2022

Representation Theory and its Applications in Physics

Jakub Bystrický

Representation Theory and its Applications in Physics

Jakub Bystricky

An Honours Thesis presented to the faculty of the
Department of Mathematics at
Colby College

Department of Mathematics
Colby College
Waterville, ME
May, 2022

Abstract

Representation theory is a branch of mathematics that allows us to represent elements of a group as elements of a general linear group of a chosen vector space by means of a homomorphism. The group elements are mapped to linear operators and we can study the group using linear algebra. This ability is especially useful in physics where much of the theories are captured by linear algebra structures. This thesis reviews key concepts in representation theory of both finite and infinite groups. In the case of finite groups we discuss equivalence, orthogonality, characters, and group algebras. We discuss the importance and implications of Maschke's and Schur's theorem. Our study of finite group representations is concluded by an example of an application of the representation of the permutation group S_3 to a system of particles. In the case of infinite groups, we devote all our attention to Lie group representations as applications of representation theory in quantum physics predominantly rely on them. We develop a way to build operators that could be used to capture invariant properties by means of representations of unitary Lie groups.

Acknowledgments

I would like to express my deepest gratitude to Professor Nora Youngs who supervised this thesis and has been incredibly supportive during the entire process and always helped me find motivation, even in the toughest of moments. Without her help this thesis would not be possible and I truly appreciate her willingness to be my supervisor and help. I would also like to express my gratitude to Professor Robert Bluhm who provided me with valuable feedback and his expertise in theoretical physics.

I would also like to thank Professor Leo Livshits, whom I discussed several challenging topics covered in this thesis with. My appreciation goes also to the entirety of the Department of Mathematics at Colby College for its welcoming atmosphere. I would like to acknowledge Huan Bui, a Colby College alum who studied mathematics and inspired me to write this thesis and also let me use his LaTex formatting for the thesis. My thanks also go to Cindy Zhang and Eduardo Sosa, current students also writing a thesis in mathematics who helped make my experience of writing this thesis more enjoyable.

I would not be able to complete this thesis without the support of: my parents - Erik and Alena, my closest friends - Iliana, Eileen, Akezhan, Yeeun, Lutie, Katie, and my girlfriend - Auroshree. Throughout my time in college they all helped me tremendously and made sure I gave my best to all that I did.

Contents

Chapter 1

Group Representations

In our first chapter we will look at some core definitions and properties of group representations. These representations are homomorphisms from a group to a general linear group of a vector space and thus allow us to study the represented groups in a new way. However, group representation might be an unfamiliar concept to many, as it is not usually touched upon in undergraduate mathematics. It is nevertheless an incredibly rich area of mathematics that makes use of many fundamental properties and ideas. The majority of the theory is algebraic, but there are plenty of analysis and even topology concepts. Before starting our discussion of group representations, we state the definition of a group. It should be noted that all the major results in this chapter can also be found in many textbooks. The one we used, and the definitions, lemmas, and theorems follow the form of Steinberg's *Representation of Finite Groups* [8], unless otherwise specified, we always assume we are working over $\mathbb{C}$.

Definition 1.0.1. *A set G is a group under a chosen operation if and only if this operation is associative and the following holds under the operation:*
i) G has an identity
ii) G is closed
iii) All elements of G have an inverse.

This definition does not specify what type of elements a group may contain. For example, it is perfectly reasonable to have a group be a set of letters with a permuting operation such that the definition of a group is satisfied. However, such a structure might not be very convenient for calculations and the study of the group. Imagine, on the other hand, if we could assign (and we can) these letters to linear operators over a vector field and find an operation between them that models the permuting operation on the set of letters.

This is what representation theory, at its core, does. So how might we $send$ the letters to linear operators over a vector space? To answer this question, we turn to the definition of a representation.

Definition 1.0.2. *A representation of a group G is a homomorphism $\Phi : G \longrightarrow GL(V)$, for some finite dimensional vector space V and its general linear group $GL(V)$. The dimension of V is called the degree of Φ. For $g \in G$, we shall write Φ_g for $\Phi(g)$ and $\Phi_g v$ for $\Phi_g(v)$, the action of Φ_g on $v \in V$.*

Now it should be clear that the reason we said $send$ instead of $match$ is that the representation is a function that does not need to be surjective. It merely needs to preserve the structure of the group, i.e. be a homomorphism. That is, it need not be an isomorphism. Consider the following example:

Example 1. *Let $\Phi : \mathbb{Z}/4\mathbb{Z} \longrightarrow GL(\mathbb{C})$ be the representation given by $\Phi([m]) = i^m$.* △

The elements of the group of cosets $\mathbb{Z}/4\mathbb{Z}$ are of the form $\{m+4x|x \in \mathbb{Z}\}$ and thus $\mathbb{Z}/4\mathbb{Z}$ only has four elements. We can choose four representative m's to be the values $(0, 1, 2, 3)$. We see that the representation, i.e. the homomorphism, is one to one, but is not onto. In addition, we can immediately observe that the operation is not the same in the $\mathbb{C}^*$ space as it was in the $\mathbb{Z}/4\mathbb{Z}$ space. What is important, though, is that the action of the elements on the remaining elements is preserved. For instance, when the element of $\mathbb{Z}/4\mathbb{Z}$ with $m = 3$ is added to the one with $m = 1$, we get the identity, $m = 0$. This action is preserved as $i^3 i^1 = i^4 = 1 = i^0$.

Thus, the group $\mathbb{Z}/4\mathbb{Z}$ is $represented$ in $\mathbb{C}^*$ by the group $(1, i, -1, -i)$ under multiplication via the representation Φ. It is important to understand this distinction in terminology. **The representation is the homomorphism**, but it allows for a group to be $represented$ by elements within another group.

There are many ways in which representation theory is used in physics. One is this sort of direct analogy between a system and a group described above. Another place where representation theory is used is a little more sophisticated than that. Quantum mechanical state spaces (we will discuss these soon) are Hilbert spaces [6], and thus are complex vector spaces. Any group action on a Hilbert space is a representation. Discussion of the states of particles therefore uses representation theory all the time, because by representing these states in this way we can study their behaviour and properties are more easily and intuitively.

1.1 Equivalence and Order of Representations

At this point, a few questions arise naturally. The first is about the uniqueness of representations, and the second one about their classification. If we were to claim a representation of a given group G is unique, we would be saying there always only exists one homomorphism of G to a new space V. This is intuitively not the case. Consider the representation in **Example 1** but this time we replace i with $-i$. This new representation is not the same as the one discussed in the example and yet, both are homomorphisms from $\mathbb{Z}/4\mathbb{Z}$ to $GL(\mathbb{C})$. We must turn to a different way of classifying. Instead, we will separate representations into equivalence classes. We shall define these as follows.

Definition 1.1.1. *Let $\Phi : G \longrightarrow GL(V)$ and $\Psi : G \longrightarrow GL(W)$ be representations. Φ and Ψ are said to be* ***equivalent*** *if there exists an isomorphism $T : V \longrightarrow W$ s.t. $\Phi_g = T\Psi_g T^{-1}$ for all $g \in G$. We write $\Phi \sim \Psi$.*

Theorem 1.1.2. *Equivalence of representations, as stated in* ***Definition 1.1.1.*** *is an equivalence relation on the representations.*

Proof. We need to prove that the relationship is reflexive, transitive, and symmetric. Since the identity transformation satisfies the definition as follows $\Phi_g = I\Phi_g I^{-1} = I\Phi_g I$, the relation is reflexive. Furthermore, if $\Phi_g = T\Psi_g T^{-1}$ and $\Psi_g = H\Upsilon_g H^{-1}$ then $\Phi_g = T(H\Upsilon_g H^{-1})T^{-1}$. Since both H and T must be isomorphisms, their composition is also an isomorphism. Noting that $(HT)^{-1} = H^{-1}T^{-1}$ we conclude the relation is transitive. To show it is also symmetric we assume $\Phi_g = T\Psi_g T^{-1}$, then since T is an isomorphism it also must hold that $\Phi_g T = T\Psi_g T^{-1}T$ and $T^{-1}\Phi_g T = K\Phi_g K^{-1} = \Psi_g$, where $K = T^{-1}$. □

What this definition really says is that representations are equivalent if they have the same codomain and there is an isomorphism between the outputs of these homomorphisms in this codomain. We want to think of equivalent representations as the same representation, but, loosely speaking, expressed in terms of a different basis system. We now have a form of uniqueness. An important observation is that the codomains must be vector spaces, by definition. Thus, if there is to be an isomorphism between them, they must have the same dimension. The dimension of the vector space V from $GL(V)$ is captured by the order of a representation.

Definition 1.1.3. *Let $\Phi : G \longrightarrow GL(V)$ be a representation, and let n be the dimension of V, then the order of Φ is $ord(\Phi) = n$.*

Combining the definition of equivalence and of the order we conclude that if two representations of a group are equivalent, they must have the same order. Consider the following example:

Example 2. *Let* $\Psi : \mathbb{Z}/n\mathbb{Z} \longrightarrow GL_2(\mathbb{C})$ *be given by* $\Psi[m] = \begin{bmatrix} \cos\left(\frac{2\pi m}{n}\right) & -\sin\left(\frac{2\pi m}{n}\right) \\ \sin\left(\frac{2\pi m}{n}\right) & \cos\left(\frac{2\pi m}{n}\right) \end{bmatrix}$ *and let* $\Phi : \mathbb{Z}/n\mathbb{Z} \longrightarrow GL_2(\mathbb{C})$ *be given by* $\Phi[m] = \begin{bmatrix} e^{\frac{2\pi mi}{n}} & 0 \\ 0 & e^{\frac{2\pi mi}{n}} \end{bmatrix}$.
We claim these are equivalent and are related by a transformation matrix $A = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$. *As it is a straightforward matrix multiplication, it can be easily verify that, indeed,* $A\Psi A^{-1} = \Phi$, *where we also need to make use of Euler's identity.* △

1.2 Decomposability, Reducibility, and Irreducibility

Thus far we have shown that representations are homomorphisms and that some of them are equivalent to each other. Now we will proceed to show that representations can be decomposed into subrepresentations. This is an extremely useful observation as this will allow us to study the properties of both the representation and its final space better. We will be able to observe that some representations cannot be decomposed any further, similarly to how we cannot decompose prime numbers. We will also identify certain subspaces of the representation's codomain's underlying vector space as invariant (or closed) under the representation!

In order to understand the decomposability of representations we must have a way of composing them. Though it is possible to think of multiple ways of $adding$ two homomorphisms, we shall only focus on their external direct products.

Definition 1.2.1. *Given the representations* $\Phi^{(1)} : G \longrightarrow GL(V_1)$ *and* $\Phi^{(2)} : G \longrightarrow GL(V_2)$, *the external direct product* $\Phi^{(1)} \oplus \Phi^{(2)} : G \longrightarrow GL(V_1 \oplus V_2)$ *is given by* $(\Phi^{(1)} \oplus \Phi^{(2)})g(v_1, v_2) = (\Phi^{(1)}(v_1), \Phi^{(2)}(v_2))$.

The vector space over which the codomain of the composed representation lies is always higher dimensional than that of the vector spaces of the codomains of either of the individual representations. To further illustrate how the direct sum works, consider the trivial representation.

Definition 1.2.2. *The trivial representation* $\Phi_{id} : G \longrightarrow GL(V),\ dim(V) = 1$ *is given by* $\Phi_{id}(g) = I$ *for all* $g \in G$.

