we have established this we will show that the space of restrictions of polynomials on S^2 coincides with the space of restrictions of harmonic polynomials. Then we know that the span of the spaces of harmonic polynomials, the V_l's, are dense in V_l.

We will now show that the restrictions to S^2 of polynomials on $\mathbb{R}^3$ form a dense subspace of $L^2(S^2)$. Before we can proof, we need an additional theorem. This theorem will not be proven, since it will diverge us too much in the area of measure theory.

Theorem 3.17. *Suppose μ is a measure on the Borel σ-algebra in a locally compact, seperable metric space X. Suppose also that $\mu(K) < \infty$ for each compact subset K of X. Then the space of continuous functions of compact support on X, denoted by $C_c(X)$, is dense in $L^p(X,\mu)$, for all p with $1 \leq p < \infty$. For our applications, the compact support is the compact set of points in X such that the function $\psi : X \to \mathbb{R}$ is non-zero[11].*

Theorem 3.18. (Stone – Weierstrass) *Let X be a compact metric space, and let A be an algebra in $\mathscr{C}(X;\mathbb{C})$. If $A_{\mathbb{C}}$ contains the constant functions and separates points and is closed under complex conjugation, then $\mathscr{A}$ is dense in $\mathscr{C}(X;\mathbb{R})$ with respect to the supremum norm. By separate points we mean that for $x, y \in X$ there exists $f \in A$ such that $f(x) \neq f(y)$.*

A proof for theorem 3.18 can be found in for example [17]. Note that the unit sphere in $\mathbb{R}^3$ is closed and bounded, hence compact by Heine-Borel. One can easily verify that the complex polynomials separate points, are closed under complex conjugation and trivially contain the constant functions. Therefore, the complex polynomials are dense in $\mathscr{C}(S^2,\mathbb{C})$. Hence, by theorem 3.17, we can see that the polynomials with complex coefficients are indeed dense in $L^2(S^2)$. The only point needed to proof is that the space of restrictions to S^2 of polynomials coincides with the space of restrictions of harmonic polynomials.

Corollary 3.19. *Let l be a non-negative integer and let $k = \frac{l}{2}$ if l is even and let $k = \frac{l-1}{2}$ if l is odd. Then each $p \in P_l$ can be decomposed in the form*

$$p(\mathbf{x}) = p_0(\mathbf{x}) + |\mathbf{x}|^2 p_1(\mathbf{x}) + |\mathbf{x}|^4 p_2(\mathbf{x}) + \ldots + |\mathbf{x}|^{2k} p_k(\mathbf{x}),$$

where each $p_j(\mathbf{x})$ is a harmonic polynomial that is homogeneous of degree $l-2j$. In particular, the restriction of p to S^2 satisfies

$$p|_{S^2} = (p_0 + p_1 + \ldots + p_k)|_{S^2},$$

where $p_0 + p_1 + \ldots + p_k$ is a (nonhomogeneous) harmonic polynomial.

Proof. We are going to proof this by induction on l. For the base case $l = 1$ we take know that all $p \in \mathscr{P}_1$ are harmonic and the decomposition we are looking for is simply $p = p_0$. Take an $l \geq 2$ arbitrary and assume that the result holds for all degrees less than l. Note that 3.13 tells us that $\mathscr{P}_l$ decomposes as an orthogonal direct sum of the kernel of Δ and the image of $\mathscr{P}_{l-2}$ under the multiplication by $|\mathbf{x}|^2$. Therefore, any $p \in \mathscr{P}_l$ can be decomposed into $p = p_0 + |\mathbf{x}|^2 q_0$, where p_0 is harmonic and q_0 belongs to $\mathscr{P}_{l-2}$. By induction, q_0 can be decomposed in the same fashion, resulting in the decomposition of $p \in \mathscr{P}_l$. □

With this proposition we conclude that the Hilbert space $L^2(S^2)$ decomposes as the orthogonal direct sum of the V_l's, as l ranges over the non-negative integers.

3.5. Representations Inside $L^2(\mathbb{R}^3)$

In the previous section we have showed that the restriction map for homogeneous polynomials on $\mathbb{R}^3$ to S^2 is injective. So the space V_l can be considered as a space of functions on S^2 or $\mathbb{R}^3$. From now on, we consider V_l to be the space of harmonic polynomials on $\mathbb{R}^3$ which are homogeneous of degree l.

Definition 3.20. *Suppose l is a non-negative integer and f is a measurable function on $(0,\infty)$ such that*

$$\int_0^\infty |f(r)|^2 r^{2l+2} dr < \infty.$$

Let $V_{l,f} \subset L^2(\mathbb{R}^3)$ denote that space of the functions ψ of the form

$$\psi(\boldsymbol{x}) = p(\boldsymbol{x}) f(|\boldsymbol{x}|)$$

where $p \in V_l$.

Note that the condition on $f(r)$ makes $\psi(\mathbf{x})$ square integrable on $\mathbb{R}^3$. In physics literature $\psi(\mathbf{x})$ is usually expressed as

$$\psi(\mathbf{x}) = Y_{lm}\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right) g(|\mathbf{x}|).$$

Here Y_{lm} is the restriction to the sphere of a harmonic polynomial that is homogeneous of degree l and $g(|\mathbf{x}|)$ accounts for the radial part. However, taking this definition will make calculations a bit harder in for example the analysis of the hydrogen atom, which we will not pursue here. In my honest opinion, it is still worthwhile to analyse the irreducible representations of SO(3) inside $L^2(\mathbb{R}^3)$.

3.5.1. Irreducible Representations Inside $L^2(\mathbb{R}^3)$

Before we can analyse the irreducible representations inside $L^2(\mathbb{R}^3)$ in depth, we need an auxiliary theorem.

Theorem 3.21. *Suppose that (X_1, μ_1) and (X_2, μ_2) are σ-finite measure spaces, meaning that the sets X_1 and X_2 can be covered by at most countably many measurable sets with finite measure. Then there is a unique unitary map*

$$p : L^2(X_1, \mu_1) \hat{\otimes} L^2(X_2, \mu_2) \to L^2(X_1 \times X_2, \mu_1 \times \mu_2)$$

such that

$$p(\phi \otimes \psi)(x, y) = \phi(x)\psi(y)$$

for all $\phi \in L^2(X_1, \mu_1)$ and $\psi \in L^2(X_2, \mu_2)$.

Here $\hat{\otimes}$ denotes the completion of the Hilbert space $L^2(\mathbb{R}^3) \otimes L^2(\mathbb{R}^3)$, as the tensor of two arbitrary complete Hilbert spaces might not be complete[11]. I did this proof in Appendix C and it is just quite technical and really cumbersome to set up the right framework, which is why we omitted it here. We are finally in the position to proof theorem that is at the heart of this section.

Theorem 3.22. *Every space of the form $V_{l,f} \subset L^2(\mathbb{R}^3)$ is invariant and irreducible under the action of $SO(3)$. Conversely, every finite-dimensional, irreducible $SO(3)$ invariant subspace of $L^2(\mathbb{R}^3)$ is of the form $V_{l,f}$ for some non-negative integer l and some f satisfying the conditions imposed in definition 3.20.*

Proof. We begin proving the first part of the theorem. Note that $f(|\mathbf{x}|)$ is invariant under rotations and thus SO(3) only affects the function p. Therefore, $V_{l,f}$ is as a representation of SO(3) isomorphic to the space V_l which is irreducible by theorem 3.11.

The other direction is a bit more involved. Note that the Lebesgue measure on $\mathbb{R}^3$ decomposes as the product of the surface area measure on S^2 with the measure $4\pi r^2 dr$ on $(0, \infty)$, which are both σ-finite[11]. Therefore, $L^2(\mathbb{R}^3)$ decomposes as the Hilbert tensor product of $L^2(S^2)$ and $L^2((0, \infty))$. A vector of the form $f \otimes g$ in the tensor product corresponds to $f\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right) g(|\mathbf{x}|)$ in $L^2(\mathbb{R}^3)$. This is quite convenient as the space $L^2(S^2)$ decomposed as the sum of spaces V_l, $l = 0, 1, 2, 3, \ldots$, just as we saw in theorem 3.11. Hence, we can decompose $L^2(\mathbb{R}^3)$ as the sum of spaces of the form $V_{l,k} := V_l \otimes g_k$. Here $\{g_k\}$ forms an orthonormal basis for $L^2((0, \infty))$.

We assume V to be a finite-dimensional, irreducible, SO(3) invariant subspace of $L^2(\mathbb{R}^3)$, with representation Π. Let $\pi_{l,k} : L^2(\mathbb{R}^3) \to V_{l,k}$ be the orthogonal projection operator and let $\rho_{l,k}$ be the restriction of $\pi_{l,k}$ to V. This map is easily seen to be an intertwining map for the action of SO(3). We visualise this with a figure.

$$\begin{array}{ccc} V & \xrightarrow{\rho_{l,k}} & V_{l,k} \\ \Pi(g) \downarrow & & \downarrow \Pi(g) \\ V & \xrightarrow[\rho_{l,k}]{} & V_{l,k} \end{array}$$

Hence, since V and $V_{l,k}$ are irreducible we see that either $\rho_{l,k}$ is an isomoprhism or zero by Schur's Lemma. Furthermore, we have seen that $V_{l,k}$'s are nonisomorphic for different values of l. Therefore, we know that $\rho_{l.k}$ and $\rho_{l',k'}$ cannot both be nonzero for $l \neq l'$. On the other hand , $\rho_{l.k}$ cannot be non zero for all values of k, l as $V_{k,l}$ span $L^2(\mathbb{R}^3)$. Hence, there must be some ρ_{l_0,k_0} for some $l_0 \neq 0$ such that ρ_{l_0,k_0} is nonzero for some k_0, but $\rho_{l,k} = 0$ for $l \neq l_0$. We use Schur's lemma one more time to note that $\rho_{l_0,k}(\rho_{l_0,k_0})^{-1}$ must be of the form $c_k I$ for each k. Let $\phi \in V$ be arbtrary, then let v be the unique element of V such that $\rho_{l_0,k_0}(\phi) = v \otimes g_{k_0}$. We then have that

$$\rho_{l_0,k}(\phi) = c_k(v \otimes g_k)$$

for every k. Since we also know that $\rho_{l,k}(\phi) = 0$ for every $l \neq l_0$, we know that ϕ us of the form $v \otimes g$, where $g = \sum_k c_k g_k$. Since this holds for all $\phi \in V$ we see that $V = V_{l_0} \otimes g$, which is exactly the form $p(\mathbf{x})f(|\mathbf{x}|)$ as in definition 3.20. We see that V is of the form claimed in the proposition. □

3.6. Spin

In this section, we introduce the physics behind the (irreducible) representations of so(3) and SO(3) we have derived. This section is all about the bridge between representation theory and physics. We already gave a definition of spin in the mathematical sense, however in physics, spin is usually seen as a number which describes an internal degree of freedom of a particle in a Hilbert space.

In the past few sections, we classified the irreducible finite-dimensional representations of the Lie algebra so(3) by their "spin" l. However, we noted that inside $L^2(S^2)$ and $L^2(\mathbb{R}^3)$ we only found irreducible representations of so(3) with integer spin. This is quite curious as in physics, spin is a quantum number with half integer values as well.

From a mathematical viewpoint, it is quite easy to see why this is the case: half integer spin representations of so(3) cannot be exponentiated to any representations of the group SO(3) as we have seen in the previous sections. They can be exponentiated to the the universal cover of SO(3): SU(2) as we have seen in the proof of proposition 3.8. However, we have seen that for a half integer spin representation Π'_l of SU(2) it holds that $\Pi'(-I) = -I$, which means that Π'_l does not factor through $SO(3) \cong SU(2)/\{I, -I\}$.

This may seem quite unfortunate, but the notion of the projective representation introduced in section 2.7.2 will help us out. As we have seen from from the first axiom of quantum mechanics in chapter 1, states that differ up to a phase factor describe the same physical state of a particle. Hence, we turn to the unitary representations introduced in chapter 2.

When we consider a projective unitary representations as defined in section 2.7.2, we see that $[-I]$ is just the identity element in PU(V). Hence, even when l is half integer, we obtain a well defined projective representation Π_l of SO(3), that is given by

$$\Pi_l\left(e^{tX}\right) = \left[e^{t\pi_l(X)}\right]$$

for all $X \in so(3)$. Note that this is in accordance with theorem 2.47. It says, applied to our case, that every irreducible ordinary representation of the Lie algebra gives rise to a representation of the universal cover SU(2) of SO(3), which in turn gives rise to a projective representation of SO(3).

The physics of the universe is invariant under SO(3), or a least that is generally believed. In quantum mechanics, rotational symmetry means that there should be a projective unitary representation of SO(3) on the Hilbert space of the universe that commutes with the Hamiltonian operator [11]. However, it is quite uncertain if there even exists such a Hilbert space of the universe, but assuming there is, it must be composed of Hilbert spaces of each type of particle.

Therefore, we expect that the Hilbert space of each particle will carry a projective unitary action of SO(3). One can describe the Hilbert space of a particle is several ways, but one of the simplest is to describe it with $L^2(\mathbb{R}^3)$ (as discussed earlier, this carries a unitary action of SO(3)). However, we now assign some internal degrees of freedom to each type of particle, known as spin. This suggests a modification of the quantum Hilbert space of each type of particle in the following form:

$$L^2\left(\mathbb{R}^3\right) \hat{\otimes} V,$$

where V is a finite-dimensional Hilbert space that carries an irreducible representation of SO(3). From proposition 2.46 we know that the space V carries an ordinary representation π of the Lie algebra so(3). The largest eigenvalue l of $L_3 = i\pi(F_3)$ is called the spin of the particle and we can label the corresponding space by V_l. For example, electrons have spin $\frac{1}{2}$ and are called fermions while protons for example have integer spin and are called bosons. Therefore, a single electron has Hilbert space $L^2(\mathbb{R}^3)\hat{\otimes} V_{\frac{1}{2}}$, $V_{\frac{1}{2}}$ is a two dimensional projective representation of SO(3).

We can think of $L^2(\mathbb{R}^3)$ as carrying a unitary representation of SU(2) that factors through SO(3), thus $\Pi(-I) = I$. While, V_l carries a unitary representation Π_l of SU(2) such that $\Pi_l(-I) = \pm I$ depending on whether l is integer or half-integer. Hence, $L^2(\mathbb{R}^3)\otimes V_l$ carries a unitary representation $\Pi\otimes\Pi_l$ of SU(2) in which $\Pi\otimes\Pi_l(-I) = \pm I$. Hence, this representation factors through SO(3) in a projective sense. For a treatment of spin in a more algebraic topological way, see Appendix A. I must warn the reader, prior knowledge concerning algebraic topology is required.

3.7. Addition of Angular Momentum

Let V_l and V_m be irreducible representations of so(3), dimensions $2l+1$ and $2m+1$. Note that $V_l\otimes V_m$ is another representation of so(3) and note that unless m or l is equal to zero, $V_l\otimes V_m$ is not irreducible! Assume that V_l is a SO(3) invariant subspace of $L^2(\mathbb{R}^3)$ and V_m is the space which represents the internal degrees of freedom of a particle. It is then interesting to decompose this space $V_l\otimes V_m$ into irreducible SO(3) invariant subspaces.

Proposition 3.23. *Let $V_{\frac{1}{2}}$ be the irreducible representation of so(3) of dimension 2, and let V_l be an irreducible representation of so(3) of dimension $2l+1$, where l is a nonnegative integer or half integer. If $l=0$, $V_l\otimes V_{\frac{1}{2}}$ is irreducible. If $l>0$, then we have*

$$V_l\otimes V_{\frac{1}{2}} \cong V_{l+\frac{1}{2}}\oplus V_{l-\frac{1}{2}}$$

where $\cong$ denotes an isomorphism of representations.

Proof. If $l=0$ this is trivial. Therefore, assume $l>0$. Let L^+, L^- and L_3 be the operators as proposed in theorem 3.2 using the representation π_l. Let $\sigma^+,\sigma^-,\sigma_3$ be the analogous operators constructed using the representation $\pi_{\frac{1}{2}}$. As we saw in section 2.7.4, we can define operators J^+, J^-, J_3 on $V_l\otimes V_{\frac{1}{2}}$ by

$$J^+ = L^+\otimes I + I\otimes\sigma^+$$
$$J^- = L^-\otimes I + I\otimes\sigma^-$$
$$J_3 = L_3\otimes I + I\otimes\sigma_3.$$

Let $\{v_0,\dots,v_{2l}\}$ be a basis as in theorem 3.2 and let e_0, e_1 be a similar basis for $V_{\frac{1}{2}}$. We can then easily see that the vectors $v_j\otimes e_k$ for a basis for the space $V_l\otimes V_{\frac{1}{2}}$. Note that the eigenvalues for these eigenvectors of J_3 are of the form

$$(l-j)+(\frac{1}{2}-k)$$

for $j=0,1,\dots,2l$, $k=0,1$. The the eigenvalues for J_3 range from $l+\frac{1}{2}$ to $-(l+\frac{1}{2})$. All the eigenvalues occur as a combination of l and k twice, as $(\lambda+\frac{1}{2})-\frac{1}{2}$ and $(\lambda-\frac{1}{2})+\frac{1}{2}$, but $l+\frac{1}{2}$ and $-(l+\frac{1}{2})$ occur as eigenvalues only once.

Note that the vector $v_0\otimes e_0$ is an eigenvector for J_3 with the largest eigenvalue of $l+\frac{1}{2}$. Hence, $J^+(v_0\otimes e_0)=0$. By proposition 3.5, when we apply J^- repeatedly, we obtain a "chain" of eigenvectors of length $2l+2$. The span of these vectors form an irreducible invariant subspace W_0 isomorphic to $V_{l+\frac{1}{2}}$.

