The Modern Representation Theory of the Symmetric Groups

Timothy Cioppa

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements for the degree of Master of Science in
Mathematics [1]

Department of Mathematics and Statistics
Faculty of Science
University of Ottawa

© Timothy Cioppa, Ottawa, Canada, 2012

[1]The M.Sc. program is a joint program with Carleton University, administered by the Ottawa-Carleton Institute of Mathematics and Statistics

Abstract

The goal of this thesis is to first give an overview of the modern approach, using the paper [VO05] of A. Vershik and A. Okounkov, to inductively parametrizing all irreducible representations of the symmetric groups. This theory is then used to answer questions concerning to matrix units $E_T^\lambda \in \mathbb{C}[S_n]$. We index units first by partitions λ, and then by so called standard tableaux T. We also present a new result and discuss future research exploring the connections between this theory and Quantum Information.

Acknowledgements

I would like to thank my supervisor, Dr. Benoit Collins, for his support in writing this thesis, and Naomi for her continued support.

This research was made possible by funding from the National Science and Engineering Research Council of Canada.

Contents

Chapter 1

Introduction

The goal of this thesis is to give an exposition of the recent work of A. Vershik and A. Okounkov on the representation theory of the symmetric groups. In particular, the branching rules of the chain $S_n \subset S_{n+1}$ (a representation theoretic notion) are shown to be equivalent to the more combinatorial branching rules given by inclusion of Young diagrams of n boxes inside those of $n+1$ boxes. This, together with algebraic properties of the group algebra $\mathbb{C}[S_n]$, provides a wealth of information concerning the structure of the irreducible S_n-modules, including their dimensions and potential S_k submodules, for $k \leq n$. The main tools and ideas used in [VO05] and [VO96] to prove this result are:

1. The simplicity of the branching rules of $S_n \subset S_{n+1}$, or equivalently, the commutativity of the centralizer $Z(\mathbb{C}[S_n], \mathbb{C}[S_{n-1}])$.

2. The maximally abelian subalgebra $GZ(n)$ in $\mathbb{C}[S_n]$ which is generated by the centers $Z(\mathbb{C}[S_k])$ for $k \leq n$, and which, as we will show, acts diagonally on all irreducible S_n-modules. The simplicity of the branching prescribes considerable structure on $GZ(n)$, as well as the structure of all irreducible S_n-modules.

3. The Jucys Murphy elements (YJM), defined as $X_k = (1, k) + \cdots + (k-1, k) \in \mathbb{C}[S_n]$ for $k = 2, ...n$, which generate $GZ(n)$. The spectral properties of these elements provides the connection between Young tableaux and irreducible S_n-modules.

4. The Hecke algebra $H(2)$, whose representation theory is used to make the connection between the spectrum of the YJM elements and Young tableaux.

This thesis will be structured as follows: We will begin by giving a detailed review of the necessary results from the representation theory of finite groups, including structure theorems for commutants and centralizers in the group algebra. These results are key to the work of Vershik and Okounkov. The main reference for this material is [CST10]. We will then discuss in detail the results of [VO05], namely that the branching diagram of the symmetric groups agrees with that of Young diagrams. After this, we will give two specific constructions of the irreducible S_n-modules. The first sort are known as Specht modules, and their construction relies heavily on the use of Young diagrams. Then we will give a more abstract construction, still using Young diagrams. The main references are [FH04], [Ful97], and [Sag01]. We explain the results of Vershik and Okounkov before giving specific constructions of the S_n-modules for two reasons. First, we can gain a surprisingly large amount of information about the irreducible S_n-modules without an explicit construction, and without relying on the combinatorics of Young diagrams. Secondly, the construction of Specht modules will appear much more natural after knowing about the branching rules of the symmetric groups.

After this, we will give a review of tensor product, alternating products, and symmetric products, as well as induction and extension problems in algebra and representation theory. Then, using these constructions as well as the results in [VO05], we will discuss problems involving the entanglement of subspaces of a tensor product, again in relation to induction from modules over S_n to modules over S_{n+1}. This section will include a new result concerning entanglement, due to the author.

The appendix will cover the necessary results from group theory and linear functional analysis.

Chapter 2

Representation Theory of Finite Groups

2.1 Basic Theory

In this section we give a review of the basic facts from the representation theory of finite groups which will be necessary for this thesis. Our primary reference is [FH04].

The basic goal of representation theory is to study groups of endomorphisms G of a vector space V, and the connection between the abstract group structure of G, and the vector space structure of V. Conversely, given an abstract group G, one can ask whether G can be viewed as a subgroup of $GL(V)$ for some complex vector space V, and to what extent the group structure of G affects the vector space structure of V.

One then defines a *representation* of a group G as follows:

Definition 1 *A representation of a finite group G is a homomorphism $\delta : G \to GL(V)$ where V is a vector space over a field $\mathbb{F}$. Equivalently, a representation of the group G is a left G-module.*

By a *left G-module*, we mean a vector space V over a field $\mathbb{F}$ together with a map $A : G \times V \to V$ which has the following properties for all $g, h \in G$, $\lambda, \beta \in \mathbb{F}$, and $v, w \in V$:

1. $A(gh, v) = A(g, A(h, v))$.
2. $A(\mathrm{Id}_G, v) = v$.
3. $A(g, \lambda v + \beta w) = \lambda A(g, v) + \beta A(g, w)$.