Note that the representation $\Phi : G \longrightarrow GL_n(V)$, for $n > 1$ given by $\Phi_g = I$ for all $g \in G$ is not equivalent to the trivial representation. That is, you cannot find an isomorphism that will transform it into the trivial representation (identity in a one dimensional space) -

there are no isomorphisms between spaces of different dimension. Rather, it is equivalent to n copies of the trivial representation, composed via the external **direct sum**. Composing k representations results in a matrix with k block matrices, coming from the original representations, along its diagonal. See example below.

Example 3. *Let $\Phi^{(1)} : \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{C}^*$ and $\Phi^{(2)} : \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{C}^*$ be given by $\Phi^{(1)}[m] : e^{\frac{2\pi im}{n}}$ and $\Phi^{(2)}[m] : e^{-\frac{2\pi im}{n}}$ respectively. Then $(\Phi^{(1)} \oplus \Phi^{(2)})[m] = \begin{bmatrix} e^{\frac{2\pi im}{n}} & 0 \\ 0 & e^{-\frac{2\pi im}{n}} \end{bmatrix}$. Note that the final representation is $(\Phi^{(1)} \oplus \Phi^{(2)}) : G \longrightarrow GL(\mathbb{C}^2)$.* △

To allow ourselves to continue our discussion we define G-invariance, which will help us understand the behaviour of the representation's outputs (e.g. the matrix given in **Example 3**, for a specific m) in the vector space they act over (e.g. $\mathbb{C}^2$ in **Example 3**).

Definition 1.2.3. *Let $\Phi : G \longrightarrow GL(V)$ be a representation. A subspace W of V is G-invariant if for all $g \in G$ and $w \in W$, one has that $\Phi_g w \in W$.*

The notion that the representation can act on elements outside the G-invariant space too can be a little uncomfortable. However, the importance of this definition is in that there exist subspaces of the vector space over which the representation's codomain lies that are closed under the group actions Φ_g for all $g \in G$. In a certain sense, we are speaking more about the properties of the given subspace than we are about the homomorphism.

The entire vector space over which the codomain (the general linear group) of a given representation is, is always G-invariant. We conclude that there must be some G-invariant subspaces in the codomain's respective vector space of a representation if it is a composition of other representations, because their individual codomains are now over subpaces of the composite representation codomain's vector space V.

If we look at **Example 3** again, we notice that any vector of the form $(z, 0) \in \mathbb{C}^2, z \in \mathbb{C}$ will maintain its form under $(\Phi^{(1)} \oplus \Phi^{(2)})[m]$ for all m. That is, the space spanned by $(z, 0)$ is a G-invariant subspace of the composed representation. The same holds for the other dimension ($(0, z) \in \mathbb{C}^2$) and thus $\Phi^{(1)}$ and $\Phi^{(2)}$ will now be called **subrepresentations** of $(\Phi^{(1)} \oplus \Phi^{(2)})$ and will be denoted $\Phi^{(i)}|\mathbb{C}$. Following the intuitions of this example we come to the following definitions.

Definition 1.2.4. *A non-zero representation Φ of a group G is **decomposable** if $V = V_1 \oplus V_2$ with V_1, V_2 non-zero G-invariant subspaces. Otherwise, it is called **indecomposable**.*

Definition 1.2.5. *A non-zero representation $\Phi : G \longrightarrow GL(V)$ is said to be **irreducible** if the only G-invariant subspaces of V are $\{0\}$ and V.*

Definition 1.2.6. *Let G be a group. A representation $\Phi : G \longrightarrow GL(V)$ is said to be* ***completely reducible*** *if $V = V_1 \oplus V_2 \oplus ... \oplus V_n$, where V_i are G-invariant subspaces and $\Phi|V_i$ is irreducible for all $i = 1, 2, ..., n$.*

One of the outcomes of these definitions is that representations from any group to $\mathbb{C}$ must be irreducible. This is the case as there is no proper non-zero subspace in $\mathbb{C}$ and thus there is no proper G-invariant subspace either. As a consequence, any representation to $\mathbb{C}$ has no subrepresentations. In addition, since $\mathbb{C}$ has dimension one, the order of the representation is also one. Composing order one representations produces a diagonal matrix (like in **Example 3**). As such, we can intuitively observe that complete reducibility of a representation $\Phi : G \longrightarrow \mathbb{C}^n$ must be related to diagonalizability of the matrices Φ_g for all $g \in G$.

We have stated earlier that we want to think of equivalent representations as the same representation via the equivalence relationship we proved. Now we would like to see that some further properties are preserved, i.e. be invariant, within an equivalence class. This should only further strengthen the notion that equivalent representations are the same representation.

Lemma 1.2.7. *Let $\Phi : G \longrightarrow GL(V)$ be equivalent to a decomposable representation. Then Φ is decomposable.*

Proof. Let $\Phi : G \longrightarrow GL(W)$ be a decomposable representation such that $\Phi \sim \Psi$. Therefore, $\exists\ T : V \longrightarrow W$ such that $\Psi_g = T^{-1}\Phi_g T$ for all $g \in G$. By decomposability of Φ, $W = W_1 \oplus W_2$ for some $W_1, W_2 \ \leq W$. Thus, we have $V_1, V_2 \ \leq V$, $V_1 = T^{-1}(W_1)$ and $V_2 = T^{-1}(W_2)$.

If $v \in V_1 \cap V_2$ then $Tv \in W_1 \cap W_2 = \{0\}$ and since T is injective it is precisely $v = 0$. If $v \in V$ then $Tv = w_1 + w_2$, with $w_1 \in W_1, w_2 \in W_2$. As such, $v = T^{-1}w_1 + T^{-1}w_2 \in V_1 + V_2$ and therefore $V = V_1 \oplus V_2$. What's left for us to show is that V_1, V_2 are G-invariant.

If $v \in V_i$ then $\Psi_g v = T^{-1}\Phi_g T v$. T must commute by definition and thus $T\Psi_g = \Phi_g T$ for all $g \in G$. However, $Tv \in W_i$ and W_i is G-invariant so $\Phi_g T v \in W_i$ for all v. Thus, $\Psi_g v \in T^{-1}(W_i) = V_i$, i.e. $\Psi_g v \in V_i$ for all v. Thus, V_i is invariant. □

A very similar proof to the preceding one is used to prove both of the following lemmas. For this reason, they are merely stated here.

Lemma 1.2.8. *Let $\Phi : G \longrightarrow GL(V)$ be equivalent to an irreducible representation. Then Φ is irreducible.*

Lemma 1.2.9. *Let $\Phi : G \longrightarrow GL(V)$ be equivalent to a completely reducible representation. Then Φ is completely reducible.*

1.3 Morphisms and Orthogonality

A lot of the concepts thus far resemble well known algebraic concepts, for example, complete reducibility and diagonalisability, or equivalence and isomorphism. It is only natural to wonder in what way we could introduce orthogonality as that is one of the key concepts of linear algebra. We will introduce orthogonality by defining an inner product between representations. However, to be able to truly understand the orthogonality relations, many other concepts must be covered first. These new concepts will be all tied up beautifully in the later chapters (**Chapter 3**).
We have discussed what it means for representations to be equivalent. However, we have not really discussed how two non-equivalent representations might be related. If we think back to the definition of equivalence, it requires the map from one representation to the other be an isomorphism. Loosening that restriction allows for relation between many more representations. These homomorphisms, together with the isomorphisms, of a given representation are called its morphisms.

Definition 1.3.1. *Let $\Phi : G \longrightarrow GL(V)$ and $\Psi : G \longrightarrow GL(W)$ be representations. A **morphism** from Φ to Ψ is a linear map $T : V \longrightarrow W$ such that $T\Phi_g = \Psi_g T$ for all $g \in G$. In other words, the diagram below commutes for all $g \in G$:*

$$\begin{array}{ccc} V & \xrightarrow{\Phi_g} & V \\ {\scriptstyle T}\downarrow & & \downarrow{\scriptstyle T} \\ W & \xrightarrow{\Psi_g} & W \end{array}$$

The space of all homorphisms between Ψ and Φ is denoted as $Hom_G(\Psi, \Phi)$.

Note that if T is invertible we have an equivalence relation. Another important observation is that $T : V \longrightarrow V \;\in Hom_G(\Phi, \Phi) \Leftrightarrow \; T\Phi_g = \Psi_g T \;\forall\; g \in G$. Since $Hom_G(\Phi, \Psi)$ is the space of homomorphisms from V to W (see diagram in **Definition 1.3.1**. What this means is that any morphism satisfying **Definition 1.3.1** is, in fact, in the space of homomorphisms between the codomains' vector spaces of the representations it relates. For the following propositions, assume Φ and Ψ are defined as in **Definition 1.3.1**.

Proposition 1.3.2. *Let $T : V \longrightarrow W$ be in $Hom_G(\Phi, \Psi)$. Then $\ker(T)$ is a G-invariant subspace of V and $T(V) = \mathrm{Im}(T)$ is a G-invariant subspace of W.*

Proof. Let $v \in \ker(T)$, that is, let v be such that $T(v) = 0$. We want to show that $\Phi_g(v) \in \ker(T)$ for all $g \in G$. Now, $T(\Phi_g v) = \Psi_g T(v)$ because $T\Phi_g = \Phi_g T \; \forall \; g \in G$. Since $T(v) = 0$ we have $\Psi_g(0) = T(\Phi_g v)$. Since Ψ_g is in $GL(V)$ for all $g \in G$, Ψ_g acting on the zero vector in V must be zero. This completes the proof for $\ker(T)$ being a G-invariant space. The proof of $Im(T)$ being a G-invariant space follows a very similar structure. □

Proposition 1.3.3. *$Hom_G(\Phi, \Psi)$ is a subspace of $Hom(V, W)$.*

Proof. Let $T_1, T_2 \in Hom_G(\Phi, \Psi)$ and let $c_1, c_2 \in \mathbb{C}$. Now, $(c_1T + c_2T)\Phi_g = c_1T\Phi_g + c_2T\Phi_g = c_1\Psi T + c_2\Psi T = \Psi_g(c_1T + c_2T)$. This completes the proof. □

Now that we have established what a morphism is and what the properties of the space of morphisms are we are ready for Schur's lemma. This lemma talks about the relationship of inequivalent representations with each other as well as the relationship of equal representations with each other. It does not talk about equivalent but not equal representations. It gives a nice entry point into the understanding of the orthogonality relations. We will prove this lemma here, but we will use it later, when proving Shur's orthogonality theorem, which is far more general.

Lemma 1.3.4 (Schur's Lemma)**.** *Let Φ, Ψ be irreducible representations of G, and $T \in Hom_G(\Phi, \Psi)$. Then either T is invertible or $T = 0$. Consequently:*
a) if $\Phi \nsim \Psi$ then $Hom_G(\Phi, \Psi) = 0$.
b) if $\Phi = \Psi$ then $T = \lambda I$ with $\lambda \in \mathbb{C}$.