By proposition 3.6 there is an inner product on V_l and $V_{\frac{1}{2}}$ such that π_l and $\pi_{\frac{1}{2}}$ are skew-adjoint: $\pi(X)^* = -\pi(X)$ for all $X\in$ so(3). If we use on $V_l\otimes V_{\frac{1}{2}}$ the inner product as in proposition 2.52, we see that $\pi_l \otimes \pi_{\frac{1}{2}}$ is unitary as well. Also, this is quite trivial, but, again, for completeness sake we include it. We essentially have

$$\begin{aligned}((\pi_l \otimes \pi_{\frac{1}{2}})(X))^* &= (\pi_l(X)^* \otimes I + I \otimes \pi_{\frac{1}{2}}(X)^*) \\ &= -(\pi_l(X) \otimes I + I \otimes \pi_{\frac{1}{2}}(X)) \\ &= -((\pi_l \otimes \pi_{\frac{1}{2}})(X)).\end{aligned}$$

Hence, the orthogonal complement of the invariant subspace W_0 is invariant as well, see the proof of lemma 2.54. Since all eigenvalues for J_3 except the largest and smallest have multiplicity 2, we see that the largest eigenvalue for J_3 in $W_0^{\perp}$ is $l - \frac{1}{2}$. In the same fashion as before, we can apply J^- repeatedly to obtain a chain of eigenvectors with length 2l. These eigenvectors span an irreducible subspace of $V_l \otimes V_{\frac{1}{2}}$ of dimension $2l$. Since

$$\dim(W_0) + \dim(W_1) = 4l + 2 = \dim(V_l \otimes V_{\frac{1}{2}}),$$

we have that $W_0 = W_1^{\perp}$. □

In the case that V_l is an SO(3) invariant subspace of $L^2(\mathbb{R}^3)$, the formula for the operator J in the proof of the previous theorem is written as $J_3 = L_3 + \sigma_3$ in physics literature. Here, L_3 acts on the first factor in the tensor product and σ_3 on the second one. Suppose we have a spin $\frac{1}{2}$ particle, such as an electron, then the L_3 is the angular momentum operator and σ_3 describes the action of $F_3 \in \mathrm{so}(3)$ on the space $V_{\frac{1}{2}}$. The formula $J_3 = L_3 + \sigma_3$ account for the physics terminology "addition of angular momentum" to describe the tensor products of representations of so(3) [11]. Here, L_3 is the orbital angular momentum and σ_3 is called the spin angular momentum.

The general result of irreducible representations of tensor products is stated here, but will not be proven in this report. The proof is similar to that of the previous theorem, it can also be found in [9]. Note that this is a beautiful example of complete reducibility, see the last section 2.8 of chapter 2.

Proposition 3.24. *For any $j = 0, \frac{1}{2}, 1, \ldots$ let V_j denote the unique irreducible representation of so(3) of dimension $2j+1$. Then for any l and m with $l \geq m$, we have*

$$V_l \otimes V_m \cong V_{l+m} \oplus V_{l+m-1} \oplus \ldots \oplus V_{l-m+1} \oplus V_{l-m}.$$

When we have two spin l and m particles and we wish to add these two spins, the result we could obtain spin $l+m$ all down to spin $|l-m|$[8]. This last proposition is equivalent to this result, only written in a more mathematical way. With this last observation, we conclude chapter 3.

4

Representation Theory for SU(3)

This chapter is mostly based on chapter 6 and section 4.6 of [10] and for most of this chapter we will follow these sections closely. Until we have only looked at representations of SU(2). However, when we look at representations of SU(3), we will see some generalizations happening in comparison to the SU(2) case. We will introduce all the necessary tools and concepts form a mathematical viewpoint in order to understand the links between SU(3) and the meson/octet and baryon/decimet model. I tried to introduce all the definitions in this chapter as general as possible.

Every finite-dimensional representation of SU(3) over a complex vector space, gives rise to a representation of su(3). By complex linearity, we know that a representation of su(3) can be extended to $\mathrm{su}(3)_{\mathbb{C}} \cong \mathrm{sl}(3;\mathbb{C})$. We can go the other direction as well. If we restrict a representation of $\mathrm{sl}(3;\mathbb{C})$ to su(3) and note that SU(3) is simply connected[10], we can apply theorem 2.29 and obtain a representation of SU(3) by exponentiation.

We can even establish a one-to-one correspondence between the irreducible representations of SU(3) and the irreducible representations of $\mathrm{sl}(3;\mathbb{C})$. For this, we need an additional theorem.

Theorem 4.1. *Let $\mathfrak{g}$ be a real Lie algebra and $\mathfrak{g}_{\mathbb{C}}$ 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. Furthermore, π is irreducible as a representation of $\mathfrak{g}_{\mathbb{C}}$ if and only if it is irreducible as a representation of $\mathfrak{g}$.*

Proof. The unique extension of π to $\mathfrak{g}_{\mathbb{C}}$ is given by $\pi(X+iY) = \pi(X) + i\pi(Y)$ for all $X, Y \in \mathfrak{g}$. By some trivial calculations, we can see that this is indeed a complex Lie algebra homomorphism.

The claim of irreducibility holds because a complex subspace W of V is invariant under $\pi(X+iY)$, with X and Y in $\mathfrak{g}$, if and only if it is invariant under the operators $\pi(X)$ and $\pi(Y)$. Hence, the representation of $\mathfrak{g}$ and its extension to $\mathfrak{g}_{\mathbb{C}}$ have the same invariant subspaces. □

This theorem together with theorem 2.39 tell us that a representation of SU(3) is irreducible if and only if the associated representation of $\mathrm{sl}(3;\mathbb{C})$ is irreducible. Moreover, as SU(3) is compact, we see that every finite-dimensional representation of SU(3), and hence also of $\mathrm{sl}(3;\mathbb{C})$, is completely reducible. This follows from the last two theorems of section 2.8. Therefore, classifying the irreducible representations of $\mathrm{sl}(3;\mathbb{C})$ is enough to study the general representations of $\mathrm{sl}(3;\mathbb{C})$. After we have classified the irreducible representations of $\mathrm{sl}(3;\mathbb{C})$, we can pass the results over to SU(3). This will be our strategy for this chapter.

4.1. Representation Theory for sl(2;ℂ)

In this section we closely follow parts of section 4.6 of [10]. Representation theory for $\mathrm{sl}(3,\mathbb{C})$ is actually just an extension of representation theory for $\mathrm{sl}(2,\mathbb{C})$. Therefore, it is essential to first understand representation theory for $\mathrm{sl}(2;\mathbb{C})$. Recall from example 2.8 $\mathrm{sl}(2,\mathbb{C})$ has a basis H, X, Y satisfying

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

We are going to start with a theorem which is somewhat analogous to theorem 3.2. Before we do this, we need an additional lemma.

Lemma 4.2. *Let π be a finite dimensional representation of $sl(2;\mathbb{C})$. Let u be an eigenvector of $\pi(H)$ with eigenvalue $\alpha\in\mathbb{C}$. Then we have*

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

Thus, either $\pi(X)u=0$ or $\pi(X)u$ is an eigenvector for $\pi(H)$ with eigenvalue $\alpha+2$. Similarly,

$$\pi(H)\pi(Y)u=(\alpha-2)\pi(Y)u,$$

so that either $\pi(Y)u=0$ or $\pi(Y)u$ is an eigenvector for $\pi(H)$ with eigevalue $\alpha-2$.

Proof. We know that $[\pi(H),\pi(X)]=\pi([H,X])=2\pi(X)$. Hence,

$$\begin{aligned}\pi(H)\pi(X)u&=\pi(X)\pi(H)u+2\pi(X)u\\&=\pi(X)(\alpha u)+2\pi(X)u\\&=(\alpha+2)\pi(X)u.\end{aligned}$$

The case for $\pi(X)$ replaces by $\pi(Y)$ is similar. □

Theorem 4.3. *For each integer $m\geq 0$ there is an irreducible complex representation of $sl(2;\mathbb{C})$ with dimension $m+1$. Any two irreducible complex representations of $sl(2;\mathbb{C})$ with the same dimension are isomorphic.*

Proof. This proof goes almost analogous to what we did in the proof of theorem 3.2, since $\mathrm{so}(3)_{\mathbb{C}}\cong\mathrm{sl}(2;\mathbb{C})$. Let π be an irreducible representation of $\mathrm{sl}(2;\mathbb{C})$ acting on a finite-dimensional complex vector space V. Since we are working in $\mathbb{C}$, $\pi(H)$ must have at least one eigenvector. Let u be the eigenvector of $\pi(H)$ with eigenvalue α. Applying the previous lemma 4.2 repeatedly, we see that

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

Since an operator on a finite-dimensional space can have only finitely many eigenvalues, the $\pi(X)^k u$'s cannot be all non-zero. Hence, there exists some $N\in\mathbb{N}$ such that

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

but

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

Then, analogous to what we did in the proof of theorem 3.2, we set $u_0=\pi(X)^N u$ and $\lambda=\alpha+2N$. We obtain

$$\pi(H)u_0=\lambda u_0, \text{ and } \pi(X)u_0=0.$$

Now we can define

$$u_k=\pi(Y)^k u_0$$

for $k\geq 0$. Again, by the previous lemma 4.2

$$\pi(H)u_k=(\lambda-2k)u_k.$$

By induction, we can see that

$$\pi(X)u_k = k[\lambda - (k-1)]u_{k-1} \quad (k \geq 1).$$

Since $\pi(H)$ can hace only finitely manby eigenvalues, the u_k's cannot all be non-zero. There must be a non-negative integer m such that

$$u_k = \pi(Y)^k u_0 \neq 0$$

for all $k \leq m$, however,

$$u_{m+1} = \pi(Y)^{m+1} u_0 = 0.$$

If $u_{m+1} = 0$, then $\pi(X)u_{m+1} = 0$ and so, we obtain

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

Since u_m and $m+1$ are non-zero, we know that $\lambda - m = 0$. Hence, $\lambda = m$. Hence, for every irreducible representation (π, V), there exists an integer $m \geq 0$ and nonzero vectors $u_0, ..., u_m$ such that

$$\begin{aligned}
\pi(H)u_k &= (m-2k)u_k, \\
\pi(Y)u_k &= \begin{cases} u_{k+1} & \text{if } k < m \\ 0 & \text{if } k = m \end{cases}, \qquad (*) \\
\pi(X)u_k &= \begin{cases} k(m-(k-1)u_{k-1} & \text{if } k > 0 \\ 0 & \text{if } k = 0 \end{cases}.
\end{aligned}$$

Again, one can note the similarities between this and theorem 3.2. Note that the eigenvalues of the eigevectors $u_0, ..., u_m$ of $\pi(H)$ are all independent. Therefore, the vectors $u_0, ..., u_m$ are all linearly independent. Moreover, the span of $u_0, ..., u_m$ is invariant under $\pi(H), \pi(X)$ and $\pi(Y)$. Hence, we can conclude that the span of $u_0, ..., u_m$ is invariant under $\pi(Z)$ with $Z \in \mathrm{sl}(2;\mathbb{C})$. Since π is irreducible, this space must be all of V. We showed that every irreducible representation of $\mathrm{sl}(2;\mathbb{C})$ is of the form of (*). Conversely, if we define $\pi(H), \pi(X)$ and $\pi(Y)$ by (*), the u_k's are basis elements for some $m+1$ dimensional vector space, then the operators defined as in (*) really do satisfy the commutation relations of $\mathrm{sl}(2;\mathbb{C})$ and is irreducible. This goes similarly to what we did in the proof of theorem 3.4 and is therefore omitted. In conclusion, every irreducible representation of dimension $m+1$ must have the form of (*), which shows that any two such representations are isomorphic. □

The following theorem is at the heart of this section and will be used frequently in our analysis of $\mathrm{sl}(3;\mathbb{C})$.

Theorem 4.4. *If (π, V) is a finite-dimensional representation of $sl(2;\mathbb{C})$, not necessarily irreducible, the following result holds. Every eigenvalue of $\pi(H)$ is an integer. Furthermore, if v is an eigenvector for $\pi(H)$ with eigenvalue λ and $\pi(X)v = 0$, then λ is a non-negative integer.*

Proof. Supppose that v is an eigenvector of $\pi(H)$ with eigenvalue λ. Then there is some $N \geq 0$ such that $\pi(X)^N v \neq 0$ but $\pi(X)^{N+1} v = 0$, where $\pi(X)^N v$ is an eigenvector of $\pi(H)$ with eigenvalue $\lambda + 2N$. From the previous proof, we can see that $m := \lambda + 2N$ must be an non-negative integer, so that λ is an integer. If $\pi(X)v = 0$ then we take $N = 0$ and $\lambda = m$ is non-negative. □

4.2. Gell-Mann Matrices and Commutor Relations

First of all, we are going to introduce the basis of $\mathrm{sl}(3;\mathbb{C})$ which was introduced by Gell-Mann and is usually referred to as the Gell-Mann matrices.

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$
$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \lambda_8 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

This basis is not the most friendly to execute calculations with, so we will introduce a new basis based on the Gell-Mann matrices which is dubbed the Dynkin basis. We set $T_i = \frac{1}{2}\lambda_i$ and introduce matrices

$$\begin{aligned} T_\pm &= T_1 \pm iT_2 \\ V_\pm &= T_4 \pm iT_5 \\ U_\pm &= T_6 \pm iT_7, \end{aligned}$$

We also introduce $H_1 = \lambda_3$ and $H_2 = \frac{\sqrt{3}\lambda_8 - \lambda_3}{2}$. This might seem a bit arbitrary, but there is a specific reason why we choose H_1 and H_2 in this way. The reason is that they make calculations later much easier. The span of $\langle H_1, T_\pm\rangle$, $\langle H_2, U_\pm\rangle$ and $\langle \frac{\sqrt{3}\lambda_8+\lambda_3}{2}, V_\pm\rangle$ are subalgebras of $\mathrm{sl}(3;\mathbb{C})$ isomorphic to $\mathrm{sl}(2;\mathbb{C})$. To see this, one can take a look at the matrices $H_1, H_2, T_\pm, V_\pm$ and $U_\pm$.

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad H_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad T_+ = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$T_- = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad V_+ = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad V_- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$
$$U_+ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad U_- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

The important commutation relations are as follows

$$\begin{aligned}
&[H_1, T_+] = 2T_+, && [H_1, V_+] = V_+, && [H_1, U_+] = -U_+, \\
&[H_1, T_-] = -2T_-, && [H_1, V_-] = -V_-, && [H_1, U_-] = U_-, \\
&[H_2, T_+] = -T_+, && [H_2, V_+] = V_+, && [H_2, U_+] = 2U_+, \\
&[H_2, T_-] = T_-, && [H_2, V_-] = -V_-, && [H_2, U_-] = -2U_-, \\
&[T_+, T_-] = H_1 && [V_+, V_-] = \frac{\sqrt{3}\lambda_8 + \lambda_3}{2} && [U_+, U_-] = H_2.
\end{aligned}$$

As a last remark, note that H_1 and H_2 commute, thus $\pi(H_1), \pi(H_2)$ commute (for any representation π of $\mathrm{sl}(3;\mathbb{C})$). Therefore, there is a chance that $\pi(H_1)$ and $\pi(H_2)$ can be simultaneously diagonalised, which is possible as we will see.

4.3. Weights and Roots

Two concepts have to be introduced which we did not yet have in the analysis of the SU(2) representations: the notions of weights and roots. We did not need these notions at that time, since they are trivial in the case for SU(2). However, they play an important role for SU(3). We are first giving the general definition of weight and later specifying it to the case which is of interest to us. We apply the same strategy for roots. The general definitions of weights and roots are basis independent, which will be of use for us later. For this general definition of weights and roots, we first need to specify an inner product of a general Lie algebra $\mathfrak{h}$ and introduce the notion of a Cartan subalgebra.

As we said, we need an inner product on a general Lie algebra $\mathfrak{h}$. The inner product we are hinting at is given by $\langle H, H' \rangle = \mathrm{Tr}(H^* H')$ for $H, H' \in \mathfrak{h}$. This is called the Hilbert-Schmidt inner product. We still need to check that this is indeed an inner product. Note that this expression is linear in the second factor. For conjugate symmetry, note that

$$\langle A, B \rangle = \mathrm{Tr}(A^* B) = \sum_{i,j=1}^{3} A^*_{kl} B_{lk} = \overline{\sum_{i,j=1}^{3} B^*_{lk} A_{kl}} = \overline{\langle B, A \rangle}.$$

Finally, we can see that positivity is satisfied as

$$\langle A, A \rangle = \mathrm{Tr}(A^* A) = \sum_{k,l=1}^{n} A^*_{kl} A_{lk} = \sum_{k,l=1}^{n} |A_{kl}|^2 \geq 0,$$

and the sum is zero if and only if every entry of A is zero. Now we have established an inner product, we proceed by introducing the notion of a Cartan subalgebra.

Definition 4.5. *A Cartan subalgebra of* $sl(3;\mathbb{C})$ *is a complex subspace* $\mathfrak{h}$ *of* $sl(3;\mathbb{C})$ *with the following properties:*

1. *For all* A_1 *and* A_2 *in* $\mathfrak{h}$, $[A_1, A_2] = 0$.
2. *If, for some* $X \in \mathfrak{g}$, *we have* $[A, X] = 0$ *for all* A *in* $\mathfrak{h}$, *then* X *is in* $\mathfrak{h}$.
3. *For all* A *in* $\mathfrak{h}$, ad_A *is diagonalisable.*

The dimension of $\mathfrak{h}$ *is called the rank of* $sl(3;\mathbb{C})$.