Intuitively, we think of the elements of G as acting linearly on the vector space V. Every left G-module V gives rise to a homomorphism $\delta : G \to GL(V)$ whereby $\delta(g)v = A(g, v)$, and conversely every homomorphism $\delta : G \to GL(V)$ gives rise to a G-module structure on V, where $A(g, v) = \delta(g)v$. Hence we can freely move between either notion of a representation with no confusion.

In this thesis we will consider the case when G is finite, $\mathbb{F} = \mathbb{C}$ and V is finite dimensional. As we shall see, such representations are completely classifiable.

Note 1 *When talking about a representation $\delta : G \to GL(V)$, it is common to write $g(v)$ or simply gv for $\delta(g)v$ when there is no ambiguity. We then think of an element $g \in G$ as being an endomorphism itself, rather than $\delta(g)$ being an endomorphism. Otherwise we will refer to the representation as (δ, V) where $\delta : G \to GL(V)$ is the homomorphism in question.*

When studying the action of a single endomorphism $T \in GL(V)$ on V, we begin by looking at subspaces which are invariant under T, and we hope that our vector space decomposes into such subspaces. We first classify the structure of invariant subspaces, and then study how the vector space decomposes into the direct sum of such subspaces. When we have an entire group of endomorphisms, we consider subspaces W of our vector space V which are invariant under each endomorphism $g \in G$. That is, whenever $g \in G$ and $w \in W$, we necessarily have that $gw \in W$. Such a subspace is called a *subrepresentation* of G, because we can consider it as a representation in its own right, by restricting the action of G to W. Subrepresentations are also referred to as *G-invariant subspaces*. Two examples of G-invariant subspaces are V itself, and the zero subspace. A representation (not equal to the zero representation) with no other G-invariant subspaces is called *irreducible*.

In order to distinguish between or to compare two representations (δ, V) and (ϵ, W) of a group G , we define the notion of a *G-module homomorphism*. If $\delta : G \to$

$GL(V)$ and $\epsilon : G \to GL(W)$ are two representations of G, a G-module homomorphism from V to W is a linear map $T : V \to W$ such that $T(\delta(g)v) = \epsilon(g)(T(v))$ for every $g \in G$ and $v \in V$. A G-module *isomorphism* is simply a G-module homomorphism which is also an isomorphism of vector spaces. If V and W are two G-modules, we denote the space of G-module homomorphisms from V to W by $\mathrm{Hom}_G(V, W)$. If $V = W$, then we use the notation $\mathrm{End}_G(V)$. G-module homomorphisms are often referred to as *intertwining maps*, or *intertwiners*.

The following proposition, known as *Schur's Lemma*, shows that irreducible representations of a finite group are in some sense the simplest (though not always the least complicated) representations, in that two irreducibles are either isomorphic or completely different. The proposition makes this more precise.

Proposition 1 (Schur's Lemma, [FH04], Proposition 1.7) *Any nonzero homomorphism between irreducible G-modules is an isomorphism. Further, if V is an irreducible G-module over the complex numbers, then $End_G(V)$ is isomorphic, as a complex algebra, to $\mathbb{C}$.*

Proof:

A simple calculation shows that the kernel and image of any G-module homomorphism are G-invariant subspaces. If V and W are both irreducible, and ψ is a non zero G-module homomorphism, $\ker \psi = \{0\}$ and $\mathrm{Im}\, \psi = W$, whence ψ is an isomorphism. Now supposing ψ is a complex G-module endomorphism of V, let λ be an eigenvalue of ψ. Then $\psi - \lambda\,\mathrm{Id}_V$ is a G-module homomorphism with non-trivial kernel, and hence is zero. Thus ψ acts by multiplication by λ. The map which associates to a map this eigenvalue is an algebra isomorphism between $\mathbb{C}$ and $\mathrm{End}_G(V)$.

∎

Note that the second part of Schur's lemma still holds if we replace $\mathbb{C}$ with an arbitrary algebraically closed field $\mathbb{F}$, since we could still find an eigenvalue of ψ in $\mathbb{F}$. The importance of Schur's lemma cannot be overstated. For example, we have the following well known result:

Corollary 1 *Let G be a finite abelian group. Then every complex irreducible G-module is one dimensional.*

Proof:

Let (ρ, V) be a complex irreducible G-module. Since G is abelian, each $\rho(g) : V \to V$ is a G-module homomorphism, and hence acts by scalar multiplication on V. In particular, each subspace of V (including all one dimensional subspaces) is G-invariant. Since V is irreducible, V must be one dimensional.

∎

Later, we will use Schur's lemma to study the structure of $\mathrm{End}_G(V)$. This space is called the *commutant space* of the representation V, and we will study its structure in more detail later in the thesis.

Note 2 *If (δ, V) and (ρ, W) are two representations of G, we will often denote the commutant space* $\mathrm{Hom}_G(V, W)$ *by* $\mathrm{Hom}_G(\delta, \rho)$*, and* $\mathrm{End}_G(V)$ *by* $\mathrm{End}_G(\delta)$*.*

Given two representations (π, V) and (μ, W) of a group G, we define the following representations by linear extension:

1. $(\pi \oplus \mu, V \oplus W)$, where $(\pi \oplus \mu)(g)(v \oplus w) = (\pi(g)v) \oplus (\mu(g)w)$.
2. $(\pi \otimes \mu, V \otimes W)$ where $(\pi \otimes \mu)(g)(v \otimes w) = (\pi(g)v) \otimes (\mu(g)w)$
3. $\mathrm{Hom}_{\mathbb{C}}(V, W)$ becomes a G-module, with $(gT)(v) = \mu(g)(T(\pi(g^{-1})(v)))$
4. (π^*, V^*), where $\pi^*(g)(f)v = f(\pi(g^{-1})v)$.