Proof. Let $\Phi : G \longrightarrow GL(V),\ \Psi : G \longrightarrow GL(W)$ and we also let $T : V \longrightarrow W$ be such that $T \in Hom_G(\Phi, \Psi)$. $T = 0$ is the trivial case and so we will assume $T \neq 0$. Since $ker(T)$ is G-invariant, and both Φ and Ψ are irreducible, $ker(T) = V$ or 0. Since the former brings us back to the trivial case, we assume $ker(T) = 0$ and thus T is one to one. Furthermore, $Im(T)$ is also G-invariant and since $T \neq 0$ it must be that $Im(T) = W$ and thus T is onto. We see that T is invertible and so we are done with the proof of the first statement.
Statement a) follows from the fact that if there is a non-zero T in $Hom_G(\Phi, \Psi)$ then by what we have just proven it is invertible, which means that $\Phi \sim \Psi$, contradiction to our assumption that $\Phi \nsim \Psi$.
To prove statement b) we let λ be the eigenvalue of T (recall we are working over $\mathbb{C}$). Because I is in $Hom_G(\Phi, \Phi)$ so is $(\lambda I - T)$. Furthermore, it is not invertible by definition. However, by the earlier statements of the lemma, all non-zero elements of $Hom_G(\Phi, \Phi)$ are invertible. Thus, $(\lambda I - T) = 0$ and so $T = \lambda I$. □

As we have stated earlier, this lemma talks about two specific cases only. It says what must be true when representations are not equivalent and when they are *equal*. It does not say anything about equivalent representations. If we think of non-equivalent representations as, in a sense, *orthogonal* and equal representations as *on the same line* we see how this lemma talks about orthogonality of representations. It would not be appropriate to try to fit equivalent representations into this geometric visualisation. This is because it could lead us to think some representations are more or less orthogonal than others, which is not true. We will see they are simply either orthogonal or not.

Chapter 2

Overview of Needed Concepts and Notation in Physics

In this chapter we will review some of the relevant physics concepts and their mathematical formulations. We will largely refrain from proofs in this chapter as the mathematical proofs are often not the source of these relationships or are not relevant for our discussion. In creating this section I used notes collected in completion of my physics major. However, the majority of these come from the notes I made in completion of my courses on general relativity and quantum mechanics, which were taught by Dr. Bluhm and Dr. Patton respectively. The textbooks used for those course and in this thesis are: *A Short Course in General Relativity* [3] and *An Introduction to Quantum Mechanics* [6].

In quantum physics, particles are not points or spheres as one may imagine them initially. Rather, they are both a particle and a wave. This dual nature of particles is what is behind all sorts of weirdness that arises in quantum physics. In essence, a particle is described by its wavefunction, which is a probability distribution function representing the fact that a particle has not one, but many locations at which it could be. The reason why particle's position cannot be known exactly is the Heisenberg uncertainty principle:

Definition 2.0.1. *The uncertainty in momentum (Δp) and the uncertainty in the position (Δx) of a particle must obey $\Delta p \Delta x \geq \hbar/2$. Here $\hbar$ is the Planck's constant h divided by 2π.*

This tells us that the more we know about the particles velocity, the less we know about where it is located. There are many interpretations of this fact and a lot of material on this topic can be easily found online, a good text is the one I used in this writing of this section [6]. The only important notion for us is that a particle can be described by a wavefunction.

2.1 Wavefunctions

The wavefunction of a particle is its probability distribution function. Namely, it is a function such that the squared norm of the wavefunction gives the probability of finding a particle in a given interval. These wavefunctions are actually solutions to the Schrödinger equation (SE):

$$\hat{H}\Psi = i\hbar\frac{\partial}{\partial t}\Psi \tag{2.1}$$

where $\hat{H}$ is the Hamiltonian operator and has the form $\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(r,t)$. For simpler situations, e.g. free particle, particle in a box, or a simple harmonic oscillator, it can be solved analytically as the potential energy function $V(r,t)$ is simple in these cases.

Example 4. *For a free particle, $V(r,t) = 0$, so then the solution to the SE (i.e. the wavefunction) is given by $\Psi(x,t) = Ae^{\frac{i(pr-Et)}{\hbar}}$, where A is a complex constant, p is momentum, E is total energy, and t is time. Let us find the probability that the particle is somewhere in space. That is, let us integrate the probability values:*
$\int_{-\infty}^{\infty} |\Psi|^2 dx = \int_{-\infty}^{\infty} \Psi^*\Psi dx = 1.$
This is what we expect as it must exist somewhere in space. $\triangle$

In **Example 4** we have only one possible wavefunction (up to scalar multiple). This is not generally the case. For example the solution to a particle in a one-dimensional box (i.e. in an interval on a line) is $\Psi(x) = \sqrt{2/L}\sin\left(\frac{n\pi}{L}x\right)$ where L is the size of the box and $n = 1, 2, ...$ is the energy level of the particle. As a result of the fact that all wavefunctions for a particular situation are solutions to the same differential equation, they live in the same space, which is a complex linear space and, moreover, is a Hilbert space. We now define an inner product:

Definition 2.1.1. *Suppose Ψ and Φ are in the same linear space of wavefunctions. We define the **inner product** between them by $\langle\Psi|\Phi\rangle = \int_{-\infty}^{\infty} \Psi^*\Phi dx$.*

The $\langle|\rangle$ notation is called the Dirac notation and its left half is called the 'bra' and the right a 'ket'. The 'braket' notation is not limited to the inner product definition. In fact, it is used to keep track of wavefunctions and linear operators that can act on them (e.g. the Hamiltonian $\hat{H}$). For example, we can think of these as not always joined together. Not joined into a 'braket', $|\Psi\rangle$ is simply Ψ, while $\langle\Psi|$ is Ψ^*. Thus, the inner product notation should be seen as a special case were the 'bra' takes on the role of an operator acting on the 'ket'. However, there are many different operators we can use to act on the bra or the ket, or in between of them. As an example, consider the position operator $\hat{x}$. To find the average position of a particle described by Ψ we find $\langle\Psi|\hat{x}|\Psi\rangle = \int_{-\infty}^{\infty} \Psi^* x\Psi dx$. It is always the case that the 'braket' implies integration as shown above. But it is possible to do calculations and have operators act on either just the bra or the ket. To illustrate this, we consider the 'ladder' operators. These are used as follows: $a_+ |\Psi_1\rangle = |\Psi_2\rangle$ and $a_- |\Psi_2\rangle = |\Psi_1\rangle$, where the subscript of the wavefunction indicates its energy level (e.g. different modes of a simple harmonic oscillator or energy levels of the particle in a box problem we discussed). The exact form of the ladder operator would depend on the situation and so we will not include it here. The notation also allows us to see the wavefunctions as elements of a linear space more easily- which allows us to use the power of linear algebra. This theory is very complex and so are the uses of this notation. The text mentioned earlier provides a very good introduction.We will satisfy ourselves with the following facts:

1. *Particles are described by wavefunctions that are solutions to the Schrödinger equation.*

2. *Wavefunctions live in Hilbert spaces with a well defined inner product (**Definition 2.1.1.**).*

3. *All physical quantities in quantum mechanics have an associated hermitian operator that has only real eigenvalues.*

To see how this connects to the theory of representations, we see that the operators despite being abstract mappings between states could, in fact, be Φ_g's for some representation Φ and a group G, while the wavefunctions are vectors in V. One might to wonder why we develop this notation and focus on a specific case of a vector space, even though we have considered the general case already. The reason is simply that these vector spaces and representations (we will see examples later) can give us information about physical objects (like particles). It is not in any way implying the space of wavefunctions is mathematically more interesting than the other spaces. It is merely one of the few 'physical' ones.

2.2 Spin and Spin States

"Spin" is a particular type of angular momentum particles can have. There are several values of spin combinations and states that can exist. We shall use the spin of an electron for illustration. Spin states are eigenstates of the spin operator $\hat{S}^2$ and $\hat{S}_z$ with $\vec{S} = (\hat{S}_x, \hat{S}_y, \hat{S}_z)$. $\hat{S}_z$ can only have two states for an electron: up or $down$. The state in which it is determines where in the atom it can be and what the energy of the system it is a part of is.

The easiest way to represent these two states is simply by the two standard basis vectors of $\mathbb{R}^2$ with $(1, 0)$ being the up state. As such, if an electron is acted on (by some mysterious force) in a way that induces spin change, we could simply represent this situation by a left matrix multiplication by $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Now, this is not how it is done in actuality and there are more matrices that are often used when dealing with spin. Ones we will encounter later are the famous Pauli matrices. These are used when working in situations where multiple particles have spin. They are often combined and create new operators to allow for operations on spin vectors (e.g. $(1, 0)$ as discussed above). For our intentions, we shall recall what they are and that particles can have different spin states - simplest case being only up or down.

Definition 2.2.1. *The Pauli matrices are the following matrices in $GL_2(\mathbb{C})$:*
$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $\sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$, $\sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

We are slowly coming to see how we can use mathematical language to illustrate physical reality. Of course, that is not at all a new concept. However, what is rather interesting is that in these examples we can see how representation theory might come into play. For example, if a system of particles can have only certain type of well known and defined interactions, it could be formulated mathematically to be a group G. Once we have a group, we are either already working with linear algebra structures or are just a representation away from it!

2.3 Tensors

One of the reasons tensors are mentioned in this chapter is the notational differences between physics and mathematics. The easiest way to understand tensors is to start building our understanding and notation from the simplest tensors. Namely, a scalar is a degree zero tensor, and has no special notation associated with it. The n-vectors are degree one tensors and we shall use the following notation: V_α where $\alpha = (1, 2, ...n)$ so that V_α is the whole vector, and V_i for a given i is the i-th entry of the vector. As we could predict, the degree two tensors are n by m matrices and we will use the following notation: $\Gamma_{\alpha\delta}$ where $\alpha = (1, 2, ...n)$ and $\delta = (1, 2, ...m)$. It is still possible to imagine degree three tensors as three dimensional data arrays n by m by k. It is not quite so visually clear what the higher degree tensors are, and it is generally not a good idea to try to $visualize$ them. We can only see in three dimensions. To aid our understanding further, we should see how tensor spaces are created (and why).

Let V and W be two vector spaces. Since we will not encounter any other types of tensors, we can assume these to be both $\mathbb{R}^n$. However, these can be any vector spaces. In this case both are n-dimensional as our goal is more easily illustrated that way, but they need not be. To make it even easier to understand, let us consider a specific case where both V and W are $\mathbb{R}^2$. We can create $\mathbb{R}^4$ via their external product $V \times W$. If we have a map L from this new space to e.g. $\mathbb{R}$ (but could be a different vector space too) that is multilinear, we can eliminate this multilinearity by creating yet another space, via the tensor product $V \otimes W$ such that we can find a linear map $\hat{L}$ from this space to $\mathbb{R}$. This space is defined by the map $F : (V, W) \longrightarrow V \otimes W$, which is designed to allow for the linearity of $\hat{L}$. See diagram below:

$$\begin{tikzcd} \mathbb{R}^2 \times \mathbb{R}^2 \arrow[r, "F"] \arrow[dr, "L"'] & \mathbb{R}^2 \otimes \mathbb{R}^2 \arrow[d, "\hat{L}"] \\ & \mathbb{R} \end{tikzcd}$$

Let $v_1 = \begin{bmatrix} a \\ b \end{bmatrix}$ and $v_2 = \begin{bmatrix} c \\ d \end{bmatrix}$ and let $L(v_1, v_2) = (a + b)(c + d)$. Here we note that L is multilinear. Now we choose an F that will allow us to later construct the linear map $\hat{L}$ to satisfy the diagram above. Let $F(v_1, v_2) = v_1 \otimes v_2 \cong v_1 \cdot v_2^\top = \begin{bmatrix} ac & ad \\ bc & bd \end{bmatrix}$. Thus the space we created using the tensor product, in this case, is the space of all two by two matrices with entries from $\mathbb{R}$. A basis can be given by $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. To complete the example, we let $T = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and we let $\hat{L} : M_{2\times 2}(\mathbb{R}) \longrightarrow \mathbb{R}$ be given by $\hat{L}(T) = a+b+c+d$. This map is linear, which we encourage the reader to check. Thus, our goal of lifting the multilinearity has been accomplished.