Condition one says that $\mathfrak{h}$ is a commutative subalgebra of $\mathfrak{sl}(3;\mathbb{C})$. Condition two says that it is the maximal commutative subalgebra. Condition three is obvious.

For our purposes, we would like to construct a Cartan subalgebra of sl$(3;\mathbb{C})$. The first two properties are enough to construct such an algebra of sl$(3;\mathbb{C})$[10]. We could do this by looking at diagonal matrices of sl$(3;\mathbb{C})$. For example, the matrices H_1 and H_2 satisfy the first two properties of a Cartan subalgebra of sl$(3;\mathbb{C})$, as can be verified by looking at the commutator relations of section 4.2 and noting that $[H_1, H_2] = 0$. We now introduce the definition of a weight.

Definition 4.6. *Let* $\mathfrak{h}$ *be the Cartan subalgebra of* $sl(3;\mathbb{C})$ *spanned by* H_1 *and* H_2 *and let* (π, V) *be a representation of* $sl(3;\mathbb{C})$. *An element* μ *of* $\mathfrak{h}$ *is called a weight for* π *if there exists a nonzero vector* v *in* V *such that*

$$\pi(H)v = \langle \mu, H \rangle v$$

for all $H \in \mathfrak{h}$. *The weight space corresponding to* μ *is the set of all* $v \in V$ *satisfying the above equation. The multiplicity of* μ *is the dimension of the corresponding weight space.*

Remember, for our analysis we take $\mathfrak{h}$ to be the space that is spanned by H_1 and H_2. Let μ be a weight for sl$(3;\mathbb{C})$ as defined above. Then, if (π, V) is a representation of sl$(3;\mathbb{C})$, a weight can be considered an ordered pair $(m_1, m_2) \in \mathbb{C}^2$ such that $m_1 = \langle \mu, H_1 \rangle$ and $m_2 = \langle \mu, H_2 \rangle$ for an $v \in V$. Note that in this way, a weight is just a pair of simultaneous eigenvalues for $\pi(H_1)$ and $\pi(H_2)$. It is of interest to know whether an arbitrary representation of sl$(3;\mathbb{C})$ has a weight or not.

Proposition 4.7. *Every representation (π, V) of $sl(3;\mathbb{C})$ has at least one weight.*

Proof. Since we are working over the complex numbers we see that $\pi(H_1)$ has at least one eigenvalue $m_1 \in \mathbb{C}$. Let $W \subset W$ be the eigenspace for $\pi(H_1)$ with eigenvalue m_1. Since $[H_1, H_2] = 0$ we know that $\pi(H_1)$ and $\pi(H_2)$ commute. Therefore, we know that $\pi(H_2)$ has to map W into itself. Then, we know that the restriction of $\pi(H_2)$ to W must have at least one eigenvector w with eigenvalue $m_2 \in \mathbb{C}$. Therefore, w is a simultaneous eigenvector of $\pi(H_1)$ and $\pi(H_2)$ with eigenvalues m_1 and m_2. □

Every representation (π, V) of $\mathrm{sl}(3;\mathbb{C})$ can be viewed, by restriction, as representations of the subalgebras $\langle H_1, T_\pm\rangle$ and $\langle H_2, U_\pm\rangle$. As noted before in section 4.2, these are isomorphic to $\mathrm{sl}(2;\mathbb{C})$.

Proposition 4.8. *If (π, V) is representation of $sl(3;\mathbb{C})$ and $(m_1, m_2) \in \mathbb{C}^2$ is a weight of V, then both m_1 and m_2 are integers.*

Proof. This follows quite easily from theorem 4.4. We just have to take the restriction of π to $\langle H_1, T_\pm\rangle$ and $\langle H_2, U_\pm\rangle$ and apply theorem 4.4. □

We have defined an inner product and we have introduced the notion of weight with some supplementary propositions. We proceed with the definition of root.

Definition 4.9. *A nonzero element α of $\mathfrak{h}$, a Cartan subalgebra of $\mathfrak{g}$, is a root if there exists a nonzero $X \in \mathfrak{g}$ such that*

$$[H, X] = \langle \alpha, H\rangle X$$

for all H in $\mathfrak{h}$. The set of all roots is denoted R. The element X is called a root vector corresponding to the root α.

Just like with weights, it is good to take a look what these definitions mean in the $\mathrm{sl}(3;\mathbb{C})$ case. When we have a root $\alpha \in \mathfrak{h}$ we know that $[H, X] = \langle\alpha, H\rangle X$ for all $H \in \mathfrak{h}$. However, $\mathfrak{h}$ is spanned by H_1 and H_2, hence for $\alpha \in \mathfrak{h}$ to be a root there has to exists a nonzero $X \in \mathrm{sl}(3;\mathbb{C})$ such that $[H_1, X] = \langle\alpha, H_1\rangle X$ and $[H_2, X] = \langle\alpha, H_2\rangle X$. Hence, we can express a root as a tuple, $(a_1, a_2) \in \mathbb{C}^2$ as well where $a_1 = \langle\alpha, H_1\rangle$ and $a_2 = \langle\alpha, H_2\rangle$. When executing calculations, it is often more convenient to think of roots in this more concrete fashion than as an element in $\mathfrak{h}$.

For $\mathrm{sl}(3;\mathbb{C})$ we have the following set of roots.

(a_1, a_2)	X	(a_1, a_2)	X
$(2, -1)$	T_+	$(-2, 1)$	T_-
$(-1, 2)$	U_+	$(1, -2)$	U_-
$(1, 1)$	V_+	$(-1, -1)$	V_-

Note that root vectors are eigenvectros of ad_{H_1} and ad_{H_2}. H_1 and H_2 are also simultaneous eigenvectors of ad_{H_1} and ad_{H_2}, but they are not root vectors since the simultaneous eigenvalues are zero. Note that we can single out the roots $\beta_1 = (2, 1)$ and $\beta_2 = (-1, 2)$ corresponding to T_+ and U_+ and note that all the other roots can be expressed as linear combinations of these two roots with integer coefficients.

$$(2, -1) = \beta_1; \qquad (-1, 2) = \beta_2; \qquad (1, 1) = \beta_1 + \beta_2,$$

with the remaining roots being the negatives of the above. The choice for β_1 and β_2 is arbitrary. The importance of roots for representation of $\mathrm{sl}(3;\mathbb{C})$ is expressed by the following lemma.

Lemma 4.10. *Let $\alpha = (a_1, a_2) \in \mathbb{C}^2$ be a root and let $Z_\alpha \in sl(3;\mathbb{C})$ be a corresponding root vector. Let (π, V) be a representation of $sl(3;\mathbb{C})$, let $\mu = (m_1, m_2)$ be a weight for π, and let $v \neq 0$ be a corresponding weight vector. Then we have*

$$\pi(H_1)\pi(Z_\alpha)v = (m_1 + a_1)\pi(Z_\alpha)v$$
$$\pi(H_2)\pi(Z_\alpha)v = (m_2 + a_2)\pi(Z_\alpha)v$$

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. Since Z_α is a root, we now by definition that $[H_1, Z_\alpha] = a_1 Z_\alpha$. Hence, we obtain the relations

$$\begin{aligned}\pi(H_1)\pi(Z_\alpha)v &= (\pi(Z_\alpha)\pi(H_1) + a_1\pi(Z_\alpha))v \\ &= \pi(Z_\alpha)(m_1 v) + a_1\pi(Z_\alpha)v \\ &= (m_1 + a_1)\pi(Z_\alpha)v.\end{aligned}$$

A similar argument can be used to compute $\pi(H_2)\pi(Z_\alpha)$. □

As a last remark, note that this lemma is analogous to lemma 4.2.

4.4. Highest Weight Representation

For the remainder of chapter 4, we closely follow section 6.4, 6.6 and 6.7 of [10]. Just as with the representations of $\mathrm{su}(2)$ and $\mathrm{sl}(2;\mathbb{C})$, we would like to label the representation of $\mathrm{sl}(3;\mathbb{C})$. We do this with the notion of a highest weight.

Definition 4.11. *Let (π, V) be a representation of $sl(3;\mathbb{C})$. Let $\beta = (2,-1)$ and $\beta = (-1,2)$ and be the roots introduced in the previous section . Let μ_1 and μ_2 be two weights of π. Then μ_1 is lower than μ_2 if $\mu_1 - \mu_2$ can be written in the form*

$$\mu_1 - \mu_2 = a\beta_1 + b\beta_2$$

with $a > 0$ and $b \geq 0$. This relationship is a partial ordering on the set of all weights of π and is written as $\mu_1 \preceq \mu_2$. A weight μ_0 for π is said to be a highest weight if for all weights μ of π, $\mu \preceq \mu_0$.

The theorem stated below is at the heart of this section and is called the theorem of the highest weight. This is theorem tells us how to classify all the finite dimensional irreducible representations of $\mathrm{sl}(3;\mathbb{C})$.

Theorem 4.12. *The theorem consists of several statements.*

1. *Every irreducible representation π of $sl(3;\mathbb{C})$ is the direct sum of its weights spaces.*
2. *Every irreducible representation of $sl(3;\mathbb{C})$ has a unique highest weight μ.*
3. *Two irreducible representations of $sl(3;\mathbb{C})$ 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 $sl(3;\mathbb{C})$ with highest weight (m_1, m_2).*

The proof of this theorem exists of a string of propositions. We warn the reader, the remainder of this section is fully dedicated to proving the theorem of the highest weight. If the reader is interested in the proofs of this theorem, we invite them to read the next section. If not, the reader may as well take this theorem for granted, and straight away skip to the last theorem of this section. In my personal opinion, the proofs of the individual points of this theorem are quite fun.

Proposition 4.13. *In every irreducible representation (π, V) of $sl(3;\mathbb{C})$, the operators $\pi(H_1)$ and $\pi(H_2)$ can be simultaneaously diagonalised. Thus, V is the direct sum of its weight spaces.*

Proof. Let W be the sum of the weight spaces in V. Or in other words, W is the space of all vectors $w \in V$ such that w can be written as a linear combination of simultaneous eigenvectors for $\pi(H_1)$ and $\pi(H_2)$. By proposition 4.7 we know that π has at least one weight, $w \neq 0$. On the other hand, lemma 4.2 tells us that if Z_α is a root vector corresponding to the root $\alpha \in \mathbb{C}^2$, then $\pi(Z_\alpha)$ maps the weight space of $\mu \in \mathbb{C}^2$ into the weight space corresponding to $\mu+\alpha$. Thus, W is invariant under the action of each of the root vectors $T_\pm, V_\pm, U_\pm$. Since W is also invariant under the action of H_1 and H_2, we know that W is invariant under all of sl$(3;\mathbb{C})$. Thus, by irreducibility, $W = V$. Weight vector with distinct weights are independent, a proof can be found [10], so is V actually the direct sum of its weight spaces. □

We will now proof the second point of the theorem of the highest weight.

Definition 4.14. *A representation (π, V) of $\mathfrak{g}$ is highest weight cyclic with highest weight $\mu \in \mathfrak{h}$ if there exists a nonzero vector $v \in V$ such that*

1. $\pi(V)v = \langle \mu, H\rangle v$ *for all* $H \in \mathfrak{h}$.
2. $\pi(X)v = 0$ *for all* $X \in \{T_+, V_+, U_+\}$.
3. *The smallest invariant subspace containing v is V.*

Proposition 4.15. *Every irreducible representation of $sl(3;\mathbb{C})$ is a highest weight cyclic representation, with a unique highest weight $\mu \in \mathbb{C}^2$.*

Proof. It has already been shown that every irreducible representation (π, V) is the direct sum of its weight spaces. Since we have a finite dimensional representation, there can be only a finite number of weights, so their is a maximal weight μ by the Lemma of Zorn. Thus for any nonzero weight vector v with weight μ we must have

$$\pi(X)v = 0,$$

with $X \in \{T_+, V_+, U_+\}$. As π is irreducible, the smallest invariant subspace containing v must be the whole space. Hence, all the conditions for being highest weight cyclic have been satisfied. □

We will proceed with the third point of the theorem of highest weight, which told us that two represenations of sl$(3;\mathbb{C})$ with the same highest weight are isomorphic. Before we can proof this, we need four lemmas. The third and fourth lemmas are the most important ones. The third lemma, lemma 4.18, tells us that when (π, V) is a completely reducible representation of sl$(3;\mathbb{C})$ that is also highest weight cyclic, then π is irreducible. The fourth lemma, lemma 4.19, gives us information about the reducibility of invariant subspaces of a completely reducible Lie algebra or Lie group representation.

Lemma 4.16. *Suppose that $\mathfrak{g}$ is any Lie algebra and that π is a representation of $\mathfrak{g}$. Suppose that $X_1, ..., X_m$ is an ordered basis for $\mathfrak{g}$ as a vector space. Then any expression of the form*

$$\pi(X_{j_1})...\pi(X_{j_N})$$

can be expressed as a linear combination of terms of the form

$$\pi(X_m)^{k_m}...\pi(X_1)^{k_1}$$

where each k_1 is a non-negative integer and where $k_1 + ... + k_m \leq N$.

Proof. We are going to tackle this by induction. If $N = 1$ there is nothing to proof. Any expression of the form $\pi(X_j)$ is of the form $\pi(X_j)^{k_j}$ with $k_j = 1$ and all the other k_l are equal to zero. Assume that this result holds for a product of at most N factors, and consider the first expression which is given in the lemma for $N+1$ factors. Then we can see that the last N factors are in the desired form. We can apply the induction hypothesis on these last N factors to obtain the following expression:

$$\pi(X_j)\pi(X_m)^{k_m}...\pi(X_1)^{k_1},$$

with $k_1 + ... + k_m \leq N$.
The strategy is to move the factor $\pi(X_j)$ to the right one step at a time until it is in the right spot. Therefore, whenever there is a $\pi(X_j)\pi(X_k)$ somewhere in the expression we can use the relation

$$\begin{aligned}\pi(X_j)\pi(X_k) &= \pi(X_k)\pi(X_j) + \pi([X_j, X_k]) \\ &= \pi(X_k)\pi(X_j) + \sum_l c_{jkl}\pi(X_l).\end{aligned}$$

The structure constants c_{jkl} are for the basis $\{X_j\}$. Each commutator term has at most N factors. Therefore, we ultimately obtain several terms with N factors, which can be handled by induction, and one term with $N+1$ factors which is in the desired form (once $\pi(X_j)$ is in the right spot). □

Lemma 4.17. *Let (π, V) be a highest weight cyclic representation of $sl(3;\mathbb{C})$ with weight $\mu \in \mathbb{C}^2$. Then the following results hold.*

1. *The representation π has highest weight μ.*
2. *The weight space corresponding to the weight μ is one dimensional.*

Proof. Let v be as in definition 4.14. Then consider the space W of V spanned by elements

$$w = \pi(Y_{j_1})\pi(Y_{j_2})...\pi(Y_{j_N})v$$

with each j_l equal to $1,2$ or 3 and $N \geq 0$. Here, $Y_1 = T_-, Y_2 = V_-, Y_3 = U_-$. We claim that W is invariant. The elements $H_1, H_2, T_\pm, U_\pm, V_\pm$ form a basis of sl(3;$\mathbb{C}$). We now make use of the previous lemma 4.16. When we apply a basis element to an arbitrary vector $w \in W$ we know that we can rewrite it as linear combination of terms in which the $\pi(T_+), \pi(V_+), \pi(U_+)$ act first, the $\pi(H_1), \pi(H_2)$ act second and the $\pi(T_-), \pi(V_-), \pi(U_-)$ act last. Of course, whenever a term contains a positive power of either $\pi(T_+), \pi(V_+), \pi(U_+)$, it is trivially zero. We know that v is an eigenvector of $\pi(H_1)$ and $\pi(H_2)$, we can replace the factors $\pi(H_1)$ and $\pi(H_2)$ by a constant. Hence, we have a linear combination of vectors of the form $\pi(Y_{j_1})\pi(Y_{j_2})...\pi(Y_{j_N})v$ with each j_l equal to $1,2$ or 3 and $N \geq 0$. So, W is invariant and contains v, thus $W = V$. Now, note that T_-, V_-, U_- are root vectors with roots $-\beta, -(\beta_1+\beta_2), -\beta_2$ respectively. Hence, by lemma 4.16 we know that each element of W with $N > 0$ is a weight vector with weight lower than μ. Hence, the only weight vector with weight μ are multiples of v, proving the second point of the theorem. □

Lemma 4.18. *Suppose (π, V) is a completely reducible representation of $sl(3;\mathbb{C})$ that is also highest weight cyclic. Then π is irreducible.*

Proof. Let (π, V) be a highest weight cyclic representation with highest weight μ and let v be a weight vector with weight $\mu \in \mathbb{C}^2$. By assumption, V decomposes as a direct sum of irreducible representations

$$V \cong \bigoplus_j V_j.$$

By proposition 4.13, each of the V_j's is the direct sum of its weight spaces. We know that the weight μ occurs in V and thus must occur in at least one of the V_j. However, by lemma 4.17 we know that v is the only vector in V with weight μ. Hence, V_j is an invariant subspace containing v, thus $V_j = V$. Hence, there is actually only one term occuring in the direct sum of V and its weight spaces: V_j. Therefore, V is irreducible. □

Do not be tempted to think that **any** representation with a cyclic vector is irreducible, that is not what this lemma is telling us. For example, one can take the Lie algebra of the diagonal matrices living in $M_2(\mathbb{C})$, one can check that this is indeed a Lie algebra analogous to example 2.8. We take the standard representation $(\pi, \mathbb{C}^2)$. Then we see that $\begin{pmatrix}1\\1\end{pmatrix}$ is indeed a cyclic vector. However, the subspaces spanned by $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$ are invariant subspaces and hence, the standard representation is not irreducible.