For a more detailed look at tensor products, refer to Chapter 6. Representation (4) is called the *dual representation*, or the *contragradient* representation of G. Note that with this definition, the Riesz map $\psi : V \to V^*$ given by $\psi(v)(w) = \psi_v(w) = \langle w, v \rangle$ is a G-module (conjugate linear) isomorphism of V and V^*. In fact, the action of G on V^* is induced by the action of G on V together with the action of the Riesz map ψ, so that ψ is a conjugate isomorphism of G-modules. As a result, V is irreducible if and only if V^* is irreducible.

2.1.1 Classifying Representations

In order to classify representations of a finite group G, we define the notion of a *semi-simple* group. We define such a group to be one such that any finite dimensional complex representation V can always be written as $V = V_1^{\oplus a_1} \oplus ... \oplus V_m^{\oplus a_m}$ where the V_i are irreducible, pairwise non isomorphic, and a_i is called the *multiplicity of* V_i *in* V. Such a decomposition of a G-module V into its irreducible factors is called an *isotypic decomposition of* V, and a representation which can be decomposed into irreducible factors is called *completely reducible*. For a semi-simple group G, we will study the representations of G as follows:

1. Understand the structure of all possible irreducible G-modules.
2. If $V = V_1^{\oplus a_1} \oplus ... \oplus V_m^{\oplus a_m}$ where the V_i are pairwise non-isomorphic and irreducible, have a method of finding the multiplicities a_i.

There are some natural questions which must be answered in order to justify this procedure. First, how many possible irreducible representations can a finite group G have? If the answer is possibly infinite, the first step in our procedure may not be feasible. Secondly, is the isotypic decomposition of a G-module unique in any sense? Third, how many finite groups are semi-simple in the above sense? Finally, assuming we have a semi-simple group G, how should we proceed with the decomposition? Luckily, the answers to these questions are finitely many, yes, all, and in a relatively simple fashion. In the next section we will discuss a tool which will be of great use, called *characters*. The characters of a finite group G have a nice geometric structure which turn out to completely classify all finite dimensional representations of G.

First, however, we need to establish that every finite group is semi-simple, and that the decomposition of a representation V is unique (that is, the V_i and a_i are unique up to reordering). A representation (δ, V) of G is called *unitary* if each $\delta(g)$ is a unitary operator on V. As we will soon see, every representation of a finite group can be given an inner product so that the representation is unitary. For this reason, we will assume that all representations of finite groups in this thesis are unitary ones. Later we will show that the same holds for finite dimensional representations of compact topological groups.

Theorem 1 ([FH04] Proposition 1.8) *Every finite dimensional complex representation of a finite group G is completely reducible. If $V = \oplus V_\lambda^{\oplus a_\lambda} = \oplus V_\lambda^{\oplus b_\lambda}$ are two decompositions of such a representation V into irreducible G-modules, then necessarily $a_\lambda = b_\lambda$ for each λ.*

Proof:

Assuming a representation is not irreducible, and W is a subrepresentation, we have a decomposition $V = W \oplus W^\perp$, with respect to the G-invariant inner product $\langle v, w\rangle_G = \sum_{g\in G}\langle gv, gw\rangle$, where the inner product on the right is arbitrary. Since each g is unitary with respect to this inner product, $W^\perp$ is also G-invariant. Thus the first part of the proposition follows by induction on the dimension of V.

By Schur's lemma, any G-module endomorphism of V must map $V_\lambda^{\oplus a_\lambda}$ into $V_\lambda^{\oplus b_\lambda}$ and vice versa. Indeed, if the image of $V_\lambda^{\oplus a_\lambda}$ under some G-module endomorphism intersected (non-trivially) some $V_\beta^{\oplus b_\beta}$ for V_λ and V_β not isomorphic, we would have a non zero intertwining map between non isomorphic G-modules, an impossibility. Applying this dual inclusion to the identity endomorphism gives the proposition.

∎

2.1.2 Character Theory for Finite Groups

In order to understand how to decompose a given finite dimensional representation of a finite group G, the geometric properties of representations will be exploited. The main tool we will use will be the complex valued functions defined on G called the characters of G. The *character* of a representation (π, V) of a finite group is the complex valued function on G defined by $g \mapsto \mathrm{Tr}(\pi(g)) \in \mathbb{C}$, where Tr denotes the trace functional on $GL(V)$. Because of the cyclic property of the trace, that $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ for all operators $A, B \in GL(V)$, the characters lie in $\mathbb{C}_{\text{class}}[G]$, the space of complex functions which are constant on conjugacy classes of G.

The following summarizes the important facts about the characters of a finite group G:

Proposition 2 ([FH04], Proposition 2.1) *Let V and W be finite dimensional complex representations of the finite group G, and let $\chi_V(g) = \mathrm{Tr}(g)$, with a similar notation for χ_W:*

1. $\chi_{V\oplus W}(g) = \chi_V(g) + \chi_W(g)$
2. $\chi_{V\otimes W}(g) = \chi_V(g)\chi_W(g)$
3. $\chi_{V^*}(g) = \overline{\chi_V(g)}$

This proposition together with Theorem 1 allows us, roughly, to define a ring structure on the set $\hat{G}$, which consists of all pairwise non-isomorphic G-modules (modulo equivalence). To be more precise, let $\chi(G)$ be the free abelian group generated by $\hat{G}$, with addition, multiplication, and involution taken with respect to the operations of direct sum, tensor product, and taking the dual of a representation. We call this ring the *representation ring of G*. $\chi(G)$ is an involutive ring, the elements of which are called *virtual representations*. We will not need to study this algebraic structure for the purposes of this thesis.