Thinking back to our discussion of higher degree tensors, we see that visualising them is not possible, but is also not necessary. We see their form may differ and thus we really shouldn't be too attached to the idea of tensors as being super-vectors. It is more important to understand that they exist to allow linearity and that their indices can be chosen to pick particular entries, analogously to the intuitive lower degree examples below. These examples are not related to our previous example or to each other, they simply illustrate the labelling of the entries:

$$V_\alpha = \begin{bmatrix} V_1 \\ V_2 \\ \dots \\ V_n \end{bmatrix}, T_{\alpha\beta} = \begin{bmatrix} T_{11} & T_{12} & \dots & T_{1m} \\ T_{21} & \dots & & \\ \dots & & & \\ T_{n1} & & & T_{nm} \end{bmatrix}.$$

One last convention we should mention is that there is an additional classification of tensors in their given curvilinear spaces. Depending on whether we choose to use the tangent, contravariant basis or the cotangent, covariant basis we write the tensors with upper or lower indices respectively. For tensors of degree more than one it is possible to have mixed indices.

Tensors, in this notation, are most used in general relativity and fluid dynamics, not the theories we will concern ourselves much with. However, they are important when talking about Lie groups and connecting them to Lie algebras, which we will use.

This chapter should make it more obvious where we may find the theory of representations to be useful. It should also sufficiently prepare us for the examples in the coming chapters.

Chapter 3

Representations of Finite Groups

In this chapter, we will shift our attention towards representations of finite groups and their use in physics. Applications of the representation theory of finite groups are often in systems of finitely many states or particles. We will also prove Schur's orthogonality relations, which in my opinion, is one of the most beautiful theorems in representation theory. For this chapter we return to using the notations and guidance of *Representation of Finite Groups* [8].

3.1 Maschke's Theorem

The codomains $GL(V)$'s of representations are over complex vector spaces V and thus they are equipped with an inner product - the usual inner product between complex numbers. Analogously to unitary matrices, unitary representations preserve the inner product.

Definition 3.1.1. *Let V be an inner product space. A representation $\Phi : G \longrightarrow GL(V)$ is said to be **unitary** if Φ_g is unitary for all $g \in G$, i.e. $\langle \Phi_g v, \Phi_g w\rangle = \langle v, w\rangle$ for all $v, w \in V$. In other words, $\Phi : G \longrightarrow U(V)$.*

Most of the time when we refer to unitary representations we are referring to representations with Φ_g's being unitary matrices. However, we must keep in mind that the Φ_g's are not limited to matrices. To illustrate the definition above, consider the following example:

Example 5. *Let $\Phi : G \longrightarrow GL_1(\mathbb{C})$ (that is, $\Phi : G \longrightarrow \mathbb{C}^*$ and for all $g \in G$ there is some $z \in \mathbb{C}$ such that $\Phi_g = z$). Now, if Φ is unitary then $\langle z * 1, z * 1\rangle = 1$, which means $z\bar{z} = 1$ implying that $|z| = 1$. That is, z is on the unit circle in $\mathbb{C}$ and so unitary representations into $\mathbb{C}^*$ are always into the unit circle.* $\triangle$

Note that **Definition 3.1.1.** is not restricted to finite groups, and neither is the following proposition revealing some key properties of unitary representations.

Proposition 3.1.2. *Let $\Phi : G \longrightarrow GL(V)$ be a unitary representation of a group. Then Φ is either irreducible or decomposable.*

Proof. Let $\Phi : G \longrightarrow GL(V)$ be unitary. Suppose that Φ is not irreducible. Thus, there exists a $W < V$ which is G invariant. $W^{\perp}$ is also a subspace of V. We need to show that $W^{\perp}$ is also G-invariant to show that Φ is decomposable. Now, let $v \in W^{\perp}$ and $w \in W$. $\langle \Phi_g v, w \rangle = \langle \Phi_g^{-1} \Phi_g v, \Phi_g^{-1} w \rangle = \langle v, \Phi_g^{-1} w \rangle = 0$, where we used the fact that Φ is unitary. This shows that $\Phi_g v$ is orthogonal to w and thus is still in $W^{\perp}$, i.e. $W^{\perp}$ is also G-invariant. $\square$

For the case of finite group representations, we have an even more interesting relationship, which is going to largely change what representations we work with when working with finite groups.

Proposition 3.1.3. *Every representation of a finite group G is equivalent to a unitary representation.*

We will prove this proposition momentarily, but we must first introduce a new inner product.

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

This is a true inner product, which we verify here:

$$\begin{aligned}
(c_1 v_1 + c_2 v_2) =& \sum_{g \in G} \langle c_1 v_1 + c_2 v_2, w \rangle \\
=& \sum_{g \in G} \langle c_1 v_1, w \rangle + \langle c_2 v_2, w \rangle \\
=& (c_1 v_1, w) + (c_2 v_2, w) \\
(w, v) =& \sum_{g \in G} \langle w, v \rangle = \sum_{g \in G} \overline{\langle v, w \rangle} \\
=& \overline{(v, w)} \\
(v, v) =& \sum_{g \in G} \langle v, v \rangle \geq 0 \\
(v, v) =& 0 \Rightarrow \langle v, v \rangle = 0 \ \forall \ g \Rightarrow v = 0
\end{aligned}$$

Recall from **Definition 3.1.1.** that a unitary representation preserves inner product in an inner product space. What **Proposition 3.1.3.** states is that we can find an **equivalent** representation preserving an inner product in the codomain's underlying vector space of the given representation.

Proof. To prove the proposition, we let $\Psi : G \longrightarrow GL(V)$ where $dimV = n$. We choose a basis B for V, and let $T : V \longrightarrow \mathbb{C}^n$ be an isomorphism. There must exist one as V and $\mathbb{C}^n$ both have n dimensions. We define $\Phi : G \longrightarrow GL_n(\mathbb{C})$ as $\Phi_g = T\Psi_g T^{-1}$ with respect to B for all $g \in G$.
Now, $(\Phi_h v, \Phi_h w) = \sum_{g\in G}\langle \Phi_g \Phi_h v, \Phi_g \Phi_h v\rangle = \sum_{g\in G}\langle \Phi_{gh} v, \Phi_{gh} w\rangle$. If we let $x = gh$, then as g ranges over G, x ranges over all the elements of G. This is because if $k \in G$ then $g = kh^{-1}$, $x = k$. Therefore, $(\Phi_h v, \Phi_h w) = \sum_{g\in G}\langle \Phi_x v, \Phi_x w\rangle = (v, w)$. Thus Φ preserves (v, w) and is unitary. □

We have just shown that $\Phi \sim \Psi$, where Ψ is unitary. We know unitary representations are either irreducible and decomposable from **Proposition 3.1.2.**. In **Chapter 1** we showed that if one representation in an equivalence class is decomposable, so are all the other representations in the class. The same holds for irreducibility. Thus, the following corollary follows.

Corollary 3.1.4. *Let $\Phi : G \longrightarrow GL(V)$ be a non-zero representation of a finite group. Then Φ is irreducible or decomposable.*

In the preceding corollary, we require the group be finite as that guarantees it is equivalent to a unitary group, which need not be the case for infinite group representations. Now we are ready to state our first theorem of this thesis: Maschke's theorem. It is beautifully concise yet gives us a powerful insight into the structure finite group representations.

Theorem 3.1.5 (Maschke's Theorem). *Every representation of a finite group is completely reducible.*

Proof. We prove the theorem by induction. Let $\Phi : G \longrightarrow GL(V)$ be a representation of a finite G. If $\dim(V) = 1$, Φ is irreducible as there is no proper subspace. Now we assume this is true for $\dim(V) \leq n$. Now we let $\Phi : G \longrightarrow GL(V)$ where $\dim(V) = n + 1$. If it is irreducible than we are done. If it is not, by the preceding corollary it is decomposable. So exist G-invariant V_1, V_2 such that $V_1 \oplus V_2 = V$. Furthermore, since the dimension of V_1, V_2 are less than or equal to n, both are completely reducible by the inductive hypothesis. Thus, $V = V_1 \oplus V_2 = U_1 \oplus U_2 \oplus ...U_s \oplus W_1 \oplus ...W_r$, where $V_1 = U_1 \oplus U_2 \oplus ... \oplus U_s$ and

$V_2 = W_1 \oplus W_2 \oplus ... \oplus W_r$ and $r = \dim(V_2),\ s = \dim(V_1)$. Therefore, Φ is completely reducible. $\square$

3.2 Schur's Orthogonality Relations

With Maschke's theorem we have shown that every representation of a finite group is completely reducible and built up of smaller irreducible subrepresentations. This is important in our study of Schur's Theorem, as this theorem discusses the orthogonality relations of irreducible representations. Furthermore, with this theorem we are also able to discern the relationship between subrepresentations of single representation and see which are orthogonal!

If we let $\Phi : G \longrightarrow GL_n(\mathbb{C})$ be a representation then we observe that Φ_g is a matrix with entries $\Phi_{ij}(g)$ where $0 < i, j \leq n$. That is, i, j refer to the row and column of the given entry, respectively. We write $\Phi_g = (\Phi_{ij}(g))$, where this notation will be used to denote a matrix with entries as given in the parentheses. Note that therefore $\Phi_{ij}(g) \in \mathbb{C}$ for all i, j as given. Thus we have n^2 functions of the form $\Phi_{ij} : G \longrightarrow \mathbb{C}$. Using this notation we can define group algebras in preparation for the statement of Schur's theorem.

Definition 3.2.1. *Let G be a group and define $L(G) = \mathbb{C}^G = \{f | f : G \longrightarrow \mathbb{C}\}$. Then $L(G)$ is an inner product space in which $(f_1 + f_2)(g) = f_1(g) + f_2(g)$, $cf(g) = c(f(g))$, and $\langle f_1, f_2 \rangle = \frac{1}{|G|} \sum_{g \in G} f_1(g)\overline{f_2(g)}$. $L(G)$ is the **group algebra** of G.*

We see that the group algebras really are inner product spaces of functions. The importance of defining these algebras is in the fact that the functions in these inner product spaces are what Schur's theorem uses to determine the orthogonality of representations. We are now ready to state Schur's theorem. Before proving it, however, we will need to work a little more.

Theorem 3.2.2 (Schur's Theorem). *Suppose $\Phi : G \longrightarrow U_n(\mathbb{C})$ and $\Psi = G \longrightarrow U_m(\mathbb{C})$ are irreducible, unitary, and not equivalent to each other. Then:*
I. $\langle \Phi_{ij}, \Psi_{kl} \rangle = 0$
II. $\langle \Phi_{ij}, \Phi_{kl} \rangle = 1/n$ if $i = k,\ j = l$, and 0 otherwise.

Note the theorem talks about the functions Φ_{ij}, which belong to $L(G)$, not the entries in the matrices $\Phi_{ij}(g)$. We start working towards our desired proof with the following proposition, in which we define and show the key properties of a new linear map. We will call this map $T^\sharp$ and as we will see, it helps us understand the effects of a morphism T between two representations.

Proposition 3.2.3. *Let $\Phi : G \longrightarrow GL(V)$ and $\Psi : G \longrightarrow GL(W)$ be representations and suppose that $T : V \longrightarrow W$ is a linear transformation. Then:*
a) $T^\sharp = \frac{1}{|G|}\sum_{g\in G} \Psi_{g^{-1}} T \Phi_g \in Hom_G(\Phi, \Psi)$
b) If $T \in Hom_G(\Phi, \Psi)$ then $T^\sharp = T$.
c) The map $P : Hom(V, W) \longrightarrow Hom(\Phi, \Psi)$ defined by $P(T) = T^\sharp$ is an onto linear map.