We will proceed with the last of the four consecutive lemmas.

Lemma 4.19. *If V is a completely reducible representation of a group or Lie algebra, then the following properties hold.*

- *For every invariant subspace U of V, there is another invariant subspace W such that V is the direct sum of U and V.*
- *Every invariant subspace of V is completely reducible.*

Proof. The proof is quite fun in its simplicity. For point 1, suppose that V decomposes as

$$V = U_1 \oplus U_2 \oplus \ldots \oplus U_k,$$

where the U_j's are irreducible invariant subspaces, and that U is any invariant subspace of V. If U is all of V, we can take $W = \{0\}$ and we are done. If not, there is some U_{j_1} such that U_{j_1} is not contained in U. Since U_{j_1} is irreducible, it must hold that $U_{j_1} \cap U = \{0\}$. So when $U_{j_1} + U = V$ then we are done, since the sum is direct by $U_{j_1} \cap U = \{0\}$. If $U_{j_1} + U \neq V$, then we can repeat the same steps to eventually obtain some family $j_1, \ldots, j_l$ such that $U + U_{j_1} + \ldots + U_{j_l} = V$ and the desired complement is $W := U_{j_1} + \ldots + U_{j_l}$.

For the second point, suppose that U is an invariant subspace of V. We are going to show that U has the invariant complement property of point one. Suppose that X is another invariant subspace of V with $X \subset U$. By point one, we can find an invariant subspace Y such that $V = X \oplus Y$. Now set $Z = Y \cap U$, which is then an invariant subspace. We want to show that $U = X \oplus Z$. For all $u \in U$ we can write $u = x + y$ with $x \in X$ and $y \in Y$. Note that $X \subset U$, so $y = u - x \in U$. Hence, $y \in Z = Y \cap U$. Therefore, we see that every $u \in U$ can be written as a sum of elements of X and Z. Also note that $U \cap Z \subset X \cap Y = \{0\}$. Hence, U is the direct sum of X and Z.

Note that U is completely reducible. If U is irreducible, we are done. If not, U has a non-trivial invariant subspace X and $U = X \oplus Z$ for some invariant subspace Z. If X and Z are irreducible, we are done, if not, we do the same steps over and over. This process has to terminate eventually, since U is finite-dimensional. □

The proof of point three and four will be provided below.

Proposition 4.20. *Two irreducible representations of $sl(3;\mathbb{C})$ with the same highest weight is are isomorphic.*

Proof. Suppose that (π, V) and (σ, W) are irreducible representations with the same highest weight μ and let v and w be the highest weight vectors for V and W, respectively. Consider the representation $V \oplus W$ and let U be the smallest invariant subspace of $V \oplus W$ which contains the vector (v, w). We know then that U is a highest weight cyclic representation. Furthermore, since $V \oplus W$ is, by definition, completely reducible, then by the previous lemma 4.19 we know that U is completely reducible. Therefore, by lemma 4.18, U is irreducible.

Consider now the maps P_1 and P_2 such that

$$P_1(v, w) = v; \qquad P_2(v, w) = w.$$

Note that P_1 and P_2 are intertwining maps from $V \oplus W$ to V and $V \oplus W$ to W and their restriction to $U \subset V \oplus W$ are also intertwining maps. Now, neither, $P_1|_U$ or $P_2|_U$ is the zero map, since both are nonzero on (v, w). Moreover, U, V and W are all irreducible. Hence, by Schur's lemma, we see that $P_1|_U$ is an isomorphism of U with V and $P_2|_U$ is an isomorphism of U with W, showing $V \cong U \cong W$. □

Proposition 4.21. *If π is an irreducible representation of $sl(3;\mathbb{C})$ with highest weight $\mu = (m_1, m_2)$, then m_1 and m_2 non-negative integers in the Dynkin basis.*

Proof. By proposition 4.8 we see that m_1 and m_2 are integers. If v is a vector with highest weight μ, then $\pi(T_+)v$ and $\pi(U_+)v$ must be zero. Thus, if we apply theorem 4.4 to the restrictions of π to $\langle H_1, T_+, T_- \rangle \cong \mathrm{sl}(2;\mathbb{C})$ and $\langle H_2, U_+, U_- \rangle \cong \mathrm{sl}(2;\mathbb{C})$, we conclude that m_1 and m_2 are non-negative. □

We will now turn to the last point of the theorem of highest weight.

Proposition 4.22. *If m_1 and m_2 are non-negative integers, then there exists an irreducible representation of $sl(3;\mathbb{C})$ with highest weight $\mu = (m_1, m_2)$ in the Dynkin basis.*

Proof. Since the trivial representation is an irreducible representation with highest weight $(0,0)$, we need only to construct representations with at least one of the m_1 or m_2 positive.

We begin with the constructions of two irreducible representations with highest weight $(1,0)$ and $(0,1)$ which we will call the fundamental representations. So we consider the standard representation of $\mathrm{sl}(3;\mathbb{C})$ acting on $\mathbb{C}^3$ in the usual way, which is an irreducible representation. Note that this representation has weight vectors e_1, e_2 and e_3 with weights $(1,0)$, $(-1,1)$ and $(0,-1)$ respectively. The highest weight is $(1,0)$. The dual of the standard representation is given by

$$\pi(Z) = -Z^{tr}$$

for all $Z \in \mathrm{sl}(3;\mathbb{C})$. This representation is also irreducible. The weight vectors e_1, e_2 and e_3 have weights $(-1,0)$, $(1,-1)$ and $(0,1)$ with highest weight $(0,1)$. Let (π_1, V_1) and (π_2, V_2) be the standard representation and its dual and let $v_1 = e_1$ and $v_2 = e_3$ be their respective highest weight vectors. Now, consider the representation π_{m_1,m_2} given by the extension of theorem 2.51 given by

$$(V_1 \otimes \ldots \otimes V_1) \otimes (V_2 \otimes \ldots \otimes V_2),$$

where V_1 occurs m_1 times and V_2 occurs m_2 times. So we can see π_{m_1,m_2} as

$$\underbrace{\pi_1 \otimes \ldots \otimes \pi_1}_{m_1} \otimes \underbrace{\pi_2 \otimes \ldots \otimes \pi_2}_{m_2}.$$

We can now check that the vector

$$v_{m_1,m_2} = v_1 \otimes \ldots \otimes v_1 \otimes v_2 \otimes \ldots \otimes v_2,$$

is a weight vector with weight (m_1, m_2) and that v_{m_1,m_2} is annihilated by $\pi_{m_1,m_2}(X), \quad X \in \{T_+, V_+, U_+\}$.

For the first claim notice that

$$\pi_{m_1,m_2}(H_1)(v_1 \otimes \ldots \otimes v_1 \otimes v_2 \ldots \otimes v_2) = m_1(v_1 \otimes \ldots \otimes v_1 \otimes v_2 \otimes \ldots \otimes v_2)$$
$$\pi_{m_1,m_2}(H_2)(v_1 \otimes \ldots \otimes v_1 \otimes v_2 \ldots \otimes v_2) = m_2(v_1 \otimes \ldots \otimes v_1 \otimes v_2 \otimes \ldots \otimes v_2).$$

The second claim follows by executing the same steps. Let W be an invariant subspace containing v_{m_1,m_2}. When we assume that π_{m_1,m_2} is completely reducible, we know that W is completely reducible as well by lemma 4.19 and lemma 4.18 tells us that W is the irreducible representation with highest weight (m_1, m_2).

The complete reducibility has to be proven. Note that for both the standard representation and its dual the following hold: $\pi(X)^* = -\pi(X)$ for all $X \in \mathrm{su}(3)$. Therefore, we can re-use the trick we used earlier in the proof of proposition 3.23 for proving skew-adjointness of the tensor product π_{m_1,m_2}. Hence, by proposition 2.54 π_{m_1,m_2} is completely reducible under the action of $\mathrm{su}(3)$ and hence, also under the action of $\mathrm{sl}(3;\mathbb{C}) \cong \mathrm{su}(3)_{\mathbb{C}}$. □

As we now have fully proved the theorem of highest weight, we will focus our attention to the connection between the irreducible representation of $\mathrm{sl}(3;\mathbb{C})$ with $\mathrm{SU}(3)$.

Theorem 4.23. *For every pair (m_1, m_2) of non-negative integers, there exists an irreducible representation Π of $SU(3)$ such that the associated representation π of $sl(3;\mathbb{C})$ has highest weight (m_1, m_2).*

Proof. The standard representation π_1 of $\mathrm{sl}(3;\mathbb{C})$ comes from the standard representation Π_1 of SU(3), and similarly for the dual of the standard representation. By taking tensor products, one can note that Π_{m_1,m_2} is a representation corresponding to the representation π_{m_1,m_2} of $\mathrm{sl}(3;\mathbb{C})$. □

With this last theorem we end our analysis of the classification of the irriducible representations of $\mathrm{sl}(3;\mathbb{C})$. Note that the proof of point five tells us how to construct the irreducible representations of $\mathrm{sl}(3;\mathbb{C})$. Considering the great number of propositions the reader has to digest, we do not blame the said reader for taking a coffee break at this point.

4.5. The Weyl Group

We begin with the definition of the central object of this section: the Weyl group.

Definition 4.24. *Let $\mathfrak{h}$ be the Cartan subalgebra spanned by H_1 and H_2 and let N be the subgroup of $SU(3)$ such that all $A \in N$ have the property that $Ad_A(H) \in \mathfrak{h}$ for all $H \in \mathfrak{h}$. Let Z be the subgroup of $SU(3)$ consisting of those $A \in SU(3)$ such that $Ad_A(H) = H$ for all $H \in \mathfrak{h}$. Then the Weyl group of $SU(3)$, denoted W, is the quotient group Z/N.*

The main goal of this section is to show that the Weyl group is a symmetry of the weights of any finite-dimensional representation of $\mathrm{sl}(3;\mathbb{C})$. Before we can understand this, we must further analyse the Weyl group.

Suppose that we have a Weyl group W and the Cartan subalgebra $\mathfrak{h}$. We can define an action of W on $\mathfrak{h}$ in the following way. For each element $w \in W$ choose an element A in the corresponding coset of N, then

$$w \cdot H = \mathrm{Ad}_A(H).$$

First of all, we must show that this action is well defined. For this end take B an element of the same coset as A, thus $B = AC$ with $C \in Z$. Then

$$\mathrm{Ad}_B(H) = \mathrm{Ad}_A(\mathrm{Ad}_C(H)) = \mathrm{Ad}_A(H).$$

We used that $\mathrm{Ad}_C(H) = H$ by definition. Note that by definition, if $w \cdot H = H$ for all $H \in \mathfrak{h}$, then w is the identity element of W. Therefore, we can identify W with the group of linear transformations of $\mathfrak{h}$ that can be expressed as $H \to w \cdot H$ with $H \in \mathfrak{h}$. The next proposition gives us even more information concerning the Weyl group.

Proposition 4.25. *The group Z consists of the diagonal matrices inside $SU(3)$, the diagonal matrices with diagonal entries $(e^{i\theta}, e^{i\phi}, e^{-i(\theta+\phi)})$ with $\theta, \phi \in \mathbb{R}$. The group N consists of the matrices $A \in SU(3)$ such that for each $j \in 1,2,3$, there exists $k_j \in \{1,2,3\}$ and $\theta_j \in \mathbb{R}$ such that $Ae_j = e^{i\theta_j} e_{k_j}$. Here, e_1, e_2, e_3 is the standard basis for $\mathbb{C}^3$. The Weyl group $W = N/Z$ is isomorphic to the permutation group on three elements.*

Proof. Take $A \in Z$ arbitrary. That means that A commutes with all elements of $\mathfrak{h}$, including H_1, which has eigenvectors e_1, e_2, e_3, with corresponding eigenvalues $1, -1, 0$. Since A commutes with H_1, it has to preserve each of these eigenspaces. Thus Ae_j must be a multiple of e_j for each j. That implies that A is diagonal itself. Now for the converse implication. Assume that $A \in \mathrm{SU}(3)$ is a diagonal matrix. Then A commutes with H_1 and H_2 trivially, and thus with all elements of $\mathfrak{h}$. The claim that the diagonal matrices inside SU(3) have entries $(e^{i\theta}, e^{i\phi}, e^{i(\theta+\phi)})$ with $\theta, \phi \in \mathbb{R}$ comes from the fact that the determinant of the matrices should be 1 and that the columns form an orthonormal basis of $\mathbb{C}^3$.

Now we are going to proof the second part of the theorem. Take $A \in N$ arbitrary. Then AH_1A^{-1} must be in $\mathfrak{h}$ and therefore must be diagonal by definition. Hence, e_1, e_2, e_3 are eigenvectors for AH_1A^{-1}, with the same eigenvalues $1, -1, 0$ as H_1, but not necessarily in the same order. On the other hand, it is easy to see that the eigenvectors of AH_1A^{-1} are of the form Ae_1, Ae_2, A_3e_3, as can be checked by direct calculation. Thus, Ae_j must be some multiple of e_{k_j}, and the constant must have absolute value 1, as A is unitary. Conversely, if Ae_j is a multiple of e_{k_j} for each j, then for any $H \in \mathfrak{h}$, we know that AHA^{-1} will again be diagonal and thus in $\mathfrak{h}$.

We are now going to proof the last part of the theorem. Take an arbitary $A \in W$, then we know from the previous observations that A maps each e_j to a multiple of e_{k_j} for some k_j, depending on j. It is not hard to see that the matrix AHA^{-1} will be diagonal with diagonal entries rearranged by the permutation $j \mapsto k_j$. For any permutation, we can choose the constants to that map taking e_j to $e^{i\theta_j} e_{k_j}$ have determinant 1 (every permutation arises in this way). Thus W, is isomorphic to the permutation group on three elements. □

As we already said, we want to show that the Weyl group is a symmetry of weights for any finite-dimensional representation of sl(3;$\mathbb{C}$). To do this, we will switch to the basis independent notion of weights, as presented originally in definition 4.5 and 4.6. Hence, we see a weight μ again as an element of the Cartan subalgebra $\mathfrak{h}$ spanned by H_1 and H_2. Note that the action of the Weyl group on $\mathfrak{h}$ is unitary, meaning that $\langle w \cdot H, w \cdot H' \rangle = \langle H, H' \rangle$, which can be checked by direct calculations. The following theorem tells us that the roots are invariant under the action of the Weyl group and that $w \cdot \mu$ is a weight of a representation V of sl(3;$\mathbb{C}$) whenever μ is a weight of V.

Theorem 4.26. *Suppose that* (Π, V) *is a finite-dimensional representation of* $SU(3)$ *with associated representation* (π, V) *of* $sl(3;\mathbb{C})$*. If* $\mu \in \mathfrak{h}$ *is a weight for* V *then* $w \cdot \mu$ *is also a weight of* V *with the same multiplicity. In particular, the roots are invariant under the action of the Weyl group.*

Proof. Suppose that μ is a weight for V with weight vector v. Then for all $U \in N$ and $\mathfrak{h}$, we have

$$\begin{aligned}\pi(H)\Pi(U)v &= \Pi(U)(\Pi(U)^{-1}\pi(H)\Pi(U))v \\ &= \Pi(U)\pi(U^{-1}HU)v \\ &= \langle \mu, U^{-1}HU \rangle \Pi(U)v.\end{aligned}$$

Note that $U \in N$ and hence $U^{-1}HU$ is again in $\mathfrak{h}$. Hence, if w is the Weyl group element represented by U, we obtain

$$\pi(H)\Pi(U)v = \langle \mu, w^{-1} \cdot H \rangle \Pi(U)v = \langle w \cdot \mu, H \rangle \Pi(U)v.$$

Hence, $\Pi(U)v$ is a weight vector with weight $w \cdot \mu$.

$\Pi(U)$ can therefore be seen as an invertible map of the weight space with weight μ onto the weight space with weight $w \cdot \mu$, whose inverse is $\Pi(U)^{-1}$. Hence, two weights have the same multiplicity. □

We will now connect the two notions of weights and roots. For example, when we want to express the weights $(1,0)$ and $(0,1)$ we note that

$$\mu_1 = \begin{pmatrix} \frac{2}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & -\frac{1}{3} \end{pmatrix}, \qquad \mu_2 = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & -\frac{2}{3} \end{pmatrix},$$

are such that

$$\begin{aligned}\langle \mu_1, H_1 \rangle &= 1, & \langle \mu_1, H_2 \rangle &= 0, \\ \langle \mu_2, H_1 \rangle &= 0, & \langle \mu_2, H_2 \rangle &= 1.\end{aligned}$$

The positive simple roots are in this case

$$\beta_1 = (2,-1) = 2\mu_1 - \mu_2 = H_1$$
$$\beta_2 = (-1,2) = 2\mu_2 - \mu_1 = H_2.$$

Note that β_1, β_2 have length $\sqrt{2}$ and $\langle \beta_1, \beta_2 \rangle = -1$. Therefore, by basic geometry, the angle between them is $\theta = \frac{2\pi}{3}$.

We are going to connect this to proposition 4.25. Let $w_{(1,2,3)}$ denote the Weyl group element which cyclically permutes the diagonal entries of each $H \in \mathfrak{h}$. We see that $w_{(1,2,3)}$ takes β_1 to β_2, β_2 to $-(\beta_1 + \beta_2)$. This can be identified with a counterclockwise rotation by $\frac{2\pi}{3}$. Similarly, we can see that $w_{(1,2)}$ can be identified with the reflection across the line perpendicular to β_1. The other permutations are calculated similarly.