2.1.3 The Group Algebra $\mathbb{C}[G]$ and the Regular Representations

Two of the most important representations of any finite group G are called the *left and right regular representations of G*. In both cases the underlying vector space is the group algebra $\mathbb{C}[G]$, the free complex vector space over G. Clearly $\mathrm{Dim}(\mathbb{C}[G]) = |G|$. Further, $\mathbb{C}[G]$ becomes a $*$ algebra by defining $g^* = g^{-1}$ and extending by conjugate linearity.

The left regular representation, denoted by $(\delta, \mathbb{C}[G])$ is given by left multiplication, that is, $\delta(a)b = ab$. The right regular representation, denoted $(\rho, \mathbb{C}[G])$, is given by $\rho(a)b = ba^*$. It is an easy exercise to show that the character of the regular representations are given by $\chi(g) = |G|$ if $g = e$ and 0 if $g \neq e$.

The group algebra $\mathbb{C}[G]$ can also be thought of as the space of complex valued functions on G, denoted $\mathbb{C}^G$. Given a function $f : G \to \mathbb{C}$, we get an element of the group algebra (as previously defined), namely $\sum_{g\in G} f(g)g$. Conversely, a linear combination $\sum_{g\in G}\lambda_g g$ induces a function $f(g)=\lambda_g\in\mathbb{C}$. Define a map $\mathbb{C}[G]\to\mathbb{C}^G$ by the rule $g\to\delta_g$, by which we mean the indicator function of the set $\{g\}$. Extending this map by linearity gives a vector space isomorphism $\mathbb{C}[G]\cong\mathbb{C}^G$. However, the group algebra, as its name suggests, is also a complex algebra, with multiplication defined first over basic elements, and then extended by bilinearity. When we move from linear combinations to functions, multiplication becomes convolution. Recall that if $f,g:G\to\mathbb{C}$ are complex valued functions on G, then the *convolution* $(f*g)$ of $f$and g is the function defined by $(f*g)(a)=\sum_{b\in G}f(b)g(b^{-1}a)$. With this multiplication, $\mathbb{C}^G$ is an associative unital algebra. See [CST10], Section 1.5 for a detailed treatment of the group algebra and Fourier transform of a finite group G.

The space of complex valued functions on G has a Hilbert space structure, with the inner product given by $\langle f_1,f_2\rangle=\frac{1}{|G|}\sum_{g\in G}\overline{f_1(g)}f_2(g)$. The algebra $\mathbb{C}[G]$ has dimensional equal to $|G|$ while the space of class functions has dimension equal to the number of conjugacy classes of G, and contains all of the characters of G-representations.

Theorem 2 ([FH04], Proposition 2.30) *With respect to the inner product defined above, the set of irreducible characters forms an orthonormal basis of* $\mathbb{C}_{class}[G]$.

Before we can prove this fact, we need to develop some preliminaries. If (ρ,V) is a representation of G, define the space of G-*invariants* to be the space $V^G=\{v\in V|\ \rho(g)v=v\ \forall g\in G\}$. Clearly this space is the isotypic component of the trivial representation in V. The following lemma will allow us to prove the theorem above.

Lemma 1 ([FH04], Chapter 2.2) *Let V and W be G-modules. Then the following hold:*

1. *With respect to the isotypic decomposition of V, the operator given by $v\mapsto\frac{1}{|G|}\sum_{g\in G}gv$ is projection from V onto V^G. Hence* $\mathrm{Dim}(V^G)=\mathrm{Tr}(\frac{1}{|G|}\sum_{g\in G}g)=\frac{1}{|G|}\sum_{g\in G}\chi(g)$
2. *With respect to the induced actions of G, $Hom_{\mathbb{C}}(V,W)$ and $V^*\otimes W$ are isomorphic G-modules, via the map $[f\otimes w](v)=f(v)w$.*
3. $\mathrm{Hom}(V,W)^G=\mathrm{Hom}_G(V,W)$.

Proof: (**Theorem 2**)
If V and W are irreducible G-modules, then by the above lemma we have that $\langle \chi_V, \chi_W \rangle = \frac{1}{|G|} \sum_{g \in G} \overline{\chi_V(g)} \chi_W(g) = \frac{1}{|G|} \sum_{g \in G} \chi_{V^* \otimes W}(g) = \frac{1}{|G|} \sum_{g \in G} \chi_{\mathrm{Hom}(V,W)}(g) = \mathrm{Dim}(\mathrm{Hom}(V,W)^G)$. By Schur's lemma, this is 1 if V and W are isomorphic, and 0 otherwise, as required.