Proof. We prove this proposition by direct computations:
a) $T^\sharp \Phi_h = \frac{1}{|G|}\sum_{g\in G} \Psi_{g^{-1}} T \Phi_g \Phi_h = \frac{1}{|G|}\sum_{g\in G} \Psi_{g^{-1}} T \Phi_{gh}$. We let $gh = x$ so that $g^{-1} = hx^{-1}$. Then, $\frac{1}{|G|}\sum_{g\in G} \Psi_{g^{-1}} T \Phi_{gh} = \frac{1}{|G|}\sum_{g\in G} \Psi_h \Psi_{x^{-1}} T \Phi_x = \Psi_h \frac{1}{|G|}\sum_{g\in G} \Psi_{x^{-1}} T \Phi_x = \Psi_h T^\sharp$.

b) $T \in Hom(\Phi, \Psi)$ implies $T\Phi_g = \Psi_g T$. Using this fact we compute that:
$T^\sharp = \frac{1}{|G|}\sum_{g\in G} \Psi_{g^{-1}} T \Phi_g = \frac{1}{|G|}\sum_{g\in G} \Psi_{g^{-1}} \Psi_g T = T$.

c)We check its linearity: $P(c_1T_1 + c_2T_2) = (c_1T_1 + c_2T_2)^\sharp = \frac{1}{|G|}\sum_{g\in G} \Psi_{g^{-1}}(c_1T_1 + c_2T_2)\Phi_g = c_1\frac{1}{|G|}\sum_{g\in G} \Psi_{g^{-1}} T_1 \Phi_g + c_2\frac{1}{|G|}\sum_{g\in G} \Psi_{g^{-1}} T_2 \Phi_g = c_1T_1^\sharp + c_2T_2^\sharp = c_1P(T_1) + c_2P(T_2)$. Now if $T \in Hom_G(\Phi, \Psi)$ then $T = T^\sharp = P(T)$. Thus, P is onto. □

By defining $T^\sharp$ we have defined a certain type of average of the quantity $\Psi_{g^{-1}} T \Phi_g$ over all $g \in G$, which we will now use in the next proposition to express the orthogonality of representations.

Proposition 3.2.4. *Let $\Phi : G \longrightarrow GL(V)$, $\Psi : G \longrightarrow GL(W)$ be irreducible representations of G and let $T : V \longrightarrow W$ be a linear map. Then:*
a) if $\Phi \nsim \Psi$ then $T^\sharp = 0$.
b) if $\Phi = \Psi$ then $T^\sharp = \frac{Tr(T)}{deg(\Phi)} I$

The form of this proposition is reminiscent of **Shur's lemma** from **Chapter 1** and it is indeed similar in its implications too. It is often considered to simply be a different form of the same lemma. It is different in what implications of the lemma it highlights and what proofs it is useful for. In fact, we use Schur's lemma to prove this proposition.

Proof. Statement a) is simply due to the fact that if $\Phi \nsim \Psi$ then by Schur's lemma $Hom_G(\Phi, \Psi) = 0$ and since $P(T)$ is onto, the result follows. To prove statement b) we also use Schur's lemma. We know that if $\Phi = \Psi$, then $T = \lambda I$ for some $\lambda \in \mathbb{C}$. Since $T^\sharp = V \longrightarrow V$ we have $Tr(\lambda I) = \lambda Tr(I) = \lambda \dim(V) = \lambda \deg(\Phi)$. Therefore $T^\sharp = \frac{Tr(T^\sharp)}{\deg(\Phi)} I$. □

These two propositions are the foundations for the Schur's theorem. We now introduce two lemmas which will provide us with a few missing pieces and help us tie our developments into a singular proof.

Lemma 3.2.5. *Let $A \in M_{rn}(\mathbb{C})$ and $B \in M_{ns}(\mathbb{C})$ and $E_{ki} \in M_{nn}(\mathbb{C})$. Then the following formula holds: $(AE_{ki}B)_{lj} = a_{lk}b_{ij}$ where $A = (a_{ij})$ and $B = (b_{ij})$.*

We will not prove this technical lemma as it is primarily notational. An example is more helpful in understanding it. Let us consider the following example.

Example 6. *Let A and B be as given in the lemma. Let $k = 1, i = 2$ so that $E_{ki} = E_{12}$. Then*

$$(AE_{ki}B) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{21} & a_{11}b_{22} \\ a_{21}b_{21} & a_{21}b_{22} \end{bmatrix}.$$

Therefore, $(AE_{ki}B)_{11} = a_{11}b_{21} = a_{lk}b_{ij}$ for the chosen k, j, l, i.
It is straightforward to check this works for all the other l, j pairs and that it also holds if we adjust values of k and i. Though, the latter requires performing the entire calculation again.

△

Now we immediately use results of **Lemma 3.2.5** in the following lemma, which is our last piece needed to prove **Schur's Theorem**.

Lemma 3.2.6. *Let $\Phi : G \longrightarrow U_n(\mathbb{C})$ and $\Psi : G \longrightarrow U_m(\mathbb{C})$ be unitary representations. Let $A = E_{ki} \in M_{mn}(\mathbb{C})$. Then $A^\sharp_{jl} = \langle \Phi_{ij}, \Psi_k l \rangle$.*

Proof. Ψ is unitary and thus $\Psi_{g^{-1}} = \Psi_g^{-1} = \Psi_g^*$. That is $\overline{\Psi_{kl}(g)} = \Psi_{lk} g^{-1}$ (otherwise it would not preserve the inner product). We can now compute:
$A^\sharp_{lj} = \frac{1}{|G|}\sum_{g \in G}(\Psi_{g^{-1}} E_{ki} \Phi_g)_{lj} = \frac{1}{|G|}\sum_{g \in G}(\Psi_{lk}(g^{-1})\Phi_{ij}(g)) = \frac{1}{|G|}\sum_{g \in G}(\overline{\Psi_{kl}(g)}\Phi_{ij}(g)) = \langle \Phi_{ij}, \Psi_{kl} \rangle$.

□

At last, we can prove the orthogonality relations of **Schur's Theorem** and discuss their implications. With the tools and machinery we have built in this section, the proof itself becomes fairly simple.

Proof. Let $A = E_{ki} \in M_{mn}(\mathbb{C})$. By **Proposition 3.2.4** $A^\sharp = 0$ and $A^\sharp_{lj} = \langle \Phi_{ij}, \Psi_{kl} \rangle$. Thus, $\langle \Phi_{ij}, \Psi_{kl} \rangle = 0$. This proves part a).

To prove part b) we let $\Phi = \Psi$ and let $A = E_{ki} \in M_n(\mathbb{C})$ where $dim(V) = n$. Then $A^\sharp = \frac{Tr(E_{ki})}{n} I$, by **Proposition 3.2.4.**. Furthermore, by **Lemma 3.2.6** $A^\sharp = \langle \Phi_{ij}, \Phi_{kl} \rangle$.

If $l \neq j$ then $I_{ij} = 0$ and thus $A^\sharp_{lj} = \langle \Phi_{ij}, \Phi_{kl} \rangle = 0$.

If $i \neq k$ then E_{ki} has only zeros on its diagonal and thus $Tr(E_{ki}) = 0$, which again implies $A^{\sharp}_{lj} = \langle \Phi_{ij}, \Phi_{kl} \rangle = 0$.

If $i = k$ and $j = l$ then E_{ki} has a single 1 on the diagonal and all other entries are zero. Thus, $Tr(E_{ki}) = 1$ and so $A^{\sharp}_{lj} = \langle \Phi_{ij}, \Phi_{kl} \rangle = \frac{1}{n}$. This completes the proof.

□

We now have a concept of orthogonality between representations (and thus a form of geometry!), we also know there are no non-trivial morphisms relating orthogonal representations. Schur's theorem is also a beautiful theorem because of the way it connects many core linear algebra concepts and objects. It makes the step allowing us to now use these concepts not only as analogies, but as definitions and rules. To complete our discussion of Schur's theorem, we present a resulting corollary.

Corollary 3.2.7. *Let Φ be a unitary irreducible representation of G of degree d. Then $\{\sqrt{d_k}\Phi^{(k)}_{ij} | 1 \leq i, j \leq d\}$ is an orthonormal set.*

We see that the corollary follows from Schur's theorem by a renormalisation. Since all Φ_{ij} are in $L(G)$, the whole orthonormal set must be in $L(G)$ too. Noting $dimL(G) = |G|$, we conclude there can be at most $|G|$ orthogonal equivalence classes, the following proposition follows.

Proposition 3.2.8. *Let G be a finite group. Let $\Phi^{(1)}$,...,$\Phi^{(s)}$ be a complete set of representatives of the equivalence classes of irreducible representations of G and let $d_i = deg\Phi^{(i)}$, then the functions $\{\sqrt{d_k\Phi^{(k)}_{ij}} | 1 \leq k \leq s, 1 \leq i, j \leq d_k\}$ form an orthonormal set in $L(G)$ and hence $s \leq d_1^2 + ... + d_s^2 \leq |G|$.*

3.3 Character of a Representation and the Regular Representation

In the previous section we shifted from studying representations to instead studying representation classes. At this point, it should be clear why we consider equivalent representations to be the same representation. We will continue in the study of equivalence classes by introducing the character of a representation, which will lead us to class functions and the regular representation. Class functions are important in understanding the representation in terms of its character, which allows us to determine orthogonality between representations in an easier way.

Definition 3.3.1. *Let $\Phi : G \longrightarrow GL(v)$ be a representation. The* ***character*** *$\chi_\Phi : G \longrightarrow \mathbb{C}$ of Φ is defined by $\chi_\Phi(g) := Tr(\Phi_g)$. The character of an irreducible representation is called an irreducible character.*

We immediately see that the character of a degree one representation is simply the representation itself. The character of a representation gives us information about the dimension of a representation.

Proposition 3.3.2. *Let Φ be a representation of G. Then $\chi_\Phi(1) = \deg \Phi$.*

Here $\deg \Phi$ refers to the degree of representation as discussed in **Chapter 1**. That is, the degree of a representation is the dimension of the vector space over which the codomain (the general linear group) of a representation lies.

Proof. $Tr(\Phi_1) = Tr(I) = \dim(V) = \deg \Phi$. □

Now we can see that if we have a representation Φ that is actually a direct sum of subrepresentations, then the character $\chi_\Phi(1) = \deg \Phi$ must be the sum of the characters of the subrepresentations. This is due to the fact that the degrees of the subrepresentations add to give the degree of the full representation.

Lemma 3.3.3. *Let $\Phi = \Psi \oplus \Upsilon$ then $\chi_\Phi = \chi_\Psi + \chi_\Upsilon$.*

Proof. Follows from **Definition 1.2.1.** and **Example 3**. □

Because equivalence classes helped us find a form of uniqueness between representations, we would like to relate character to these classes. In particular, we will be able to define an equivalence class by its character and thus check for a representation's membership in the given equivalence class.