4.6. Properties of Representations

We have now established a classication of the irreducible representations of $\mathrm{sl}(3;\mathbb{C})$. However, we still do not know

1. The other weights that occur besides the highest weight.
2. The multiplicities of those weights.
3. The dimension of the representations.

In this section we will just state and discuss the theorems in order to prevent a string of propositions and proofs as in the last section. We begin with discussing the first point. For doing this, we need the notion of a convex hull.

Definition 4.27. *If $v_1, ..., v_N$ are elements of a real or complex vector space, the convex hull of $v_1, ..., v_N$ is the set of all vectors of the form*

$$c_1 v_1 + ... + c_N v_N,$$

where the c_j's are non-negative real numbers satisfying $\sum_j c_j = 1$.

Another way to see the convex hull, is to note that the convex hull of $v_1, ..., v_n$ is the smallest convex set such that it contains all of the v_j's.

There are different kind of weights for a representation (π, V) of $\mathrm{sl}(3;\mathbb{C})$. For example, if $\mu \in \mathfrak{h}$ is such that $\langle \mu, H_1 \rangle$ and $\langle \mu, H_2 \rangle$ are integers, then μ is said to be integral. μ is dominant when $\langle \mu, H_1 \rangle \geq 0$ and $\langle \mu, H_2 \rangle \geq 0$.

Theorem 4.28. *Let μ be a dominant integral element and let V_μ be the irreducible representation with highest weight μ. Suppose that λ is an integral element satisfying the following two conditions.*

1. *$\mu - \lambda$ can be expressed as an integer combination of roots.*
2. *λ belongs to the convex hull of $W \cdot \mu$, the orbit of μ under the action of the Weyl group W.*

Then λ is a weight of V_μ.

The proof of this theorem can be found in [10]. This theorem basically says that there are no "unexpected holes" in the set of weights V_μ. For the dimension of each representation, we will refer to the following theorem for which the proof can be found in [10] as well.

Theorem 4.29. *The dimension of the irreducibel representation with highest weight (m_1, m_2) is*

$$\frac{1}{2}(m_1+1)(m_2+1)(m_1+m_2+2).$$

Lastly, the multiplicities of the weights in a finite-dimensional representation of sl$(3;\mathbb{C})$ can be calculated by the Konstant multiplicity formula[10]. However, this is out of the scope of this report. For the representations of sl$(3;\mathbb{C})$ which we will use in this report, we will explicitly indicate the multiplicities of them. In the next section, a few of these weight-root diagrams are provided for representations of sl$(3;\mathbb{C})$. They will make the last few notions of this chapter more insightfull.

5

The Quark Model

In this chapter we follow closely the works of [19], [1] and [3]. In the last few chapters we have seen quite some representation theory. We started with representation theory of finite groups, we continued our journey with representation theory of Lie groups and Lie algebras and gave an application regarding spin in quantum mechanics. In the last chapter we discussed representation theory of SU(3) and in this chapter we are going to give a quite interesting application of representation theory of SU(3): the quark model. In this introduction, we will explain the intuition behind this model.

Before we can proceed, we need to introduce the electromagnetic, strong and weak force, since these play an important role in this model. We begin with the strong force. The atomic nucleus constitutes of protons and neutrons. In nature two particles with the same charges repel, hence, two protons are repelled from each other. Despite the repelling protons, the atomic nucleus stays together. Therefore, physicists hypothesized a new force, the strong force, strong enough to overcome the electric repulsion of the proton. It must be strongest only at short distances, order of magnitude 10^{-15}m, and then it has to fall of rapidly, as protons are repelled by one another unless their seperation is small enough. Neutrons also experience the strong force as they are bound to the nucleus as well. Around the nucleus, electrons are bound by the electromagnetic force. Lastly, the neutron can decay into a proton due to the weak force. In this report we will only focus on the strong force and the particles which interact via the strong force: hadrons. This does not mean that the particles cannot interact via the electromagnetic or weak force, but we pretend that they only interact by means of the strong force.

The hadrons can be split into two groups: mesons and baryons. Mesons have integer spin while baryons have half integer spin. The lightest mesons and baryons can be described by three elementary quarks, the up, down and strange quark. Actually, for the sake of being complete, there are six quark flavours with the addition of the charm, bottom and top quarks. The idea is that the three lightest elementary quarks, the up, down and strange quarks, are related to SU(3) representations. They form a vector space that transforms in the fundamental representation of this group. Their corresponding anti-particles transform in the dual of the fundamental representation. The composite states, for example the mesons and baryons, transform in tensor products of these representations. Theoretically, mesons and baryons in the same irreducible representation of SU(3) should have the same masses. However, this is not the case because the elementary quarks have slightly different masses and charges. The mass (rest energy) of a composite system does not only depends on the rest energy of its constituents, but also on their interaction energy. In the case of the quark model, these include their strong, weak and electromagnetic interactions. This shows that the SU(3) flavour symmetry is only an approximate symmetry. This is roughly the sketch of the quark model.

5.1. Isospin

Before we can proceed to the quark model, we need to introduce two concepts: isospin and hypercharge. In this section and the next one we will explain how these two concepts were formed. We begin with isospin and for this, we must travel back to 1932 after the discovery of the neutron. At first sight, the mass of these particles were almost the same. The mass of a proton is $938,280$ MeV/c^2 where the mass of the neutron is 939.56MeV/c^2. When we assume that the mass difference is due to the difference in electromagnetic interaction, their mass is almost the same with respect to the strong force. Therefore, Werner Heisenberg, pioneer in quantum mechanics, proposed the following idea. He thought of the proton and the neutron as two states of the same particle: the nucleon. When we take the simplest non-trivial Hilbert space for the both the proton and the neutron: $\mathbb{C}$, then the Hilbert space for the nucleon should be $\mathbb{C}^2 \cong \mathbb{C} \oplus \mathbb{C}[1]$. The proton and the neutron should then correspond to the basis vectors of $\mathbb{C}^2$:

$$p = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad n = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Hence, we can see a nucleon as in a superposition of the two particles, see example 1.37, $\alpha p + \beta n \in \mathbb{C}^2$ where $|\alpha|^2 + |\beta|^2 = 1$. So there is probability $|\alpha|^2$ that a measurement of the nucleon will result in finding the proton, and $|\beta|^2$ for finding a neutron. In order for this to be interesting, there should be a process that can turn protons and neutrons into different states of the nucleon, since otherwise there would be no point in having the full $\mathbb{C}^2$ space of states. Conversely, when there is a process that which can change protons into neutrons and back, we need the full $\mathbb{C}^2$ space to describe them.

This was Heisenbergs first try, and although it proved to be a breakthrough, his first ideas turned out to be wrong. He was thinking about the forces between the nucleons in nuclei and it was mathematically convenient to write the proton and neutron as a nucleon doublet as we did above. He was thinking about the neutron as a sort of tightly bound state of proton and electron and imagined that forces between nucleons could arise by exchange of electrons. He thought of this because of a process similar to this one in nuclear physics. Nevertheless, it turned out to be a quite poor analogy.

Around this time, spin was already well understood in quantum mechanics and in 1936 there was a paper by Cassen and Condon [2] suggesting that the nucleon's Hilbert space $\mathbb{C}^2$ is acted on by the symmetry group SU(2), the same as we did when treating spin! Pushing the analogy even further, the property that distinguishes the proton and neutron states of a nucleon was called isospin, the proton has isospin $T_3 = \frac{1}{2}$ and the neutron has isospin $T_3 = -\frac{1}{2}$. Note that these values are half integer and that the proton and neutron live inside a two dimensional representation of su(2). We also say that these particles form an isospin doublet.

Note that the "spin" in isospin reflects on the similarities of isospin with spin. Analogous to our treatment of spin, this means that the proton and neutron live in the Hilbert space $L^2(\mathbb{R}^3)\hat{\otimes} V_{\frac{1}{2}}$ where $V_{\frac{1}{2}}$ is the two dimensional projective representation of SO(3). This is when we are only considering the extra internal degree of freedom caused by isospin. In the case of spin, we labeled it by means of the largest eigenvalue l of the operator $L_3 = i\pi(F_3)$, which could be integer of half integer. Isospin is labeled exactly the same way as spin, but instead of l we use T to label the dimension of the projective representation of SO(3) and instead of the angular momentum operator T_3 we rename it to the isospin operator T_3. For the sake of consistency, we will use the representation of SU(2) in the remainder of this chapter when we are talking projective representations of SO(3) or representations of SU(2).

The proton was assigned isospin up, meaning that it is represented by the basis vector of $V_{\frac{1}{2}}$ with eigenvalue $\frac{1}{2}$ of the T_3 operator. The neutron was assigned isospin down, meaning that it is represented by the basis vector of $V_{\frac{1}{2}}$ with eigenvalue $-\frac{1}{2}$ of the T_3 operator. Or to put it more compactly, the neutron is isospin down with $T_3 = -\frac{1}{2}$ and the proton is isospin up with $T_3 = \frac{1}{2}$, as was stated before.

Isospin proved to be useful since it formalised the idea that the strong force, unlike the electromagnetic force, is the same whether the particles are protons or neutrons. So protons and neutrons are interchangeable, as long as we neglect small differences in mass and more importantly, neglect electromagnetic effects. So to phrase this in terms of group theory and analogous to what we have seen with spin, the Hamiltonian $\mathscr{H}$ exclusively describing the strong force is invariant under the action of SU(2).

As we have seen in the intermezzo of section 2.7, we assume that the Hamiltonian $\mathscr{H}$ commutes with the representation of the elements of the group SU(2) and hence also with the Lie algebra operators associated with isospin. Therefore, each eigenspace for $\mathscr{H}$ is invariant under the isospin Lie algebra operators. In the calculations of section 1.7 we saw that the eigenvalues of the eigenvectors of the Hamiltonian do not change after applying a commuting operator. From this we can see that total isospin T is a conserved quantity after a process that involves the strong force.

So in conclusion, two particles, the neutron and proton each living in a different representations $\mathbb{C}$ of the trivial group, were unified into the nucleon, with representation $\mathbb{C}^2$ of SU(2). This holds for the strong force, but not for the electromagnetic force, it breaks the SU(2) symmetry.

We are now at the end of our treatment of isospin. There are other hadrons as well, called pions, which are mesons. We listed the pions together with their relevant properties.

	Mass (MeV/c^2)	T_3
π^-	139.57	-1
π^0	134.97	0
π^+	139.57	1

These values for T_3 can be obtained by looking at interactions with other particles, especially the neutron and protons, but we will not dive into this theory. We can see that the considering the isospin of the pions, we can take the representation V_1, which is a 3 dimensional representation of SU(2) with eigenvalues -1, 0 and 1 for the T_3 operator. The pions correspond to the basis vectors of this space with the corresponding eigenvalues. The other particles which are of interest are the $K^\pm$ and the K^0 and the $\overline{K}^0$ mesons. But we will provide the properties of these particles at the end of the next section.

5.2. Hypercharge

We are now going to skip a few decades to the 1950s to the discovery of strangeness. By now isospin was a well established approximate symmetry, broken by the weak and the electromagnetic interactions. Around this time, experimenters began to notice a strange new class of particles which were produced in pairs by scattering of strongly interacting particles, and decayed much more slowly back into ordinary particles[6]. A new quantum number was hypothesised, which was conserved by the strong and electromagnetic interaction, but not by the weak interactions, which was called strangenes: S. Particles which carry a non-zero strangeness are called strange particles. Just as with isospin, strangeness could be obtained by looking at particle interactions, but in this report we will not do that.

Another important quantum number is the baryon number B. It is used to distinguish the hadrons into mesons and baryons. Mesons and anti-mesons have baryon number 0, baryons and antibaryons have baryon number 1 and -1 respectively. We can now define the hypercharge as follows:

$$Y = B + S,$$

where B is the baryon number and S is the strangeness of a particle. Lastly, we introduce the Gell-Mann-Nishijima formula, which relates the hypercharge and the isospin with the "classical" charge of a particle, this formula was introduced empirically:

$$Q = T_3 + \frac{Y}{2}.$$

Here, T_3 is the three component of the isospin of a particle, and Y is its hypercharge. With the introduction of this hypercharge, we may suspect that we enlarge the symmetry group of the strong interaction to SU(3) instead of just SU(2) which described isospin.

8 mesons and the relevant particle properties are listed below.

	Mass (MeV/c^2)	T_3	Y
$\pi^\pm$	139.57	± 1	0
π^0	134.97	0	0
$K^\pm$	493,67	$\pm\frac{1}{2}$	± 1
K^0	497,61	$-\frac{1}{2}$	1
$\overline{K}^0$	497,61	$\frac{1}{2}$	-1
η	547,86	0	0

These are the spin 0^- mesons, meaning that the particles have total angular momentum 0 and odd parity, denoted by P. Parity is another concept which will take time to explain, and therefore left out. Let just say that it is another particle property by which we can classify them. When we plot these particles into the $Y - T^3$ plane something interesting happens, as we will explain in the following section.

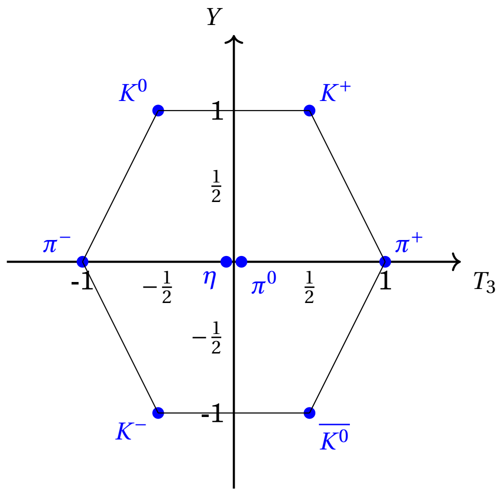


Figure 5.1: $J^P = 0^-$ mesons plotted according to their three component of the isospin and their hypercharge.

5.3. Fundamental Representations and Quarks

Recall our original Gell-Mann basis of sl(3;$\mathbb{C}$) given by the Gell-Mann matrices in section 4.2. We noted that H_1 and H_2 are basis of a Cartan subalgebra of sl(3;$\mathbb{C}$). However, any linear combination of them would suffice as basis of the Cartan subalgebra, and hence, $T_3 = \frac{1}{2}\lambda_3 = \frac{1}{2}H_1$ and $Y = \frac{2}{\sqrt{3}}T_8 = \frac{1}{3}(H_1 + H_2)$ would suffice as well. So the Cartan subalgebra can be defined as span$\{T_3, Y\}$. There is a reason why we labelled the basis vectors of the Cartan subalgebra by T_3 and Y, as will become clear in a moment. The root vectors for this basis of the Cartan subalgebra are the same as the root vectors for the previous basis of the Cartan subalgebra. However, the commutator relations are slightly different.

$$[T_3,T_+]=T_+, \qquad [T_3,V_+]=\frac{1}{2}V_+, \qquad [T_3,U_+]=-\frac{1}{2}U_+,$$
$$[T_3,T_-]=-T_-, \qquad [T_3,V_-]=-\frac{1}{2}V_-, \qquad [T_3,U_-]=\frac{1}{2}U_-,$$
$$[Y,T_+]=0, \qquad [Y,V_+]=V_+, \qquad [Y,U_+]=U_+,$$
$$[Y,T_-]=0, \qquad [Y,V_-]=-V_-, \qquad [Y,U_-]=-U_-.$$
$$[T_+,T_-]=2T_3 \qquad [V_+,V_-]=\frac{2}{3}Y+T_3 \qquad [U_+,U_-]=\frac{2}{3}Y-T_3$$

Note that $\langle T_3, T_\pm\rangle$ is isomporphic to sl(2;$\mathbb{C}$) as one can see from the commutator relations and we denote these set of operators as sl(2;$\mathbb{C}$)$_I$. We will use this when we are branching sl(3;$\mathbb{C}$) representations into sl(2;$\mathbb{C}$) representations in section 5.4.1. We can see that the roots are

$$\beta_{T_\pm}=(\pm1,0), \qquad \beta_{V_\pm}=\left(\pm\frac{1}{2},\pm1\right), \qquad \beta_{U_\pm}=\left(\mp\frac{1}{2},\pm1\right).$$

Hence, this means that the $T_\pm$ affects the isospin of states in a representation by one unit, while the $V_\pm$ and $U_\pm$ raise or lower the isospin of states by half a unit and the hypercharge by 1 unit.