To prove that the characters form a complete orthonormal set, we will show that if a function is orthogonal to all irreducible characters, it must be identically zero. If f is such a class function, then for any irreducible G-module V, the map $\sum_{g \in G} f(g)g$ is a G-module endomorphism of V. By Schur's lemma, this map is multiplication by $\lambda = \mathrm{Tr}(\sum_{g \in G} f(g)g)/\,\mathrm{Dim}(V) = \sum_{g \in G} f(g)\chi_V(g)/\,\mathrm{Dim}(V) = |G|\overline{\langle f, \chi_{V^*} \rangle}/\,\mathrm{Dim}(V)$. The Riesz Representation Theorem provides a G-module (conjugate linear) isomorphism between V and V^*, whence V^* is irreducible, and so $\lambda = 0$.

Considering $\sum_{g \in G} f(g)g$ as an element of the group algebra acting upon itself by left multiplication, where the elements $g \in G$ are independent, $\sum_{g \in G} f(g)g$ is zero when restricted to any irreducible subrepresentation. By complete reducibility, this implies $\sum_{g \in G} f(g)g = 0$ and hence $f = 0$ as required.

∎

Hence the finite dimensional representations of a finite group G are completely characterized by their multiplicities. Specifically, we have the following:

1. A character χ is irreducible if and only if it is a unit vector. Equivalently, V is irreducible if and only if χ_V is a unit vector.

2. Any character χ can be written uniquely as $\chi = \sum \langle \chi, \chi_{V_i} \rangle \chi_{V_i}$. Moreover, χ is the character of $V = \oplus V_i^{\langle \chi, \chi_{V_i} \rangle}$. Hence, a representation is completely characterized by the Fourier coefficients of its character.

Applying this to the regular representation, whose character we denote χ_R, we see that for any irreducible G-module V, $\langle \chi_R, \chi_V \rangle = \mathrm{Dim}(V)$. This gives the following important and well known result:

Corollary 2 *As a G-module, $\mathbb{C}[G] = \oplus V^{\oplus \mathrm{Dim}(V)}$, indexing over all irreducible G-modules.*

The association of a representation to its character can be thought of an an imbedding of $\chi(G)$ into $\mathbb{C}_{\text{class}}[G] \cong Z(\mathbb{C}[G])$. Further, since the latter space is endowed with an inner product, we can think of the representation ring as also having an inner product. Hence we have a map $\chi : \chi(G) \to Z(\mathbb{C}[G])$ which of course associates to a virtual representation its virtual character in the space of class functions on G. By the properties of the characters, and by our previous results, this map is an injective ring homomorphism. This statement, in fact completely summarizes all of our results into one statement. The geometric structure of group representations we mentioned at the beginning of this section is of course given by the geometric structure of their characters. Irreducible representations can be thought of as orthonormal vectors in a Hilbert space and the decomposition of a representation into irreducible components is equivalent to its orthogonal decomposition in terms of those orthonormal vectors. In fact, representation theory of a fixed group G is equivalent to the Fourier Analysis of its class functions.

As an example we will analyze the representations of $\mathbb{Z}_n = \{\exp(2\pi ik/n)|k = 0, .., n-1\}$. Since this is in particular an abelian group, Schur's lemma implies all irreducible representations are one dimensional, and there is one for each element of the group. Hence an irreducible character is equal to its corresponding representation, and is simply a homomorphism $\mathbb{Z}_n \to \mathbb{C}^*$. The only such homomorphisms possible are $\chi_m(z) = z^m, m = 0, ..., n-1$. Indeed, these are all of the characters of $\mathbb{Z}_n$. Now because $\mathbb{Z}_n$ is abelian, every complex function is a class function. Hence, any function $f : \mathbb{Z}_n \to \mathbb{C}$ can be written as $f(z) = \sum_i \langle f, \chi_i \rangle z^i$. We will later show a similar formula when G is the torus itself. In this way, representation theory provides a canonical way for studying the square integrable functions defined on a compact group.

All of the results in this section provide a fundamental technique for classifying the representations of a finite group, which we now summarize:

1. Determine the conjugacy classes of your finite group G.

2. To each conjugacy class, associate an irreducible representation of G, so that different classes give non-isomorphic representations.

3. Determine the character of each irreducible representation.

4. Given a representation V of G, determine all Fourier coefficients of χ_V with respect to the irreducible characters.

2.1.4 Representation Theory of Compact Groups

In this section we will look at how the representation theory of finite groups also applies to the representation theory of compact topological groups. For a finite group G, understanding a finite dimensional representation came down to finding the Fourier coefficients of its character function. This depended on the fact that the irreducible characters formed an orthonormal basis of the space of class functions on G, and that all finite dimensional representations of a finite group are completely reducible (since, in particular, every representation possesses a G-invariant inner product). This sort of analysis does not work for an arbitrary group however, as there are non-compact groups for which not all finite dimensional representations are completely reducible. (see [FH04], pg. 6)

For compact topological groups, however, the case is very similar to the finite case. By a *compact topological group* we mean a group G equipped with a topology with respect to which G is both compact and Hausdorff. We also require that multiplication $G \times G \to G$ is continuous with respect to the product topology, and inversion $G \to G$ is continuous.

A representation of a compact group is a pair (π, V) where V is a Banach space and π is a continuous homomorphism $\pi : G \to GL(V)$. When V is finite dimensional, we define the characters of V to be the complex valued functions on G defined by $g \mapsto \mathrm{Tr}(\pi(g))$. Clearly, the characters are continuous.