Proposition 3.3.4. *Let $\Phi : G \longrightarrow GL(V)$ and $\Psi : G \longrightarrow GL(V)$ be equivalent representations of G. Then $\chi_\Phi = \chi_\Psi$.*

Proof. Since the two representations are equivalent, there exists a T in $GL(V)$ such that $\Phi_g = T\Psi_g T^{-1}$ for all $g \in G$. Using the fact that $Tr(AB) = Tr(BA)$, we find that $\chi_\Phi(g) = Tr(\Phi_g) = Tr(T\Psi_g T^{-1}) = Tr(\Psi_g TT^{-1}) = Tr(\Psi_g) = \chi_\Psi(g)$. □

Intuitively, we may already see that since $g = hgh^{-1}$ for $g, h \in G$, then it should hold that $\chi_\Phi(g) = \chi_\phi(hgh^{-1})$ too. This can be proved rather quickly using the definition of a character and the property $Tr(AB) = Tr(BA)$. This is an important property of the characters and so we shall give the functions with the property a name too.

Definition 3.3.5. *A function $f : G \longrightarrow \mathbb{C}$ is called a* ***class function*** *if $f(g) = f(hgh^{-1})$ for all $g, h \in G$, or equivalently if f is constant on conjugacy classes of G. The space of class functions is denoted $Z(L(G))$.*

The notation of the space of class functions coincides with the notation of the center of $L(G)$, which is the group algebra as defined in **Definition 3.2.1**. This is because they actually are the same. We will not show that in detail as it is not pertinent to our discussion. We do claim that $Z(L(G))$ is a subspace of $L(G)$, as that is something we need to be aware of.

Proposition 3.3.6. *$Z(L(G))$ is a subspace of $L(G)$.*

The formal proof of this proposition is omitted, as the result follows immediately from **Definition 3.2.1** and **Definition 3.3.5**. A second fact we need is stated in the form of the following proposition, which tells us how to define a basis for this space of class functions.

Proposition 3.3.7. *The set $B = \{\delta_C | C \in Cl(G)\}$ is a basis for $Z(L(G))$. Consequently, $dim(Z(L(G))) = |CL(G)|$, where C is a conjugacy class of G, $Cl(G)$ is the set of all conjugacy classes of G, and $\delta_C = 1$ if $g \in C$ and 0 otherwise.*

Proof. We see that $B \subseteq Z(L(G))$ as all δ_C's are constant on conjugacy classes, and thus are class functions. B spans $Z(L(G))$ since if $f \in Z(L(G))$ then $f = \sum_{C \in CL(G)} f(C)\delta_C$. To see that B is also orthogonal, we check the following.
Let $C, C' \in CL(G)$ then $\langle \delta_C, \delta_{C'} \rangle = \frac{1}{|G|} \sum_{g \in G} \delta_C \overline{\delta_{C'}} = |C|/|G|$ if $C = C'$ or 0 if $C \neq C'$.
Thus we have a basis for $Z(L(G))$ and we have that $|B| = |Z(L(G))|$, which completes the proof. $\square$

We notice that for an abelian group, the basis B as defined above would be a basis of the entire $L(G)$. Realising this B is a basis of the center of the group algebra is a much more impressive fact than it might seem at first. This is because the group algebra is actually a set of all functions from G to $\mathbb{C}$, which is a very large space and depends on G.
At this point we have shown how the character relates to the representation and also how it relates to other character within the inner product equipped space of characters. Thinking back to Schur's theorem we hope to transfer some of the ideas directly into the language of characters. The following theorem captures these ideas, but it is largely a consequence of Schur's theorem.

Theorem 3.3.8. *Let Φ, Ψ be irreducible representations of G. Then* $\langle \chi_\Phi, \chi_\Psi \rangle = \begin{cases} 1 & \Phi \sim \Psi \\ 0 & \Phi \nsim \Psi \end{cases}$
Thus, the irreducible characters of G form an orthonormal set of class functions.

Proof. Without any loss of generality we can let $\Phi : G \longrightarrow U_n(\mathbb{C})$ and $\Psi : G \longrightarrow U_m(\mathbb{C})$. We can now compute the inner product:

$$\begin{aligned}
\langle \chi_\Phi, \chi_\Psi \rangle =& \frac{1}{|G|} \sum_{g \in G} \chi_\Phi \overline{\chi_{\Psi'}} \\
=& \frac{1}{|G|} \sum_{g \in G} \sum_{i=1}^{n} \sum_{j=1}^{m} \Phi_{ii}(g) \overline{\Psi_{jj}(g)} \\
=& \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{1}{|G|} \sum_{g \in G} \Phi_{ii}(g) \overline{\Psi_{jj}(g)} \\
=& \sum_{i=1}^{n} \sum_{j=1}^{m} \langle \Phi_{ii}(g), \Psi_{jj}(g) \rangle
\end{aligned}$$

At this point we can use Schur's theorem and see that if $\Phi \nsim \Psi$ we have $\langle \Phi_{ii}(g), \Psi_{jj}(g) \rangle = 0 = \langle \chi_\Phi, \chi_\Psi \rangle$. On the other hand, if $\Phi \sim \Psi$ then by **Proposition 3.2.4.** we see that $\Phi = \Psi$. Using Schur's theorem again, we compute: $\langle \chi_\Phi, \chi_\Psi \rangle = \sum_{i=1}^{n} \langle \Phi_{ii}(g), \Phi_{ii}(g) \rangle = \sum_{i=1}^{n} \frac{1}{n} = 1.$ □

An immediate consequence of this theorem is that there can be at most $|CL(G)|$ equivalence classes of irreducible representations of G, because they form an orthonormal basis and are in $Z(L(G))$, which has $dim(Z(L(G)) = |CL(G)|$.

We have now established several properties of irreducible representations and we have also showed all representations of finite groups are completely reducible. It is time to look at whether we can relate our developed theory to representations that are not irreducible but can be reduced. We are interested in extending our theory to these representations as many representations are not irreducible, but are a composition of irreducible representations. First, we define the multiplicity of an irreducible representation within the decomposition. Note that this definition uses the results discussed in the preceding paragraph.

Definition 3.3.9. *If $\Phi \sim m_1\Psi^{(1)} \oplus m_2\Psi^{(2)} \oplus ... \oplus m_s\Psi^{(s)}$, where $\Psi^{(i)}$'s are irreducible, then m_i is the* ***multiplicity*** *of $\Psi^{(i)}$ in Φ. If $m_i > 0$ then $\Psi^{(i)}$ is an* ***irreducible constituent*** *of Φ.*

Using this definition, we state the following theorem:

Theorem 3.3.10. *Let $\Phi^{(1)}, ...\Phi^{(s)}$ be a complete set of representations of the equivalence classes of irreducible representations of G and let $\Phi \sim \Psi^{(1)} \oplus m_2\Psi^{(2)} \oplus ... \oplus m_s\Psi^{(s)}$. Then, $m_i = \langle \chi_\Phi, \chi_{\Psi^{(i)}} \rangle$. Consequently, the decomposition of Φ into irreducible constituents is unique and Φ is determined up to equivalence by its character.*

Proof. From **Lemma 3.3.3.** we know that $\chi_\Phi = m_1\chi_\Psi^{(1)} + ... + m_s\chi_\Psi^{(s)}$ and by the orthogonality relations of characters, we conclude that $\langle \chi_\Phi, \chi_{\Psi^{(i)}} \rangle = m_1\langle \chi_{\Psi^{(1)}}, \chi_{\Psi^{(1)}} \rangle + ... + m_1\langle \chi_{\Psi^{(s)}}, \chi_{\Psi^{(i)}} \rangle = m_i$, which proves the first part of the theorem. The rest is implied by **Proposition 3.3.4.**.

□

An immediate consequence of the theorem and of the orthogonality relations is the following corollary.

Corollary 3.3.11. *A representation is irreducible if and only if $\langle \chi_\Phi, \chi_\Phi \rangle = 1$.*

This is a very useful result, as otherwise it might be quite difficult to determine whether a representation is irreducible. Namely, we would have to check for all possible G-invariant subspaces and see if there are any non-trivial ones. This is also why until now we always assumed or were given an irreducible representation, or we used a very simple one. We no longer have this issue. This is one of the major results of this section and chapter as a whole.

Up to this point, we have observed how representations relate to one another and form spaces. We will now make, or tighten, the connection between properties of the group itself and its representations. We will define the regular representation, which is the last concept in our discussion of finite groups. First, let:

$$\mathbb{C}X = \{\sum_{x \in X} c_x x | c_x \in \mathbb{C}\}$$

such that

$$\sum_{x \in X} a_x x = \sum_{x \in X} b_x x \Leftrightarrow a_x = b_x \ \forall x \in X$$
$$\sum_{x \in X} a_x x + \sum_{x \in X} b_x x = \sum_{x \in X} (a_x + b_x) x$$
$$\langle \sum_{x \in X} a_x x, \sum_{x \in X} b_x x \rangle = \sum_{x \in X} a_x \overline{b_x}.$$

Note that the space $\mathbb{C}X$ is a vector space. We create the vector space from the combinations of elements of X. Now we can define the regular representation.

Definition 3.3.12. *Let G be a finite group. The* ***regular representation*** *of G is the homomorphism $L : G \longrightarrow GL(\mathbb{C}G)$ defined by $L_g \sum_{h\in G} c_h h = \sum_{h\in G} c_h gh = \sum_{x\in G} c_{g^{-1}x} x$ for $g \in G$.*

Since the group is assumed to be finite in this definition, the regular representation must be at least equivalent to a unitary representation. This is satisfied as in fact, it itself is a unitary representation, which we can prove by direct computation.

Proposition 3.3.13. *The regular representation is unitary.*

Proof. To prove the regular representation is unitary, we directly compute
$\langle L_g \sum_{h\in G} c_h h, L_g \sum_{h\in G} k_h h\rangle = \langle \sum_{x\in G} c_{g^{-1}x} x, \sum_{x\in G} c\rangle = \sum_{x\in G} c_{g^{-1}x}\overline{k_{g^{-1}x}}$.
Setting $y = g^{-1}x$ gives the final result:
$\sum c_y \overline{k_y} = \langle \sum_{y\in G} c_y y, \sum_{y\in G} k_y y\rangle$. □

We have explored a lot of properties of finite group representations, and we shall conclude this chapter by looking at how understanding of some of these properties is helpful in applications of representation theory.

3.4 The State Permutation Group, S_3

It is finally time to talk about our first direct example of the application of representation theory to physics. Let us first consider why the permutation group might be interesting to a physicist. What permutation groups do is they permute a set of elements according to some bijective function from the group into itself. Thus, any system in which only a finite set of states is available, such that they can be changed by some (group) action could be described by these groups. Why we would be interested in the representation, as opposed to simply the bijective function that works is that this way we can then find irreducible representations, invariant spaces, etc. All of these could have physical meaning, e.g. G-invariance might correspond to a set of temporally stable states which only interchange with each other (not the other states).

Now let us consider 3 particles in three single particle states i_1, i_2, i_3. We can choose a *normal* order of these states. This order is arbitrary but once fixed it must stay consistent. Most natural is to pick the normal order in which the first particle is in the first state, the second particle in the second, and so on. As such, we can write the normal order as:

$$\Psi(X) = \Psi(x_1, x_2, x_3) = \psi_{i_1}(x_1)\psi_{i_2}(x_2)\psi_{i_3}(x_3) = |i_1 i_2 i_3\rangle$$

Where we can notice the ket notation being used. We can now use it as a vector, so that if the order of the i's is switched, it corresponds to switch in the functional expression. Now, to tie this to an example, these three states could be the three energy sublevels $m = 1, 0, -1$ of an electron shell with the orbital angular momentum of 1. That is, $|i_1 i_2 i_3\rangle = |10-1\rangle$.