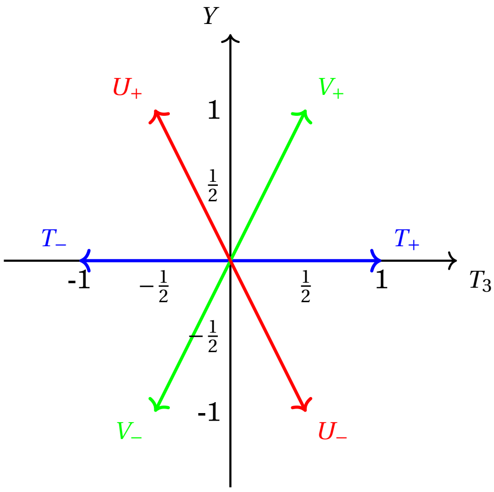


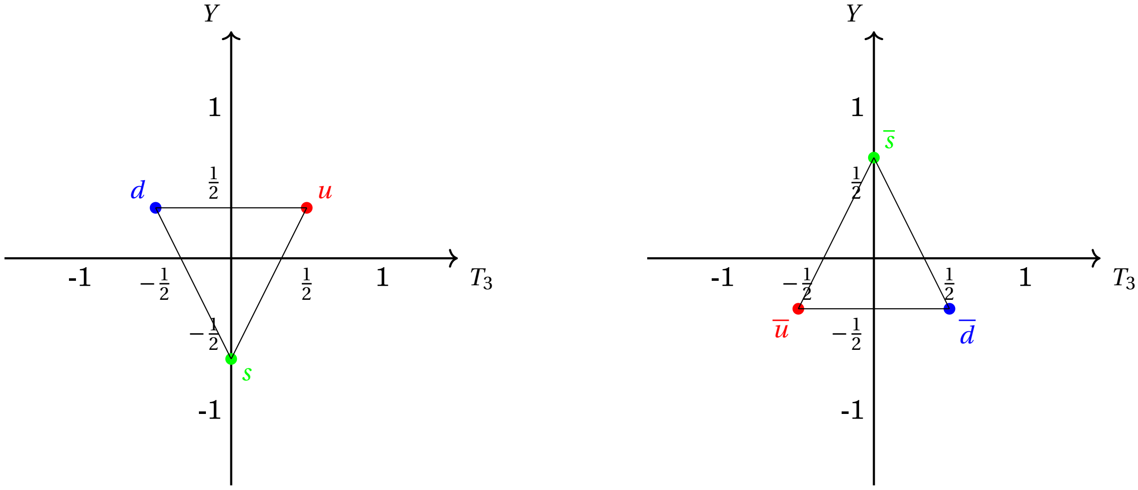
Figure 5.2: Root diagram of sl(3;$\mathbb{C}$) with T_3 and Y as basis for the Cartan subalgebra.

The weights we introduced to label the irreducible representation of sl(3;$\mathbb{C}$) must be modified for the new basis $\{T_3, Y\}$ of the Cartan subalgebra. This can be done in the following way. Suppose that (m_1, m_2) labels an irreducible representation in the original basis $\{H_1, H_2\}$ of the Cartan subalgebra. Then the new weights can be obtained by realising that $\{T_3, Y\}$ are linear combinations of H_1, H_2. Therefore $\left(\frac{1}{2}m_1, \frac{1}{3}(m_1+m_2)\right)$ in the $\{T_3, Y\}$ basis corresponds to the weight (m_1, m_2) in the basis $\{H_1, H_2\}$. So, the fundamental representation, corresponding to the highest weight $(\frac{1}{2}, \frac{1}{3})$ in our new basis $\{T_3, Y\}$ and denoted by , has a basis of states

$$u=\begin{pmatrix}1\\0\\0\end{pmatrix}, \qquad d=\begin{pmatrix}0\\1\\0\end{pmatrix}, \qquad s=\begin{pmatrix}0\\0\\1\end{pmatrix},$$

corresponding tot the up, down and strange quarks. These vectors have weights $(\frac{1}{2}, \frac{1}{3})$, $(-\frac{1}{2}, \frac{1}{3})$ and $(0, -\frac{2}{3})$ corresponding to the weight vectors u, d and s respectively. The dual of the standard representation, has a basis of states $\overline{u}$, $\overline{d}$ and $\overline{s}$ corresponding to the anti-quark states. Recall that the dual of a standard representation of (π, V) of sl(3;$\mathbb{C}$) is given by $\pi(Z)=-Z^{tr}$ for all $Z \in$ sl(3;$\mathbb{C}$). Therefore, we can easily see that the weights of the representation are $(-\frac{1}{2}, -\frac{1}{3})$, $(\frac{1}{2}, -\frac{1}{3})$ and $(0, \frac{2}{3})$

for u, d and s respectively. The weight diagrams are given below.



(a) Weight diagram of the $\left(\frac{1}{2}, \frac{1}{3}\right)$ representation.

(b) Weight diagram of the $\left(0, \frac{2}{3}\right)$ representation.

Figure 5.3: Weight diagrams for the fundamental representation and the dual of the fundamental representation.

We listed the quarks together with their important properties in the table below.

Quark	B	T	T_3	σ	S	Y	Q
u	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{3}$	$\frac{2}{3}$
d	$\frac{1}{3}$	$\frac{1}{2}$	-$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{3}$	-$\frac{1}{3}$
s	$\frac{1}{3}$	0	0	$\frac{1}{2}$	-1	-$\frac{2}{3}$	-$\frac{1}{3}$
$\overline{u}$	-$\frac{1}{3}$	$\frac{1}{2}$	-$\frac{1}{2}$	$\frac{1}{2}$	0	-$\frac{1}{3}$	-$\frac{2}{3}$
$\overline{d}$	-$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0	-$\frac{1}{3}$	$\frac{1}{3}$
$\overline{s}$	-$\frac{1}{3}$	0	0	$\frac{1}{2}$	1	$\frac{2}{3}$	$\frac{1}{3}$

Note that the quarks all have spin $\frac{1}{2}$ and that they have baryon number $\frac{1}{3}$ because one needs three quarks to form a baryon. Note as well that mesons consist of a quark and anti-quark, hence the anti-quarks should then have baryon number $-\frac{1}{3}$. The u and d quark have zero strangeness, so their hypercharge can be calculated by $Y = B + S$ and the results can be found in the table above. The same can be said about the strange quark with strangeness -1. The charge of the quarks can be computed by the Gell-Mann-Nishijima formula, which yield:

$$Q = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{-1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & -\frac{2}{3} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & -\frac{1}{3}. \end{pmatrix}$$

This calculation shows the up quark has $\frac{2}{3}$ electric charge, while the down and the strange quark have -$\frac{1}{3}$ electric charge.

5.4. Mesons

We begin with the simplest composite states: mesons, these consist of a quark and an anti-quark. Therefore, they lie in the tensor product representation $(\mathbf{3} \otimes \overline{\mathbf{3}})_{\mathrm{sl}(3;\mathbb{C})} = (\mathbf{8} \oplus \mathbf{1})_{\mathrm{sl}(3;\mathbb{C})}$[6]. The integer denotes the dimension of the irreducible representation and the subscript denotes whether it is an $\mathrm{sl}(3;\mathbb{C})$ or $\mathrm{sl}(2;\mathbb{C})$ representation. The bar denotes that we are considering a dual representation.

So we expect to find an octet of mesons with similar masses and an additional singlet meson. As we know from the proof of theorem 4.21, the weights in a tensor product representation correspond to the sum of the weights in the fundamental representation. Hence, we can obtain a basis for the $\mathbf{3}\otimes\overline{\mathbf{3}}$ representation. From this point on, when we write ud for example, $u\otimes d$ is meant.

States	$u\overline{u}, d\overline{d}, s\overline{s}$	$u\overline{d}$	$u\overline{s}$	$d\overline{u}$	$s\overline{s}$	$s\overline{u}$	$s\overline{d}$
Weights	$(0,0)$	$(1,0)$	$(\frac{1}{2},1)$	$(-1,0)$	$(-\frac{1}{2},1)$	$(-\frac{1}{2},-1)$	$(\frac{1}{2},-1)$

The associated weight diagram is given below. As one can see, there are three states with weight $(0,0)$. However, only two linear combinations belong to the octet.

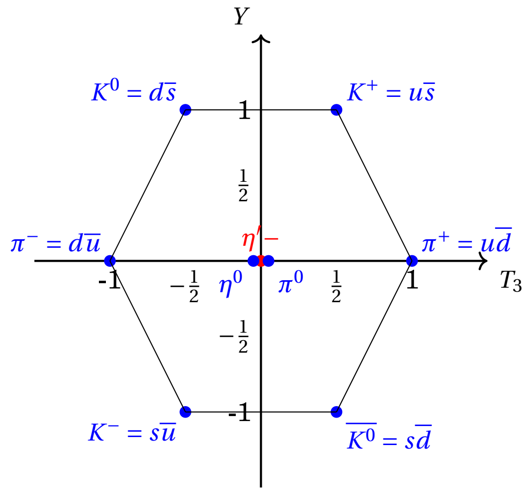


Figure 5.4: Weight diagram of the $\mathbf{3}\otimes\overline{\mathbf{3}}$ representation, including the names of the mesons they correspond to. The particles in this octet all have spin 0. These are the $J^P = 0^-$ mesons as mentioned earlier.

To determine which two linear combinations are in the octet and which one belongs to the singlet we note the following. Suppose we start of at $K^+ = u\overline{s}$ in our diagram and apply the $V_-\otimes -V_-^{tr}$ operator and we do the same with the $K^0 = d\overline{s}$ state and the $U_-\otimes -U_-^{tr}$ operator, then

$$V_-\otimes -V_-^{tr}(u\overline{s}) = (V_-u)\otimes\overline{s} + u\otimes\left(-V_-^{tr}\overline{s}\right) = s\overline{s} - u\overline{u},$$
$$U_-\otimes -U_-^{tr}(d\overline{s}) = (U_-d)\otimes\overline{s} + d\otimes\left(-U_-^{tr}\overline{s}\right) = s\overline{s} - d\overline{d}.$$

These states are in no way orthogonal, but we can make them by choosing the right linear combinations:

$$\pi^0 = \frac{1}{\sqrt{2}}(u\overline{u} - d\overline{d}), \qquad \eta^0 = \frac{1}{\sqrt{6}}(u\overline{u} + d\overline{d} - 2s\overline{s}).$$

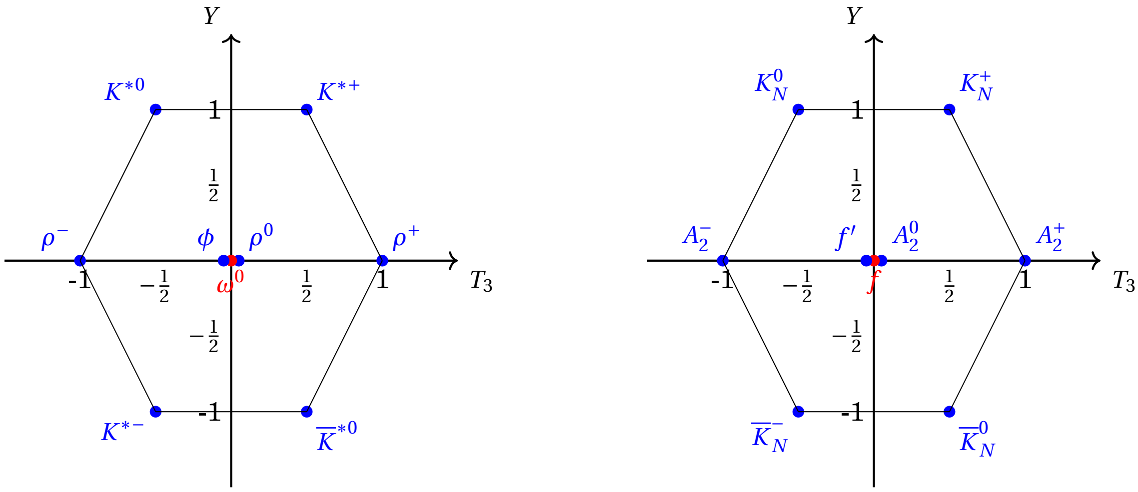
These two states yield the two $(0,0)$ mesons in the octet. The singlet meson is a linear combination of π^0 and η^0 that is orthogonal to both π^0 and η^0, yielding

$$\eta' = \frac{1}{\sqrt{3}}(u\overline{u} + d\overline{d} + s\overline{s}).$$

The exact values of the scalars of the linear combinations can be obtained by Clebsch-Gordan decomposition, but that is beyond the scope of this report. It is not hard to see that this state remains invariant under the $\mathrm{sl}(3;\mathbb{C})$ representation, it is killed by all operators $U_\pm$, $V_\pm$ and $T_\pm$. Note that by example 1.37, we see that the probability of finding a certain quark-anti quark combination is given by the square of the coefficient in front of that combination. So we see that when we measure the quark content of π^0 we have $\frac{1}{2}$ chance to find $u\overline{u}$ and $\frac{1}{2}$ chance to find $d\overline{d}$.

This is just one example of such meson octet and singlet. Some of them have been completed thus far, but the particles of the lowest meson mass octet are stable under the strong interaction, the others are not. They decay to members of the lowest octets in times on the order of 10^{-24} seconds[3] and are dubbed the meson resonances. The other meson resonances differ in spin, parity and mass.

We will just provide the different weight-root diagrams corresponding to these multiplets based on empirical data. We will list a few, but there are more.



(a) The $J^P = 1^-$ mesons.

(b) The $J^P = 2^+$ mesons.

Figure 5.5: Two meson octets.

It may be confusing that particles with the same quarks are in different multiplets, meaning that they have different total angular momentum. We will provide some intuition for this. Recall the results of section 3.7 where we treated the addition of angular momentum in quantum mechanics. Presumably, the energy (and thus the mass) increases rather drastically with the orbital angular momentum.

The spin 0 particles can be formed by a quark and anti-quark with anti-parallel spins and zero orbital angular momentum, or when the quark spins are parallel and the quarks have one unit of orbital angular momentum directed opposite to the spins. With anti parallel spins we mean that the spins are represented by a basis vectors of the two dimensional representation of SU(2) with eigenvalues $-\frac{1}{2}$ and $\frac{1}{2}$. However, this results in higher energies than the first octet of particles[5].

Spin one mesons are formed by the following compositions of quarks.

1. Quark spins parallel and zero orbital angular momentum.
2. Quark spins anti-parallel and one unit of orbital angular momentum.
3. Quark spins parallel and two units of orbital angular momentum directed opposite to the spins.

5.4.1. Branching of sl(3;ℂ) Representations into sl(2;ℂ) Representations

The postulates and mathematical reasoning behind the quark model and the mathematical reasoning seems to be invalid physically. This can be seen as follows. If the strong interaction is invariant under the sl(3;ℂ) operators, then all the hadrons of an octet or a decimet should have the same mass. The operators of sl(3;ℂ) change one quark into the other, or equivalently, one member of a supermultiplet (singlets, octets, decimets) into the other without changing any of the interactions. If the quark model is to be valid, it must be that the interaction responsible for the supermultiplets is not the strong interaction, but some other interaction which is invariant under the operations of sl(3;ℂ). Gell-Mann realized this and proposed the following modification of the strong force:

- **Very strong interactions:** Invariant under SU(3) symmetry under which the light baryons and mesons transform like octets and decuplets.
- **Medium strong interactions:** This breaks SU(3) symmetry, but conserve isospin and hypercharge.

In section 4.2, we have seen that sl(3;ℂ) contains the subgroup sl(2;ℂ) in three possible ways. One of the possibilities is to look at $\{T_3, T_\pm\}$ or equivalently $\{\lambda_1, \lambda_2, \lambda_3\}$ in the Gell-Mann basis. They correspond to the subgroup

$$X = \begin{pmatrix} U_2 & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{sl}(3;\mathbb{C}).$$

Here, $U_2 \in \mathrm{sl}(2;\mathbb{C})_I$. Note that this embedding leaves the strange quark invariant while performing sl(2;ℂ) transformations in the sub-space spanned by the u and d vectors in the fundamental representation. Hence, these from a sl(2;ℂ) isospin doublet and do not change the hypercharge of the whole system, which can be seen from the table of quark properties. We can therefore see that we have the decomposition [19]

$$\mathbf{3}_{\mathrm{sl}(3;\mathbb{C})} \to (\mathbf{2} \oplus \mathbf{1})_{\mathrm{sl}(2;\mathbb{C})_I}.$$

Here, u and d are mapped to the basis vectors in the two dimensional represenation of sl(2;ℂ), while s is mapped to the one dimensional representation of sl(3;ℂ). For a more in depth-look on induced representations and branching, see [13]. The arrow denotes the decomposition and $\mathbf{3}_{\mathrm{sl}(3;\mathbb{C})}$ denotes the 3 dimensional fundamental representation of sl(3;ℂ) with highest weight $(\frac{1}{2}, \frac{1}{3})$. Similarly, we can see that

$$\overline{\mathbf{3}}_{\mathrm{sl}(3;\mathbb{C})} \to (\overline{\mathbf{2}} \oplus \mathbf{1})_{\mathrm{sl}(2;\mathbb{C})_I}.$$

Here the bar denotes that the dual representation should be considered. Hence, $\overline{\mathbf{3}}_{\mathrm{sl}(3;\mathbb{C})}$ is the dual of the fundamental representation with highest weight $(0, \frac{2}{3})$. $\overline{\mathbf{2}}$ and $\mathbf{2}$ denote the two dimensional irreducible representations of sl(2;ℂ) and its dual and hence, are isomorphic to one another by theorem 4.3. We can apply this as well to the case of mesons:[19]

$$\begin{aligned}
(\mathbf{3} \otimes \overline{\mathbf{3}})_{\mathrm{sl}(3;\mathbb{C})} &\to (\mathbf{2} \oplus \mathbf{1}) \otimes (\mathbf{2} \oplus \mathbf{1})_{\mathrm{sl}(2;\mathbf{C})_I} \\
&\to (\mathbf{2} \otimes \mathbf{2} \oplus \mathbf{2} \otimes \mathbf{1} \oplus \mathbf{1} \otimes \mathbf{2} \oplus \mathbf{1} \otimes \mathbf{1})_{\mathrm{sl}(2;\mathbf{C})_I} \\
&\to (\mathbf{3} \oplus \mathbf{1} \oplus \mathbf{2} \oplus \mathbf{2} \oplus \mathbf{1})_{\mathrm{sl}(2;\mathbb{C})_I}.
\end{aligned}$$

Here we made use of proposition 3.24. We note that an sl(3;ℂ) singlet will be an $\mathrm{sl}(2;\mathbb{C})_I$ singlet as well [19], so we can conclude

$$\mathbf{8}_{\mathrm{sl}(3;\mathbb{C})} \to (\mathbf{3} \oplus \mathbf{2} \oplus \mathbf{2} \oplus \mathbf{1})_{\mathrm{sl}(2;\mathbb{C})_I}.$$

Each of these multiplets must have a constant hypercharge, as the only T_3 can be contained in the $\mathrm{sl}(2;\mathbb{C})_I$ subgroup of sl(3;ℂ), and not Y. Hence, for the mesons we see that

$$\mathbf{3}_{\mathrm{sl}(2;\mathbb{C})} = (\pi^+, \pi^0, \pi^-)_{Y=0}, \qquad \mathbf{2}_{\mathrm{sl}(2;\mathbb{C})} = (K^+, K^0)_{Y=1},$$
$$\mathbf{2}_{\mathrm{sl}(2;\mathbb{C})} = (K^-, \overline{K}^0)_{Y=1}, \qquad \mathbf{1}_{\mathrm{sl}(2;\mathbb{C})} = (\eta^0)_{Y=0}.$$

That the η^0 is an isospin singlet should be clear from its explicit form $\eta = \frac{1}{\sqrt{2}}(u\overline{u} + d\overline{d})$. We see that η is invariant under the exchange of the u and d quarks and of their anti-particles. The same applies to the η' singlet. When we compare the masses of these triplets, doublets and singlets, we see that the masses of particles in each triplet, doublet and singlet are relatively close to one another. As one could calculate from the quark content of (π^+, π^0, π^-), the π^-, π^0, π^+ have charges

$-1, 0$ and 0 respectively. Hence, the electromagnetic interaction breaks this isospin further and explains why the mass of the π^0 is different from the π^+ and π^- masses. We implicitly assumed that the electromagnetic and the strong interactions are invariant under charge conjugation.