One of the main differences between the finite and compact case is that we replace summation over the group by integrating with respect to a translation invariant Borel probability measure on G. What is required in both cases is the concept of averaging over the group, which is no problem in the finite case, and which is also possible in the compact case by replacing summation with integration. In particular, every compact topological group has a unique Borel probability measure μ with the following properties:

1. μ is translation invariant, in that $\mu(gU) = \mu(Ug) = \mu(U)$ for every $g \in G$ and

every measurable set U.

2. μ is regular, in that $\mu(X) = \inf\{\mu(U)|$ U open with $X \supset U\} = \sup\{\mu(K)|$K compact with $K \supset X\}$ for all measurable X.

3. Every nonzero open set has nonzero measure, and every continuous, non-negative, non-vanishing, and compactly supported function has positive integral.

The measure μ is called the *normalized Haar measure on* G (see, for example, [Coh80], section 9.2, for a construction of Harr measure on a compact group, as well as the proofs of the claims of the proposition above). In general, a Haar measure need not be a probability measure, although all such measures are proportional. Hence we can refer to *the* normalized Haar measure. It will be central to all that follows. Compact groups are examples of groups which are called *amenable*. An amenable group is one which supports a norm 1 linear functional $\alpha : L^\infty(G) \to \mathbb{C}$ which sends positive functions to positive real numbers, which sends the unit to 1, and which is invariant under translation by G. For compact groups, such a map is given by integrating a bounded integrable function against the Haar measure.

In order to analyze representations of a compact group, we begin by reflecting on how we proceeded with finite groups, since finite groups are certainly compact groups, and we would like the theory of compact groups we develop to reduce to the theory we already know of finite groups. In order to prove that all finite dimensional representations of a finite group are completely reducible, we appealed to the fact that we can induce a G-invariant inner product on a given representation V by defining, for $v, w \in V$, $\langle v, w\rangle_G = \frac{1}{|G|}\sum_{g\in G}\langle gv, gw\rangle$ where $\langle , \rangle$ is an inner product on V. Because G may not be finite, or even countable, the above G-invariant inner product is not well defined. However, $\langle v, w\rangle_G = \int_G \langle gv, gw\rangle d\mu$ is a G-invariant inner product on V, and so we have the following:

Theorem 3 ([Bum04], Chapter 2, Proposition 2.2) *All finite dimensional complex representations of a compact group G are completely reducible.*

Proof: If V is a G-module, and U is a submodule, then $U^\perp$ is a G-invariant complimentary subspace. The theorem follows by induction on the dimension of V.

∎

Thus any finite dimensional representation V can be decomposed as $V = \oplus_i V_i^{\oplus a_i}$ where the V_j are the irreducible representations of G, and all but finitely many multiplicities are zero.

2.1.5 Orthogonality of Characters

Just as with the finite case, the irreducible finite dimensional characters of a compact group form an orthonormal system in $C(G)$, the space of continuous functions on G, with respect to the inner product $\langle f_1, f_2 \rangle = \int_G f_1(g)\overline{f_2(g)} d\mu g$. Since all representations are completely reducible, any character χ can be written as $\sum \langle \chi, \chi_V \rangle \chi_V$ where we index over all irreducible representations, and all but finitely many Fourier coefficients are zero. Thus the problem of understanding representations of a compact group reduces to describing the irreducible characters (now possibly infinite in number), and finding the Fourier coefficients of the representation's character with respect to the irreducible characters. The irreducible characters of course form an orthonormal generating set of the involution ring of virtual characters in $C(G)$.

In order to prove the above claims, we have to use slightly more sophisticated techniques. Compare the following proposition to the finite group case:

Proposition 3 *Let (π, V) be a finite dimensional continuous representation of the compact group G. Then the map $v \mapsto \int_g \pi(g) v d\mu(g)$ is projection of V onto V^G. Consequently, $\int_G \chi(g) dg = \mathrm{Dim}(V^G)$.*

Proof: Supposing that $\pi(g)v = v$ for every $g \in G$, we have that $\int_g \pi(g) v d\mu(g) = \int_g v d\mu(g) = v$. For $v \in V$, let $w = \int_G \pi(g) v d\mu(g)$. Then by linearity of each $\pi(g)$, and by the invariance of the Haar measure, we have that $\pi(g)w = w$, and hence $w \in V^G$. Hence the map is indeed projection onto V^G. Taking the trace of the map will hence give the dimension of V^G. Since V is finite dimensional, the trace is $\sum \langle \int_G g e_i d\mu(g), e_i \rangle = \int_G (\sum \langle g e_i, e_i \rangle) d\mu(g)) = \int_G \chi(g) d\mu(g)$ as required. ∎

See [Bum04], Proposition 2.8 for an alternative proof of this Proposition. The proof we provide of this fact is slightly different: in particular, we use the projection onto V^G, so that the only difference between the proof in the finite case is that we integrate rather than use summation.

It is clear that Schur's lemma also holds for finite dimensional complex representations of a compact topological group (see Theorem 2.2, pg 10, of [Bum04]), and so $\mathrm{Dim}(\mathrm{Hom}_G(V, W)) = \mathrm{Dim}(\mathrm{Hom}(V, W)^G)$ is 1 if V and W are isomorphic, and 0 if not. But since $\chi_{\mathrm{Hom}(W,V)} = \chi_V \overline{\chi_W}$, the character orthogonality formula follows.

Because of its importance in the finite group case, our next goal is to generalize the notion of the regular representation to compact groups. Recall that the underlying vector space of the regular representation is $\mathbb{C}[G]$, the group algebra, a complex Hilbert space, which we could interpret as either linear combinations of elements of G, or complex valued functions on G. G acts on the group algebra by left multiplication. When viewed as a function space, multiplication in the group algebra corresponds to convolution of functions.