An element of a permutation group usually refers to a permutation on the particles, that is $(23)\,|10-1\rangle = |1-10\rangle$. However, we could also be interested in permuting the states. To differentiate between the two, we let the ordinary permutations be denoted as $p()$ and the state permutations as $s()$. The difference between these two is not obvious for the normal order state. But consider a situation where we have already permuted the normal order, e.g. $|i_2 i_1 i_3\rangle$. In that case, $p(23)\,|i_2 i_1 i_3\rangle = |i_2 i_3 i_1\rangle$ while $s(23)\,|i_2 i_1 i_3\rangle = |i_3 i_1 i_2\rangle$. We can represent the permutation group as three-dimensional square matrices:

$\psi(12) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\psi(123) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $\psi(132) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$, $\psi() = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,
$\psi(23) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, $\psi(13) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

We satisfy **Proposition 3.1.3** since this representation is already unitary. By **Maschke's Theorem** it should also be completely reducible or irreducible. It is not irreducible, which we will show below and because of the fact that S_3 has three conjugacy classes we expect three inequivalent irreducible representations - one of the outcomes of this chapter! It is not so clear what the character is here, so we express this representation as matrices (we choose the standard representation) and the kets as proper n-vectors:

It is enough to look at these matrices and the state permutation and see that it is itself just a permutation of this ordinary standard permutation. We use the inner product defined in the section on characters to compute that $\langle \chi_\psi, \chi_\psi \rangle = \frac{1}{6}(1+0+0+9+1+1) = 2$ and by **Corollary 3.3.11** we conclude that the representation ψ is not irreducible. This means that there exist non-trivial G-invariant subspaces. One of these subspaces is the line defined by $(1,1,1)$.

Note that these matrices and vectors represent the permutation and not the states directly. If we look at the mentioned G-invariant space, it is saying that the only triplet of particles that is invariant under the permutation is one in which they are all in the same state (or one particle is in all three states). Now this is not possible with our example as all three states are single particle states and a given particle can only be in one state. Basically, what we found is that identical particles can be interchanged without any problems - not very surprising. It is important we took time looking at the invariance anyway.

Chapter 4

Representations of Infinite Groups

The representations of finite groups we discussed earlier are helpful in developing the theory, but they are a little limited in their physical applications. The motivation of this thesis comes from this chapter, the infinite group representations and their use in particle physics. It is much more difficult to describe a representation of an infinite group than a finite group. However, we will find our way to do that here. Namely, we will focus most of our attention will be on Lie groups, as they truly are the heart of mathematics in particle physics.

This section makes use of some of the notation from **Chapter 2**. It is key to remember the indices in tensors can be chosen from an index list to then represent a particular entry. In development of the Lie group and Lie algebra theory we consulted and used the notation of *Group Representation Theory for Physicists* [2]. For the interested reader, *Representation Theory* [4] is another good resource we consulted. The latter of these two has a more mathematical approach to the subject matter and less focus on in the applications.

4.1 Lie Groups

A Lie group is a special infinite group that also happens to be a differentiable manifold. Manifold is a topological structure that is locally *like* $\mathbb{R}^n$ (or $\mathbb{C}^n$). That is, at least locally we can find a continuous function from the manifold to $\mathbb{R}^n$ (or $\mathbb{C}^n$). To give this object more clarity we state its definition:

Definition 4.1.1. *Let $R_{(a)} = R_{(a^1,a^2,\ldots,a^r)}$ be an element of a group G, where the parameters a^i vary over finite or infinite range and are from $\mathbb{R}$. Then G is a* ***Lie group*** *of* <u>*order*</u> *r if $R_{(a)}$ obeys:*
I. There exists $R_{(a_0)}$ such that $R_{(a_0)}R_{(a)} = R_{(a)}R_{(a_0)} = R_{(a)}$ for all $R_{(a)} \in G$.

II. For all $R_{(a)}$, there exists $R_{(\overline{a})}$ such that $R_{(a)}R_{(\overline{a})} = R_{(a_0)}$.
III. If $R_{(c)} = R_{(a)}R_{(b)}$ then c is in the space of parameters and $c = f(a,b)$, f is a real function.
IV. $[R_{(c)}R_{(a)}]R_{(b)} = R_{(c)}[R_{(a)}R_{(b)}]$.
V. Both c and $\overline{a}$ are analytic functions.

A Lie group is said to be ***compact*** *if its parameters are bounded.*

To give face to the name and to understand the notation we immediately consider an example of a Lie group: $GL(2,\mathbb{R})$. The elements of this group have the form $R_{(a)} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ where a_{ii} are the real paramaters so that the order of this Lie group, as defined in **Definition 4.1.1** is 4. Note that for $GL(2,\mathbb{C})$, the order is 8 as every parameter has a complex and a real part, thus two real parameters for each. The reason why we use this parameter notation is that Lie groups need not be matrices or objects easily described visually. However, it is once again the case that we will be primarily interested in the cases of matrix Lie groups.

If we want to think about morphisms between Lie groups we must consider both their group and manifold structure. Thus a morphism between Lie groups is a map that is a group homomorphism and is differentiable. We also note that every compact Lie group is automatically abelian.

4.2 Lie Algebras

4.2.1 Characterisation Near Identity

We see that the concepts of a set of generators and relations that we had while studying the characters of finite group representations and their orthogonality do not really apply here. However, we can learn a lot about a Lie group from a neighbourhood of the identity. We start by letting $R_{(a)} = R_{(0)} + a^\rho X_\rho + \ldots$ where $X_\rho = (\frac{\partial R_{(a)}}{\partial a^\rho})_{a=0}$. The X_ρ are called **infinitesimal generators**. For a Lie group of order n we will have n generators. In the process of finding these generators we are actually locally linearising the group. This formulation uses the tensor notation.

Now, if a^ρ only differs in one parameter from the identity $R_{(0)}$, then $R_{(a)} = 1 + \varepsilon X_\rho$ and $R_{(b)} = 1 + \varepsilon X_\sigma$. Additionally, $R_{(a)}R_{(b)} = R_{(c)} = 1 + C^\tau X_\tau$, $R_{(b)}R_{(a)} = R_{(c')} = 1 + C'^\tau X_\tau$.

We will use this notation and definition of generators to describe Lie algebras. Similarly to the importance of group algebras in the finite case, these will be very important.

4.2.2 The Structure Constants

The structure constant of a Lie group is a tensor that holds information about the relationships between the infinitesimal generators. To find the structure constant of a Lie group we need to first know how to find the commutator between two generators. Recall that the commutator between two operators, or other objects, is given by $[A, B] = AB - BA$.
We compute:
$[R_{(a)}, R_{(b)}] = R_{(a)}R_{(b)} - R_{(b)}R_{(a)} = R_{(c)} - R_{(c')} = 1 + C^\tau X_\tau - 1 - C'^\tau X_\tau = (C^\tau - C'^\tau)X_\tau$
and we also note:
$[R_{(a)}, R_{(b)}] = \varepsilon^2[X_\rho, X_\sigma]$
which leads us to define **the structure constant** of a Lie group:
$C^\tau_{\rho\sigma} = (C^\tau - C'^\tau)/\varepsilon^2$ so that $[R_{(a)}, R_{(b)}] = \varepsilon^2 C^\tau_{\rho\sigma} X_\tau$.

If we try to visualise these constants, one way is to think of them as n-dimensional vectors with τ being the index determining the position in the vector. The ρ and σ merely show what generators we used to get the constant. These structure constants have the following properties:

I. They are anti-symmetric, $C^\tau_{\rho\sigma} = -C^\tau_{\sigma\rho}$.
II. The following is true: $C^\mu_{\rho\sigma}C^\nu_{\mu\tau} + C^\mu_{\sigma\tau}C^\nu_{\mu\rho} + C^\mu_{\tau\rho}C^\nu_{\mu\sigma} = 0$.

We now observe that the set of generators $\{X_\rho\}$ is closed under linear combinations and multiplications defined by the commutator $[,]$. Thus, it is an algebra. We will call it the Lie algebra to its associated Lie group. It may not be immediately obvious, but we have just reduced our search for irreducible representations of a Lie group to a finite number of elements.

Definition 4.2.1. *A **Lie algebra** is the set of infinitesimal generators of a Lie group, $\{X_\rho\}$, equipped with the commutator $[,]$ operation.*

In physics, Lie algebras often arise very naturally and often the Lie group itself is harder to use and does not match the physical system as well as the Lie algebra. For this reason we mostly deal with the Lie algebras directly. It is appropriate for us to consider a simple example.

Example 7. *We again consider $GL(2,\mathbb{R})$. Its generators are the usual basis matrices $e_{11}, e_{12}, e_{21}, e_{22}$. The commutator and thus the structure constants can be found as follows: $[e_{\alpha\beta}, e_{\gamma\delta}] = \delta_{\beta\gamma}e_{\alpha\delta} - \delta_{\alpha\delta}e_{\gamma\beta}$, where δ's are delta functions. We encourage the reader to check that this indeed works by choosing arbitrary combinations of $\alpha, \beta, \gamma, \delta$ and calculating the result both as a classic matrix commutator and also using the right hand side.* $\triangle$

4.2.3 Lie Algebra and Lie Group Correspondence

As we want to use Lie algebras instead of Lie groups for the representations, such that we will have representations $\Phi : g \longrightarrow V$ where g is the Lie algebra and V is a vector space, we need to show and stress a few parallels. The full discussion and proof of this connection is beyond the scope of this thesis. However, it is a very interesting one and a good explanation of it can be found in *Representation Theory* [4]. The following are the important outcomes.

a) Simple groups (no invariant subgroups) have simple algebras (no invariant subalgebras)
b) Semi-simple groups will have semi-simple algebras (no invar. abelian subalgebras)
c) Subgroups correspond to their own algebras which are subalgebras of the Lie algebra
d) Invariant subgroup has its own invariant algebra that is a subalgebra of the Lie algebra
e) An abelian group will have an abelian algebra
f) Compact Lie groups will have compact Lie Algebras

Note also that the matrices we have been visualising in our examples and discussion are already the representations. The matrices represent the group actions on the vectors in the n-dimensional space they operate in. This is a rather tricky idea and so it is important to fully grasp this before moving onwards to the applications.

4.3 Representations used in Physics

In this section we will take a look (at last) at the most commonly used representations in physics. We have already considered some finite group representations but the most commonly used ones are not finite but infinite (and Lie). There is some overlap between the theories developed in the finite group representation sections and this section. Namely, orthogonality, irreducibility, and invariance are examples of properties that carry over to the infinite case almost unharmed.

Many of these examples, if not for the reminders, would seem to be applications of linear algebra in physics, not really representation theory. However, it is the connection between Lie algebras (the matrices and linear algebra structures) and Lie groups that makes it so that we are still making use of representation theory. In the end, if we only looked at the following examples through the lenses of linear algebra, we would not understand their whole structure and origin.

In this section, there will be some mathematical concepts we didn't cover in detail.

It should be noted that the following applications were not taken from a singular source nor fully developed in this thesis. Rather, they were informed by all of the following texts and articles, and then adjusted to fit the needs of this thesis:
Quantum Theory, Groups, and Representations - An Introduction [10]
Gauge Fields, Knots, and Gravity [1]
Crystallography: Symmetry Groups and Group Representations [5]
Algebraic Topology [7]

4.3.1 The Unitary Group, U(1)

As discussed in **Section 2.1**, particles in physics are described by wavefunctions that live in Hilbert spaces with defined inner product. Computation of the inner product can often tell us important information about the particle, e.g. position. We thus often wish to preserve it when adding additional, position independent characteristics to the particle, e.g. charge. That is, we sometimes wish to create operators that could be used to retrieve some information about the particle without affecting the inner product. If we are to use representation theory, we naturally look to unitary representations.