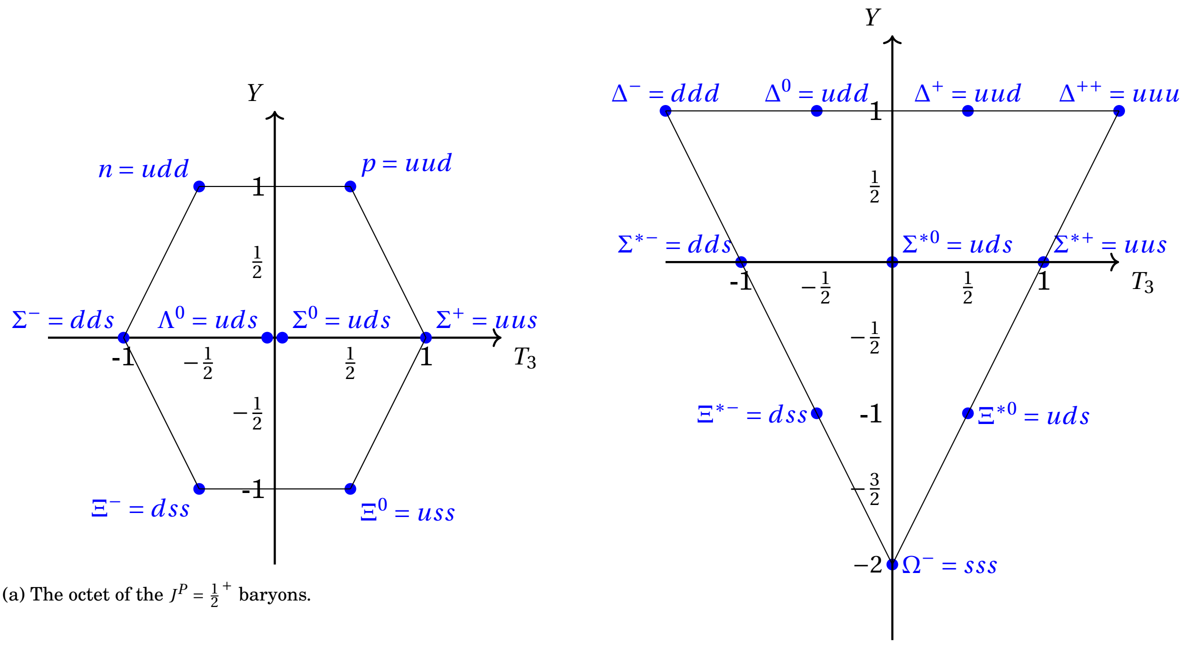
As a last remark note that all of this is an example of what we discussed in section 1.7, where we discussed the situation of lower symmetry groups contained in higher symmetry groups. In our case, the medium strong interaction causes a perturbation of the Hamiltonian of the very strong interactions, causing that the masses of different isospin multiplets of the same octets are different.

5.5. Baryons

We can also consider baryons, which are composed of three quarks and the antibaryons having three anti-quarks. Note that we expect baryons to transform according to the tensor representation[6]

$$(\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3})_{\mathrm{sl}(3;\mathbb{C})} = (\mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1})_{\mathrm{sl}(3;\mathbb{C})}.$$

Not all of these irreducible representations occur in nature as quarks are spin-$\frac{1}{2}$ fermions that also carry an additional SU(3) colour charge, but this is far beyond the scope of this report. The irreducible representations that do occur are the **10**, **8** and the **1** representation, e.g. there is only one baryon octet possible. The baryon decuplet includes the spin $\frac{3}{2}$ baryons, while the baryon octet contain only the states with spin $\frac{1}{2}$. We will first provide the baryon octets and decuplet.



(a) The octet of the $J^P = \frac{1}{2}^+$ baryons.

(b) Decuplet of the $J^P = \frac{3}{2}^+$ baryons.

Figure 5.6: The baryon octet and decuplet.

We must make an important remark concerning the udd and uud particles. They can live as basis vectors of the 2 and 4 dimensional representations of SU(2), when only regarding isospin. Considering isospin exclusively, the up and down quarks live inside a 2 representation of SU(2) denoted by **2**. Again, the integer denotes the dimension of the representation and for this case we assume that the representations are all representations of SU(2). So to put it in more familiar terms, they have $T = \frac{1}{2}$ or $T = \frac{3}{2}$. This is no wonder, as $\mathbf{2} \otimes \mathbf{2} \otimes \mathbf{2} \cong (\mathbf{3} \oplus \mathbf{1}) \otimes \mathbf{2} \cong \mathbf{4} \oplus \mathbf{2} \oplus \mathbf{2}$. Here we made use of proposition 3.24. Because of the additional colour charge of the barions, there is no second baryon octet. Therefore, we see that udd and uud are either basis vectors of the 2 dimesnional representation of SU(2) or the 4 dimensional representation of SU(2) augmented by the $\Delta^- = ddd$ and $\Delta^{++} = uuu$ particles.

We end this section with discussing the power of this model. In the 1960's, Gell-Mann, the man who developed this quark model, noticed that there was a glaring hole on the spot where now the Ω^- particle is situated. He calculated its mass and other properties by means of this model and urged experimental physicists to find this particle. In 1964 this particle was ultimately found, showing the predicting power of this model and in 1969 he even received the nobel price for this work!

6
Conclusion

This thesis has given an introduction to finite dimensional representation theory for the Lie groups SU(2) and SU(3). This thesis also provided two applications of representation theory in physics: the spin and quark models. We provided a rough sketch of how these models work.

In chapter 1 we looked at representation theory of finite groups and introduced notions such as representation, intertwining maps, character theory and complete reducibility to create intuition for representation theory. At the end of chapter 1 we mentioned some applications of representation theory in quantum mechanics such as the use of symmetry groups in pertubation theory.

In chapter 2 we treated representation theory of Lie groups and Lie algebras. An emphasis was put on the Lie groups SU(2) and SO(3) since these play an important role in the description of spin. We specifically showed that SU(2) is in fact the double cover of SO(3) as can be found in theorem 2.35. Furthermore, we introduced the notion of projective unitary representations in section 2.7.2 and showed that by means of tensor products and direct sums of representations we can construct new representations from old ones.

In chapter 3 we gave a description of spin in quantum mechanics by means of representation theory. At the end of this chapter, we came to the conclusion that half integer spin representations of the Lie algebra so(3) cannot be exponentiated to SO(3), but they can be exponentiated to its universal cover: SU(2).

We also argued what we can modify the quantum Hilbert space in which a particle lives: $L^2(\mathbb{R}^3)$ to account for the extra internal degrees of freedom introduced by spin. We do this by modifying the Hilbert space of the particle as follows: $L^2(\mathbb{R}^3)\hat{\otimes}V$, where $\hat{\otimes}$ denotes the completion of the Hilbert space $L^2(\mathbb{R}^3)\otimes V$. Here, V is a finite dimensional Hilbert space that carries an irreducible representation of SO(3). We know from proposition 2.46 that V carries an ordinary representation π of the Lie algebra so(3). We label V by the largest eigenvalue l of the angular momentum operator L_3 as defined in section 3.1. Furthermore, we saw that V_l carries a projective unitary representation of SO(3) when l is half integer and V_l carries an ordinary unitary representation of SO(3) when l is integer.

Lastly, in chapter 3 we discussed the addition of total angular momentum in quantum mechanics from a mathematical viewpoint. We discussed the decomposition of the tensor product of irreducible representations V_l and V_m of so(3) into SO(3) invariant subspaces of $L^2(\mathbb{R}^3)$.

In chapter 4, we treated representation theory for SU(3), which is closely connected to representation theory for the Lie algebra sl(3;$\mathbb{C}$). We showed that we can classify all the irreducible representations by means of their highest weight. Furthermore, we introduced the notion of the Weyl group and showed that the Weyl group is a symmetry of weights of any finite-dimensional representation of sl(3;$\mathbb{C}$). Lastly, we discussed a few other properties of irreducible representations of sl(3;$\mathbb{C}$), such as the dimension of irreducible representations and the other weights that occur besides the highest weight.

In chapter 5, we discussed the quark model and to be more precise: we discussed flavor symmetry for the lightest baryons and mesons. These are particles which also interact by means of the strong force. We assumed that the building blocks of these particles, up, down and strange quarks, form a vector space that transforms in the fundamental representation of the group SU(3) and their antiparticles in the dual of the fundamental representation. With this description, their composite states, mesons and baryons, transform in the tensor products of these representations.

We argued that theoretically, the mesons and baryons in the same irreducible representation of SU(3) should have the same masses. However, as we have seen for example for the lightest meson octet, this is not the case. Instead, we introduced the modification of the strong force, into the very strong interactions, which preserved the total SU(3) symmetry, and the medium strong interaction, which broke the SU(3) symmetry, but conserved isospin and hypercharge. With this we branched the sl(3;$\mathbb{C}$) representations into sl(2;$\mathbb{C}$) representations and made subgroups of the particles in the octet of the lightest mesons with masses which were much more similar to one another.

However, this model is not complete. As mentioned several times in chapter 5, Quantum Chromo Dynamics (QCD) also plays a role in the description of mesons and baryons. QCD is also described by the symmetry Lie group SU(3). For further study, one can study how QCD is incorporated into the description of mesons and baryons. Also another topic for further study is to look at the whole symmetry group of the standard model, which describes the structure and the interaction between the elementary particles: U(1)$\times$SU(2)$\times$SU(3).

7
Acknowledgements

First and foremost, I would like to thank Dr. Wolter Groenevelt for his patience and for guiding me during this project. I am thankful for the time and effort he put in explaining the concepts behind the abstract theory of Lie groups and Lie algebras. I am especially thankful for pulling me back on earth whenever I wanted to do way too complex, far above my level mathematics. It was a fun and enlightening journey! I am as well thankful for the guidance of professor Yuli Nazarov for providing feedback and ideas from the viewpoint of a physicist. I would also like to thank Dr. David de Laat and Dr. Anton Akhmerov for being part of my thesis committee. Furthermore, I would like to thank Hidde de Bos for pre-reading parts of my report and giving me valuable feedback. I would like to thank my family for supporting me during my bachelor degree. Lastly, I would like to thank my girlfriend, Isabel, for always supporting me.

Bibliography

[1] John C. Baez and John Huerta. *The Algebra of Grand Unified Theories*. 2010. arXiv: 0904.1556 [hep-th].

[2] B. Cassen and E. U. Condon. “On Nuclear Forces”. In: *Phys. Rev.* 50 (9 Nov. 1936), pp. 846–849. DOI: 10.1103/PhysRev.50.846. URL: https://link.aps.org/doi/10.1103/PhysRev.50.846.

[3] J.R. Christman, ed. *SU(3) and the quark model*. 2001. URL: http://www.physnet.org/modules/pdf_modules/m282.pdf.

[4] Marius Crainic, ed. *Inleiding Topologie*. Mathematisch Instituut Universiteit Utrecht, 2017.

[5] M.S. Dresselhaus, ed. *Applications of groups to the physics of solids.* MIT readers, 2002. URL: http://web.mit.edu/course/6/6.734j/www/group-full02.pdf.

[6] H. Georgi, ed. *Lie algebras in Particle Physics*. Perseus Books, 1999. ISBN: 0-7382-0233-9.

[7] D. Gijswijt, ed. *Algebra 1*.

[8] David J. Griffiths and Darrell F. Schroeter, eds. *Introduction to quantum mechanics*. Cambridge University Press, 2018. ISBN: 978-1-107-18963-8. DOI: 10.1017/9781316995433.

[9] W.G.M. Groenevelt, ed. *Special Functions an Representation Theory*.

[10] Brian C. Hall, ed. *Lie groups, Lie algebras and representations*. Springer, 2003. ISBN: 978-3-319-13466-6. DOI: 10.1007/978-3-319-13467-3.

[11] Brian C. Hall, ed. *Quantum theory for mathematicians.* Springer, 2013. ISBN: 978-1-4614-7115-8. DOI: 10.1007/978-1-4614-7116-5.

[12] A. Hatcher, ed. *Algebraic Topology.* Cambridge university press, 2001. ISBN: 978-0-521-79540-1.

[13] W. Fulton J. Harris, ed. *Representation theory a first course.* Springer, 1991. ISBN: 0-387-97495-4.

[14] Alexander Kirillov Jr, ed. *An Introduction to Lie Groups and Lie Algebras*. Cambridge University Press, 2008. DOI: 10.1017/CBO9780511755156.

[15] N. Miller, ed. *Representation theory and quantum mechanics*. 2018.

[16] James R. Munkres, ed. *Topology*. Prentice Hall, 1975. ISBN: 0-13-181629-2.

[17] J. van Neerven, ed. *Applied Functional Analysis*. 2021.

[18] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511976667.

[19] J. G. Rosa, ed. *Lie Groups and Lie Algebras in Particle Physics*. URL: http://gravitation.web.ua.pt/sites/default/files/migrated2016/LieGroups.pdf.

[20] W. Rudin, ed. *Principles of Mathematical Analysis, 3rd edn. International Series in Pure and Applied Mathematics.* McGraw-Hill, New York-Auckland-Düsseldorf, 1976.

[21] J.-P. Serre, ed. *Linear Representations of Finite Groups.* Springer-Verlag, 1977. ISBN: 978-1-4684-9460-0.

[22] Benjamin Steinberg, ed. *Representation Theory of Finite Groups.* Springer, 2012. ISBN: 978-1-4614-0775-1.

[23] J. Thijssen, ed. *Lecture notes on classical mechanics*. TU Delft, 2017.

[24] J. Vermeer, ed. *Lineaire Algebra, Deel 2 voor Technische Natuurkunde*.

A

Algebraic topology

In this appendix the notions of connectedness and simply connectedness are explained in a proper way. However, first we have to introduce the notion of topology:

Definition A.1. *A topology on a set X is a collection $\mathscr{T}$ of subsets of X, satisfying the following axioms:*

- *$\emptyset$ and X belong to $\mathscr{T}$.*
- *The intersection of any two sets from $\mathscr{T}$ is again in $\mathscr{T}$.*
- *The union of any collection of sets of $\mathscr{T}$ is again in $\mathscr{T}$.*

A topological space is a pair $(X,\mathscr{T})$ consisting of a set X and a topology $\mathscr{T}$ on X. A subset $U \subset X$ is called closed in the topological space $(X,\mathscr{T})$ if $X \setminus U$ is open.

Now one has the notion of topology one has to define what is meant with continuous function:

Definition A.2. *Given two topological spaces $(X,\mathscr{T}_X)$ and $(Y,\mathscr{T}_{\mathscr{Y}})$ and a function $f: X \to Y$, we say that f is a continuous funtion if:*

$$f^{-1}(U) \in \mathscr{T}_X, \ \forall U \in \mathscr{T}_X.$$

We are now in the position of defining what connectedness and path connectedness are.

Definition A.3. *A topological space $(X,\mathscr{T}_X)$ is said to be path connected if X cannot be written as the union of two disjoint non-empty opens $U,V \subset X$.*

Definition A.4. *We say that a topological space $(X,\mathscr{T}_X)$ is path-connected if for any $x, y \in X$ there exists a path γ connecting x and y, i.e. a continuous map $\gamma : [0,1] \to X$ such that $\gamma(0) = x$ and $\gamma(1) = y$.*

It is interesting to note that any path connected space is as well connected.

In topology, one is interested in continuous functions between topological spaces. However, such as one can define equivalence classes in algebra, one can also define equivalence classes in a set of continuous functions between two spaces. The equivalence class hinted at is being homotopic equivalent:

Definition A.5. *Two functions $f : X \to Y$ and $g : X \to Y$ are homotopic to one another when there is a family of maps $h_t : X \to Y$, $t \in [0,1]$ such that the associated map $F : X \times [0,1] \to Y$ given by $F(x,t) = f_t(x)$ is continuous and $h_0 = f$ and $h_1 = g$.*

When one has two topological spaces $(A, \mathcal{T}_A)$ and $(B, \mathcal{T}_B)$ one can idenify all continuous functions up to homotopy. That is one can define a set of continuous functions between these two spaces, $[A,B]$ up to homotopy. That is, two functions are in the same equivalence class if there exists a homotopy between them.