In order to generalize the group algebra, we first ask that it be a vector space of complex valued functions on G, with inner product $\langle f, g\rangle = \int_G \overline{f(h)} g(h) d\mu(h)$, and multiplication given by $(f * g)(h) = \int_G f(x) g(x^{-1}h) d\mu(x)$. If we ask that the inner product be well defined for all functions, we will require that $|f|^2$ be integrable for all functions f in our group algebra. Thus, for the definition of the group algebra, we will take $L^2(G)$, the Banach space of square integrable functions.

To see how G may act on $L^2(G)$, we look again to the finite case. If f was a complex valued function on G and $g \in G$, we defined $(gf)(h) = f(g^{-1}h)$. This is still well defined when G is compact, so we take this to be the regular action of G on $L^2(G)$. In the same way that a representation of G (a finite group) extends to one of the group algebra $\mathbb{C}[G]$, so too does any representation of a compact group G extend to one of $L^2(G)$ in a similar fashion. Indeed, if G acts on V, we define an action of $L^2(G)$ on V by letting $fv = \int_G f(g) g(v) d\mu(g)$ (see Section 7.2.6 for the definition of such an integral).

There are even more similarities between the representation theory of finite and compact groups. Recall that if G is finite then the group algebra decomposes into the direct sum of all irreducible G-modules, with the multiplicity of V_λ being the dimension d_λ. This is also true for compact groups, and is called the Peter-Weyl Theorem. However, while in the finite case we only had to make use of character theory, the tools needed for the compact case are much more complex, namely the theory of matrix coefficients, compact operators on a Hilbert space, and spectral theory from functional analysis. We will state the generalization without proof (see [Bum04], Theorem 4.3, part 1, section 4).

Theorem 4 (Peter-Weyl Theorem) *Let G be a compact topological group.*

1. *Every unitary representation $\pi : G \to GL(H)$, where H is a Hilbert space, decomposes into the direct sum of finite dimensional, irreducible, unitary representations.*

2. *The group algebra $L^2(G)$ decomposes as $\oplus_\lambda H_\lambda^{\oplus d_\lambda}$, with all irreducible finite dimensional unitary representations appearing with multiplicity $d_\lambda = \mathrm{Dim}(H_\lambda)$.*

In the next section we will apply this theory to the representation theory of the torus, and recover well known results concerning Fourier analysis on S^1.

2.1.6 Irreducible Characters of the Torus

In this section we will prove that the characters of the circle group S^1 are all of the form $\chi_n(z) = z^n$ for $n \in \mathbb{Z}$. Because the circle is an abelian group, every irreducible representation is one dimensional. Hence any irreducible representation of the circle is a continuous homomorphism $\pi : S^1 \to \mathbb{C}^*$.

The following results are presumably well known in Functional Analysis literature (as they are simple to prove, as we shall do), however we were unable to find a specific reference for the results.

Proposition 4 *If $\pi : S^1 \to \mathbb{C}^*$ is an irreducible representation of the circle, then $\pi(S^1) \subset S^1$.*

Proof:

If we had an element $\pi(z)$ with norm greater than 1, then $\pi(z^n) = \pi(z)^n$ would be an unbounded sequence in $\pi(S^1)$, contradiction the compactness of the image of

π. Hence $\pi(S^1)$ is contained in the closed unit disc. Similarly if $\pi(z)$ had norm less than 1, the sequence $\pi(z^n) = \pi(z)^n$ would converge to zero. Since S^1 is compact we can find a subsequence $z^{n_j} \to z \in S^1$ implying, by continuity of π, that $\pi(z) = 0$, a contradiction. Hence π maps the circle to itself. ∎

Proposition 5 *The characters of the circle group S^1 are all of the form $\chi_n(z) = z^n$ for $n \in \mathbb{Z}$.*

Proof: Indeed, any such χ is a character of an irreducible representation of the circle. Conversely, any representation π is of the form $\pi(e^{i\theta}) = e^{i\alpha(\theta)}$ for some function $\alpha : \mathbb{R} \to \mathbb{R}$ by the above proposition. Now $e^{i\alpha(\theta+\beta)} = \pi(e^{i(\theta+\beta)}) = \pi(e^{i\theta}e^{i\beta}) = e^{i\alpha(\theta)}e^{i\alpha(\beta)} = e^{i(\alpha(\theta)+\alpha(\beta))}$ implying that α is linear in $\mathbb{Q}$, hence we can write $\alpha(\theta) = \lambda\theta$ for some $\lambda \in \mathbb{R}$ and for all $\theta \in \mathbb{Q}$. In other words, for all rational points z on the circle, and by continuity of the representation π, for all points z on the circle, $\pi(z) = z^\lambda$. Letting ω be the canonical n^{th} root of unity, we must have that $(\pi(\omega))^n = 1$, hence ω^λ is also an n^{th} root of unity. But this implies that $n\lambda \in n\mathbb{Z}$, so that $\lambda \in \mathbb{Z}$ as required. ∎

If we use the notation $e_n(z) = z^n$ then the elements $\{e_n | n \in \mathbb{Z}\}$ form an orthonormal basis of $L^2(S^1)$ by the Peter-Weyl Theorem. Hence for any square integrable function f we can write $f(z) = \sum_n (f, e_n) z^n$ where $(f, e_n) = \int_{S^1} f(z) z^{-n} d\mu(z)$, integrating against the Haar measure on the circle. This is often proved in a course on Functional Analysis by showing that the e_n constitute a maximally orthonormal family of elements in $L^2(S^1)$. Since $L^2(S^1)$ is of course a Hilbert space, the result follows.