Let us consider an n-dimensional state space H (where wavefunctions live - also a Hilbert space). Let $U(1)$ be irreducible on H, then it has to be one dimensional. This is true since if a group G is commutative, then for $g, h \in G$ then $\Phi(g)\Phi(h) = \Phi(h)\Phi(g)$. From this we use Schur's Lemma to conclude that all of the $\Phi(h)$ matrices are simply scalars (i.e. one-dimensional).

Since the representation of $U(1)$ is irreducible, one dimensional, and the elements of the group are points on a unit circle, we can express it as $e^{i\theta}$. The whole state space is n-dimensional and so we will need n of these irreducible representations summed together via the external direct sum, which will give us a diagonal matrix. If we give each of these scalar representations a multiplication factor q_i in their exponent to allow for different angles, then we get the final representation:

$$\Phi(e^{iQ\theta}) = \begin{bmatrix} e^{iq_1\theta} & 0 & 0 & ... \\ 0 & e^{iq_2\theta} & 0 & ... \\ ... & & & \\ 0 & ... & 0 & e^{iq_n\theta} \end{bmatrix}$$

Now, this is actually a map between manifolds from $U(1)$ to $GL(n, \mathbb{C})$, where the identity of $U(1)$ is taken to the identity of $GL(n, \mathbb{C})$. That is, Φ is a representation of $U(1)$. Since $U(1)$ is a Lie group we can find a differential of the map, namely $Qie^{iQ\theta}$ where the matrix Q is:

$$Q = \begin{bmatrix} q_1 & 0 & 0 & ... \\ 0 & q_2 & 0 & ... \\ ... & & & \\ 0 & ... & 0 & q_n \end{bmatrix}$$

This matrix is the "generator of the $U(1)$ action on H" [10]. It is also one of the generators of the Lie group near identity (i.e. when $\theta \longrightarrow 0$) as we discussed them in the previous sections. In this example, we will call it the charge operator as we wish to encode charge in it (hence we named it Q), but it could be a different property of a particle that we would wish to encode. The key observation here is that Q is an example of an observable. That is, the wavefunctions in $H_{q_i} \subset H$ will be eigenvectors of this matrix and will have eigenvalues corresponding to the q_i's, where these eigenvalues are values of charge we will physically measure if we experiment on the particles represented by the wavefunctions in the state subspace. It is also another representation where we have the map from $U(1)$ to $GL(n, \mathbb{C})$ and then another map to $GL(n, \mathbb{C})$ ($\Phi(e^{iQ\theta})$ and Q live in the same space). It is important to note that Q is not necessarily unitary, only $\Phi(e^{iQ\theta})$ is.

In this case it is charge, but we could have chosen other physical property that we would wish to encode into a unitary representation (like we encoded Q into $\Phi(e^{i\theta})$) to allow ourselves to perform operations on wavefunctions without changing their inner product. That is, we can do calculations that impact the particle's charge without, say, altering its position.

4.3.2 The Special Unitary Group, SU(2)

One of the first famous applications of representation theory in physics is isospin, now a rather outdated theory in physics. We will talk about it anyway as it can show the ways symmetry is encoded in these representations. The idea of isospin was developed by Heisenberg when he tried to figure out why neutrons and protons are so similar and "hang out" together in the nucleus so much [9]. He came up with the idea that they really are the same particle, which he named nucleon, but in a different isospin state. Namely, the proton is a spin up nucleon, while neutron is a spin down. It is a very simple idea, which is maybe why it turned out to be not exactly correct.

How does this lead us to $SU(2)$? There are two spin states and thus the states of the particle will live in a two dimensional state space H. If we want to use a representation to represent spin change actions we will use some subgroup of the general linear group of the 2D space. Now, note that the strong force (the force that holds nucleus together) acts the same way no matter what state the nucleons are in. This hints that the isospin needs a representation that does not interact with the strong force. The information of the effects of the strong force are found by means of an inner product between the nucleon wavefunctions, thus we will again need a unitary representation. We say that the strong force is **symmetric**, as it treats all states of nucleons the same. Since nucleons have only two different states, we are limited in our choices to the group of unitary matrices of dimension 2, $U(2)$, and its subgroups.

We construct a basis for $U(2)$ by using exponents, not exactly analogously but similarly to previous section. One of the basis matrices is simply the identity matrix exponentiated to give unitary matrix, where θ is some angle that defines the orientation of the representation:

$$U_1 = \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{i\theta} \end{bmatrix}$$

Now, we need three more matrices if we wish to have a set of generators for the group $U(2)$, as that is a 4 dimensional space of 2 by 2 matrices. Three matrices that work are the exponentiated Pauli matrices:

$$U_2 = \begin{bmatrix} 0 & e^{i\theta} \\ e^{i\theta} & 0 \end{bmatrix},\ U_3 = \begin{bmatrix} 0 & e^{\theta} \\ e^{-\theta} & 0 \end{bmatrix},\ U_4 = \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}$$

We have already hinted that the exponentiated matrices come from the Pauli matrices and the identity matrix; this is because those are the generators of the group action on H near identity. That is, they are the differentials of the generators of $U(2)$ near identity, just like the matrix Q was for our representation of $U(1)$. At this point we can simply list the four generators of the group action:

$$\hat{U}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \hat{U}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \hat{U}_3 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \hat{U}_4 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

These are multiplied by i in the differential form, but we drop that as that is simply a scalar multiple. We can leave the multiplication factor out, as it still remains a set of generators of the group action and in this form the eigenvalues of all of these are always real. This is something we need if we are trying to determine an observable quantity - something we can later measure in an experiment. We already have four generators of group action, which could be the observables. They all have determinant equal to one and, in fact, form a basis for generators of group action of $SU(2)$ on H and are a basis for Lie algebra of $SU(2)$ with the bracket operation determining the commutation relations as described in the section on Lie algebras.

Returning to our application, let us consider a situation where spin up nucleon (a proton) is represented by $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and spin down nucleon (neutron) is represented by $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. We will now use our generators of group action to construct operators that will mimic the act of raising and lowering the spin of the nucleons. We will also build the charge operator, which is used to count the charge.

Let $\hat{Q}$ be the charge operator that only "counts" the charged particle and let $\hat{a}_{+/-}$ be the raising/lowering operators that turn a neutron into a proton and vice versa. Note also that if we apply the lowering operator to a neutron (or raising operator to proton) we wind up with no particle.

$$\begin{aligned} \hat{Q} &= \frac{1}{2}(\hat{U}_1 + \hat{U}_4) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ \hat{a}_+ &= \frac{1}{2}(\hat{U}_2 + i\hat{U}_3) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\ \hat{a}_- &= \frac{1}{2}(\hat{U}_2 - i\hat{U}_3) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \end{aligned}$$

This representation is taken to much greater depth in the isospin theory. However, this provides us with a great illustration of the use of the representation of $SU(2)$. Namely, the matrices of exponents form a basis of $SU(2)$, while the Pauli matrices and identity form a representation of the $SU(2)$ group. We used the fact that $SU(2)$ is a Lie group when finding the representation: that is how we knew the number of generators, and that they exist in the given form.

4.3.3 The Rotation Group in 3D, SO(3)

$SO(3)$ is the rotation group in three dimensions. Much like in the previous two applications, we will start by identifying a basis. In this case, it is directly the group $SO(3)$ and its basis is given by the three matrices below and the Identity.

$$O_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin\theta \\ 0 & \sin\theta & 1\cos(\theta) \end{bmatrix} \; O_3 = \begin{bmatrix} \cos(\theta) & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos(\theta) \end{bmatrix}, \; O_4 = \begin{bmatrix} \cos(\theta) & -\sin\theta & 0 \\ \sin\theta & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

Once again, to find the infinitesimal generators we find the differentials of these near identity (as $\theta \longrightarrow 0$). This leaves us with the following infinitesimal generators, i.e. the generators of the group action of $SO(3)$ on the state space H:

$$\hat{O}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; \hat{O}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \; \hat{O}_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \; \hat{O}_4 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Again, these form a basis for the Lie algebra of the $SO(3)$ group, where the bracket determines the operation. The applications for this algebra are numerous and range from very straightforward to complex. This is because what these matrices do is simply rotate basis vectors by $\pi/2$ in their particular defined direction. An example of use is position of atoms in molecules. Often, especially in organic chemistry, the orientation and position of an atom in a molecule matters. Namely, sometimes molecules that are non-superimposable mirror images of each other (have different chirality) have very different properties. Thus, capturing which particular position an atom is in a molecule is important. Using representation theory we can find a way to encode this information into matrices and operators that do not disturb our calculations and inner products.

The representation of this group was also used in the derivation of Dirac equation, which is relativistic adjustment of the Schrödinger equation. Special theory of relativity is a theory that deals with changing reference frames and how that affects observed quantities. A rotation group could be used to capture some of these relativistic notations. In particular, we note that the $\hat{O}_i$ matrices are also unitary and thus we would use them to capture quantities that are invariant under reference frame changes.

In general, we note that once we found a way to build operators from representations of unitary groups, we have allowed ourselves to represent a large variety of invariant physical properties. We listed some examples in this thesis, but beyond these particular examples, it is important to look at the "recipe" for finding these desired operators and what mathematical background and structure we need to do that.

Chapter 5

Conclusion

In this thesis, we explored the theory of representations. We managed to derive and observe many important properties of the group representations, i.e. of the homomorphisms from the group to the general linear groups of our chosen vector space. Namely, we derived and studied Schur's and Maschke's theorems, equivalence relations, we also survey the theory of characters of representations, and Lie group representations. We have uncovered many parallels between linear algebra and representation theory. For example, we saw how diagonalisability of a matrix is analogous to the reducibility of a representation into irreducible subrepresentations.

Close attention was given to the application of representation theory in physics. There are many branches of mathematics that are used in physics. Most of the time, we think of mathematical physics as mathematics related to differential equations and continuous functions. However, in this thesis we explored another large section of mathematical physics, which is one that dominates quantum mechanics and uses primarily linear algebra. As discussed, representation theory takes us to these linear spaces and we managed to use it to encode important information into operators over these spaces.

This thesis is only a quick look at the vast theory of representations and its uses in physics. There are many textbooks developing this theory in greater depth, but in different directions. For the interested reader, *Representation Theory* [4] is a good resource for further exploration of this theory, while *Group Representation for Physicists* [2] provides in depth explanation of many applications.

Bibliography

[1] J. Baez and J. P. Munian. *Gauge Fields, Knots, and Gravity*. World Scientific Publishing, 1994.

[2] J.-Q. Chen, J. Ping, and F. Wang. *Group Representation for Physicists*. World Scientific Publishing, 2002.

[3] J. Foster and J. D. Nightingale. *A Short Course in General Relativity*. Springer New York, 3rd edition, 2006.

[4] W. Fulton and J. Harris. *Representation Theory*. Graduate Texts in Mathematics. Springer-Verlag, 1991.

[5] B. Grenier and R. Ballou. Crystallography: Symmetry groups and group representations. *EDP Sciences*, 2012.

[6] D. J. Griffiths and D. F. Schroeter. *Introduction to Quantum Mechanics*. Cambridge University Press, 3rd edition, 2018.

[7] A. Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.

[8] B. Steinberg. *Representation Theory of Finite Groups: An Introductory Approach*. Universitext. Springer New York, 2011.

[9] R. Webb. Spin's nuclear sibling. *Nature Physics*, 4(1):S8–S9, 2008.

[10] P. Woit. *Quantum Theory, Groups, and Representations - An Introduction*. Springer New York, 2017.