From this the notion of simply connectedness follows quite easily.

Definition A.6. *A topological space $(X, \mathcal{T}_X)$ is simply connected when* $[\mathbb{S}^1, X] = \{.\}$

The intuition is that every "loop", e.g. continuous function from $\mathbb{S}^1$ into X is homotopy equivalent to a point. Hence, every loop in X can be continuously shrunk into a point in the space X. It is interesting to note that any simply connected space is path-connected. Another notion which is of great importance is that of the concatenation of paths, the fundamental group and the pushforward.

Definition A.7. *If f is a path in a space X from x_0 to x_1, and g is a path in X from x_1 to x_2, we define the product $f * g$ of f and g to be the path h given by the equations*

$$h(s) = \begin{cases} f(2s) & \text{for } s \in [0, \frac{1}{2}] \\ g(2s-1) & \text{for } s \in [\frac{1}{2}, 1] \end{cases}$$

Definition A.8. *Let X be a space, let x_0 be a point of X. A path in X that begins and ends at x_0 is called a loop based at x_0. The set of path homotopy classes of loops based at x_0, with the operation $*$, is called the fundamental group of X relative to the base point x_0. It is denoted by $\pi_1(X, x_0)$.*

Definition A.9. *Let $h : (X, x_0) \to (Y, y_0)$ be a continuous map. Define*

$$h_* : \pi_1(X, x_0) \to \pi_1(Y, y_0)$$

by the equation

$$h_*([f]) = [h \circ f].$$

The map h_* is called the pushforward of h, relative to the basepoint x_0.

It can be proven that the fundamental group is indeed a group [12]. We need two more definition concerning the free group and deformation retractions.

Definition A.10. *Let G be a group, let $\{G_\alpha\}_{\alpha \in J}$ be a family of subgroups of G of G that generates G. Suppose that $G_\alpha \cap G_\beta$ consists of the identity element alone whenever $\alpha \neq \beta$. We say that G is the free product of the groups G_α if fore each $x \in G$, there is only one reduced word in the groups G_α that represents x. In this case, we write*

$$G = \prod_{\alpha \in J}^{*} G_\alpha$$

or in the finite cases, $G = G_1 * ... G_n$.

Definition A.11. *Let A be a subspace of the topological space X. We say that A is a deformation retract of X if the identity map of X is homotopic to a map that carries all of X into A, such that each point of A remains fixed during the homotopy. This means that there is a continuous map $H : X \times I \to X$ such that $H(x,0) = x$ and $H(x,1) \in A$ for all $x \in X$, and $H(a,t) = a$ for all $a \in A$. The homotopy H is called a deformation retraction of X onto A.*

A.1. The fundamental group of SO(3)

In chapter 2 we analysed some topological properties of the groups SU(2) and SO(3). In this section we consider SO(3) and prove what the fundamental group of this particular object is, which will turn out to be $\mathbb{Z}/2\mathbb{Z}$. A more glamorous result is the fact that for all $n > 3$ the fundamental group of SO(n) is actually the same as for SO(3) [10], however we will not prove this.

The strategy for proving that the fundamental group of SO(3) is equal to $\mathbb{Z}/2\mathbb{Z}$ is the following. We first start with another topological object, namely the real projective space $\mathbb{R}P^2$ and calculate the fundamental group of that space by using Van Kampen, theorem A.12. Then we are going to proof that there is an homeomorphism from SO(3) to $\mathbb{R}P^3$. Lastly we are going to show that the fundamental group of $\mathbb{R}P^3$ can be calculated by using the fundamental group of $\mathbb{R}P^2$.

Theorem A.12. ***(Seifert-Van Kampen)*** *Let $X = U \cup V$ be a pointed topological space, where U and V are open in X; assume U, V and $U \cap V$ are path connected; let $x_0 \in U \cap V$. Let*

$$j : \pi_1(U, x_0) * \pi_1(V, x_0) \to \pi_1(X, x_0)$$

be the homomorphism of the free product that extends the homomorphisms j_1 and j_2 induced by the inclusion. Then j is surjective, and its kernel is the least normal subgroup N of the free product that contains all elements represented by words of the form

$$(i_{U_*}(g)^{-1}, i_{V_*}([g])),$$

for $[g] \in \pi_1(U \cap V, x_0)$. Here, $i_{U_*} : \pi_1(U \cap V, x_0) \to \pi_1(X, x_0)$ is the homomorphism induced by the injection map $i_U : U \cap V \to U$ and similarly i_{V_*} is defined similarly for V.

We explain the notion of free product and pushforward in Appendix A.

A.1.1. The fundamental group of $\mathbb{R}P^2$

We begin with noting that the space $\mathbb{R}P^2$ is the same as the 2-sphere whose antipodal points are identified. Actually, the space $\mathbb{R}P^n$ is homeomorphic to $D^n/(x \sim -x)$ where $x \in \partial D^n \cong S^{n-1}$ [4]. Here D^n denotes the n-ball. Hence, it is homeomorphic to a disc whose antipodal points along the boundary are identified, which gives the planar representation of $\mathbb{R}P^2$. Note that we can write the planar representation of $\mathbb{R}P^2$ as the union of an annulus A and an open disc B whose intersection is again an annulus. The open disc, the annulus and their intersection are path connected and opens in the standard topology, which means that the requirements to use Van Kampen are satisfied. We clearify the sets A and B in figure A.1.

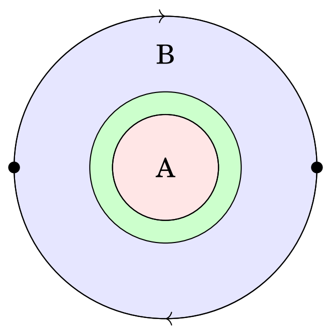


Figure A.1: Planar representation of $\mathbb{R}P^2$ with the sets $A \setminus B$ in red, $B \setminus A$ in blue and their intersection in green.

Before we can use Van Kampen, we still need to specify the relations of the annulus and the open disc which make up the space $\mathbb{R}P^2$. Note that when a topological space X deformation retracts to a topological space Y, they are homotopy equivalent and thus has the same fundamental group [16]. Hence, it is of interest to look what spaces the open disc and the annulus can deformation retract to. We can easily see that the disc A can be deformation retracted to the point and therefore it has fundamental group $\pi_1(A) = \langle 1 | \emptyset \rangle$. The annulus $A \cap B$ can be deformation retracted to the circle, and hence it has fundamental group equal to $\pi_1(A \cap B) = \langle a|\emptyset\rangle \cong \mathbb{Z}$. However, as one can see from figure 2.2, B is homeomorphic to the Möbius strip and hence has fundamental group $\pi_1(B) = \langle b|\emptyset\rangle \cong \mathbb{Z}$.

We have now identified the fundamental groups of the three topological spaces which are of interest to us A, B and $A\cap B$ and we are therefore in the position to calculate the fundamental group of $\mathbb{R}P^2$ by means of Van Kampen:

$$\pi_1(\mathbb{R}P^2) = 1 \underset{\mathbb{Z}}{*} \mathbb{Z}.$$

Now the question remains: what is the group $1 \underset{\mathbb{Z}}{*} \mathbb{Z}$ exactly? For answering this, we turn to Van Kampen. We have to use the old relations from $\pi_1(A)$ and $\pi_1(B)$ with the new relations which we will obtain by inspecting $\pi_1(A\cap B)$:

$$\pi_1(\mathbb{R}P^2) = \langle 1, b|, i_{A_*}([\gamma])i_{B_*}([\gamma])^{-1}\rangle.$$

As we know $\pi_1(A) = 1$, therefore we can conclude that i_{A_*} must be the trivial map $i_{A_*}([a]) = 1$. However, i_{B_*} is a bit more difficult to obtain. For this end note that the loop b corresponds to two loops a, as can be seen from figure 2.2, hence

$$i_{B_*}([a]) = b^2$$

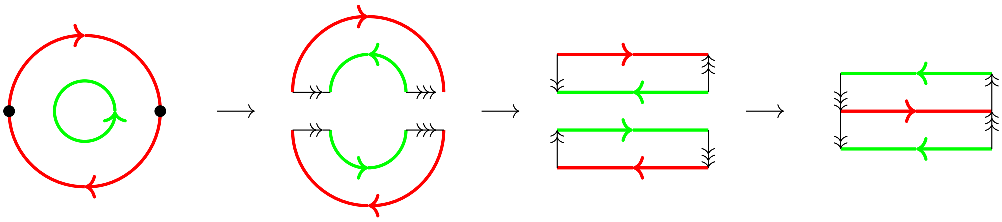


Figure A.2: The two generators of the loop space of set A and B in red, b, and green, a, respectively. The most right picture of this figure is a planar representation of the Möbius strip. We see that the red loop corresponds to two green loops in the Möbius strip.

Therefore, we can write the final result

$$\pi_1(\mathbb{R}P^2) = \langle 1, b|1 = b^2\rangle = \mathbb{Z}/2\mathbb{Z}.$$

A.1.2. SO(3) and $\mathbb{R}P^3$

In this section, we mostly follow parts of 1.3.4 of [10]. We are now going to show that the topological spaces SO(3) and $\mathbb{R}P^3$ are homeomorphic. For this to be true there must exists a continuous bijection between the two spaces with a continuous bijective inverse as well. This is what we are going to show. However, for proving this we first must show that every $R \in SO(3)$ is actually a rotation by an angle θ in the plane orthogonal to a certain eigenvector of R.

We begin with the claim that every element $R \in SO(3)$ has eigenvalue equal to 1. This claim is equivalent with the statement that $\det(R - I) = 0$. Note the following:

$$\det(R-I) = \det\big(R - R^{tr}R\big) = \det\big(I - R^{tr}\big)\det(R) = (-1)^3\det(R-I) = -\det(R-I).$$

Hence, we can conclude that $\det(R - I) = 0$. So there exists an $v \in \mathbb{R}^3$ such that $Rv = v$. Then take $w \in \mathbb{R}^3$ arbitrary such that $\langle v, w\rangle = 0$. We want to show that Rw is still in the plane orthogonal to the vector v. This follows quite easily.

$$\langle v, Rw\rangle = \langle R^{tr}v, w\rangle = \langle v, w\rangle = 0.$$

Therefore we can conclude that Rw is still in the plane orthogonal to v. Every vector can be decomposed along the eigenvector v and the plane orthogonal to it. Therefore, we can see $R \in \mathrm{SO}(3)$ as a rotation by some angle θ around the axis v.

Let v be a unit vector in $\mathbb{R}^3$ and let $R_{v,\theta}$ be the element of SO(3) consisting of a right handed rotation by an angle θ in the plane orthogonal to v. It is not hard to see that $R_{-v,\theta} = R_{v,-\theta}$ and we showed that every element of SO(3) can be expressed as $R_{v,\theta}$ for some v and θ with $-\pi \leq \theta \leq \pi$. Furthermore, by either choosing v or $-v$ we can set θ to be in the range 0 to π. If $R = I$ then $R = R_{v,0}$ for any unit vector $v \in \mathbb{R}^3$. If R is a rotation by an angle π around an axis v then R can both be expressed by $R_{v,\pi}$ and $R_{v,-\pi}$. If $R \neq I$ then R has a unique representation as $R_{v,\theta}$ with $0 \leq \theta \leq \pi$.

Let B^3 be the closed ball of radius π in $\mathbb{R}^3$ and consider the map $\Phi : B^3 \to \mathrm{SO}(3)$ given by

$$\begin{aligned} \Phi(u) &= R_{\hat{u},\|u\|}, \quad u \neq 0, \\ \Phi(0) &= I. \end{aligned}$$

Here, $\hat{u} = \frac{u}{\|u\|}$ is the unit vector in the u-direction. It is easy to see that the map Φ is continuous, also at I. This is seen by the fact that $R_{\hat{u},\|u\|}$ approaches the identity as $\|u\|$ approaches zero. From the previous observations concerning the representations of the elements of SO(3), we can conclude that Φ maps B^3 onto SO(3). It is not hard to see that the map Φ is injective as well, exept for the antipodal boundary points of B^3, as they have the same image $R_{v,\pi} = R_{-v,\pi}$. Hence Φ is a continuous injective map from $\mathbb{R}P^3$ onto SO(3). As both spaces $\mathbb{R}P^3$ and SO(3) are compact, the inverse is continuous as well [20] and hence SO(3) and $\mathbb{R}P^3$ are homeomorphic to each other.

A.1.3. Calculating the fundamental group of SO(3)

We have now showed that SO(3) and $\mathbb{R}P^3$ are homeomorphic as topological spaces and we have calculated the fundamental group of $\mathbb{R}P^2$. It is now time to put all the together and calculate the fundamental group of $\mathbb{R}P^3$. As we have noted earlier $\mathbb{R}P^3$ is homeomorphic to $D^n/(x \sim -x)$ where $x \in \partial D^3 \cong S^2$. Take an arbitrary smaller open 3-ball U contained in our original D^3 and let V be $D^3 \setminus A$, where A is an open in D^3. Note that $V \cong S^2/(x \sim -x)$ by deformation retraction. Note that $V \cap U \cong S^2$ and hence $\pi_1(U \cap V) = 1$, so no new relations arise from $\pi_1(A \cap B)$. Also note that U, V and $U \cap V$ are all path connected and hence, we are in the position to apply Van Kampen.

$$\pi_1(\mathbb{R}P^3) = \pi_1(U) * \pi_1(V)/\pi_1(U \cap V) = \pi_1(V) = \pi_1(\mathbb{R}P^2) = \mathbb{Z}/2\mathbb{Z}.$$

Thus, we see that the fundamental group of $\mathbb{R}P^3$ and $\mathbb{R}P^2$ are the same! Also, since $\mathbb{R}P^3$ and SO(3) are homeomorphic to each other, we see that the fundamental group of $\mathbb{R}P^3$ and hence the fundamental group of SO(3) is equal to $\mathbb{Z}/2\mathbb{Z}$.

We can also analyse spin in another way which involves some algebraic topology, which is quite interesting from a mathematical viewpoint. To recap most important results of last chapter concerning the fundamental groups of SO(3), we note that $\pi_1(\mathrm{SO}(3)) = \mathbb{Z}/2\mathbb{Z}$. We also showed that the SU(2) is a double cover of SO(3) (and that it is the universal cover of SO(3)). Therefore, as $\pi(\mathrm{SO}(3)) = \mathbb{Z}/2\mathbb{Z}$ we can divide the space of loops starting and ending at the same point in SO(3) in two distinct classes: the trivial loops and the non-contractible ones. The contractible loops in SO(3) lift to a contractible one in SU(2) while the non-contractible ones lift to a loop starting at I and ending at $-I$ in SU(2)[15]. Hence, they can correspond to a projective representation of SO(3)!

B

Tensor Product

In this part of the appendix we will clarify the notion of tensor products. Firstly, we will introduce the definition of a tensor product, then we will state the universal property of tensor products and we will conclude with some elementary notions of tensor products.

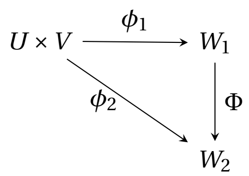
Definition B.1 The tensor product of two vector spaces V and W over a field is a vector space $V \otimes W$ equipped with a bilinear map

$$V \times W \to V \otimes W, \qquad v \times w \mapsto v \otimes w,$$

which is universal: for any bilinear map $\beta : V \times W \to U$ to a vector space U, there is a unique linear map from $V \otimes W$ to U that takes $v \otimes w$ to $\beta(v, w)$. When one would like to show which ground field is actually used, one can denote $V \otimes_K W$.

Firstly, one wold like to know whether one can construct a tensor product from finite dimensional vector fields V and W and whether the whether the tensor product is unique op to canonical homomorphism.

Theorem B.2 If U and V are any finite-dimensional real or complex vector spaces, then a tensor product (W, ϕ) exists. Furthermore, (W, ϕ) is unique up to canonical isomorphism: if (W_1, ϕ_1) and (W_2, ϕ_2) are two tensor products. then there exists a unique vector space isomorphism $\Phi : W_1 \to W_2$ such that the following diagram commutes.



Suppose that (W, ϕ) is a tensor product and that $e_1, e_2, ..., e_n$ is a basis for U and $f_1, f_2, ..., f_m$ is a basis for V. Then

$$\{\phi(e_j, f_k) | 1 \le j \le n, 1 \le k \le m\}$$

is a basis for W.

An elementary result is that if $\{e_j\}$ and $\{f_j\}$ are bases for V and W, the elements $\{e_i \otimes f_j\}$ form a basis for $V \otimes W$. Also $\dim(U \otimes V) = \dim(U) \otimes \dim(V)$.

One has also the tensor product $V_1 \times ... \times V_n \to V_1 \otimes ... \otimes V_n$, which sends elements $v_1 \times ... \times v_n \mapsto v_1 \otimes ... \otimes v_n$.

We are going to introduce one last proposition concerning operators on tensor products.

Proposition B.3 Let U and V be finite dimensional real or complex vector spaces. Let $A : U \to U$ and $B : V \to V$ be linear operators. Then 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 $v \in V$. If A_1, A_2 are linear operators on U and B_1, B_2 are linear operators on V, then

$$(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1 A_2) \otimes (B_1 B_2)$$