2.2 Commutant Theory

We now return to the representation theory of finite groups. This section is devoted to studying, in greater detail, the structure of a representation V and its commutant $\mathrm{End}_G(V)$. Our main purpose in the section is to be able to view the group algebra $\mathbb{C}[G]$ as a direct sum of full matrix algebras. This is carried out in general for a finite dimensional $*$-algebra in [CST10], Chapter 7, but we will restrict ourselves to the group algebra in order to stay as close to our goal as possible, that goal being the classification of all irreducible representations of the symmetric groups.

2.2.1 Classification of $\mathrm{End}_G(V)$

Recall that if a representation V is irreducible, then $\mathrm{End}_G(V)$ is isomorphic, as a complex algebra, to the complex numbers. In general we can always write V uniquely as $\oplus_\lambda V_\lambda^{\oplus m_\lambda}$ where we index over all possible irreducible G-modules. Recall that for $\lambda \in \widehat{G}$ we mean the irreducible representation (λ, V_λ) represented by the homomorphism $\lambda : G \to GL(V_\lambda)$. Thus our decomposition of V is given by $\oplus_{\lambda\in\widehat{G}} V_\lambda^{\oplus m_\lambda}$. If the underlying homomorphism is $\delta : G \to GL(V)$, this means that there are, for each λ, a set of m_λ operators $I_{\lambda,i}$, $i = 1, \cdots, m_\lambda$ which are necessarily independent and injective, such that:

1. $I_{\lambda,i} \in \mathrm{Hom}_G(\lambda, \delta)$
2. $V_\lambda^{\oplus m_\lambda} \cong \oplus_{i=1}^{m_\lambda} I_{\lambda,i}(V_\lambda)$.

We will call the operators $I_{\lambda,i} \in \mathrm{Hom}_G(\lambda, \delta)$ the *imbedding operators* corresponding to (δ, V). These operators are uniquely determined (up to rescaling and reordering) by the choice of decomposition $V = \oplus V_i^{\oplus a_i}$. Indeed, if T and S are G-module isomorphisms between two irreducible modules V and W, then TS^{-1} and ST^{-1} are multiplication by a scalar.

We will denote by $E_{\lambda,i} \in \mathrm{End}_G(V)$ the orthogonal projection of V onto $I_{\lambda,i}(V_\lambda)$. Note that this notation allows us to keep track of the different components in the isotypic decomposition of V, and also emphasizes that the components in the decomposition of V are *isomorphic* copies of the V_λ. The first important fact we will need is the following proposition.

Proposition 6 ([CST10],Lemma 1.2.5) *The set* $\{I_{\lambda,i} | i = 1, ..., m_\lambda\}$ *is a basis for* $\mathrm{Hom}_G(\lambda, \delta)$*. In particular,* $\mathrm{Dim}(\mathrm{Hom}_G(\lambda, \delta)) = m_\lambda$*.*

Proof:

The operators must be independent by their definition. If $T \in \mathrm{Hom}_G(\lambda, \delta)$ then we can write $T = \sum_i E_{\lambda,i}T$. By Schur's lemma, each $E_{\lambda,i}T = a_i I_{\lambda,i}$ for some $a_i I_{\lambda,i} \in \mathbb{C}$, as required. ∎

Note that this proposition gives an alternative proof to the uniqueness of the isotypic decomposition of a G-module. From this fact we can come up with a basis for $\mathrm{Hom}_G(\delta, \delta)$ which will allow us to think of the space as the direct sum of full matrix algebras.

Denote by $T^{\lambda}_{i,j} \in \mathrm{Hom}_G(\delta, \delta)$ the operator defined by $T^{\lambda}_{i,j}((I_{\lambda,j})(w)) = I_{\lambda,i}(w)$ for $w \in V_\lambda$ and which is zero elsewhere. Then $T^{\lambda}_{i,j}$ simply maps the j^{th} copy of V_λ in V to the i^{th}. Note also that $T^{\lambda}_{k,j}T^{\beta}_{l,m} = \delta_{\lambda,\beta}\delta_{j,l}T^{\lambda}_{k,m}$.

Theorem 5 ([CST10], Theorem 1.2.14) *Let (δ, V) be a G-module, where $V = \oplus_{\lambda \in \widehat{G}} V_\lambda^{\oplus m_\lambda}$. Then we have the following:*

1. *The operators $T^{\lambda}_{i,j}$ constitute a basis of $\mathrm{Hom}_G(\delta, \delta)$.*

2. *As an algebra, $\mathrm{Hom}_G(\delta, \delta) \cong \oplus_{\lambda \in \widehat{G}} M_{m_\lambda}(\mathbb{C})$.*

2.2.2 Representation Theory of $\mathrm{End}(V)$

A substantial amount of information concerning the representation theory of a group algebra relies on the fact that the group algebra can be viewed as the direct sum of full matrix algebras acting by matrix multiplication on the group algebra, with respect to the canonical decomposition discussed in Section 2.1.3, Corollary 2. In this section we will discuss the representation theory of a full matrix algebra $M_n(\mathbb{C})$, and show that the only vector space on which it acts irreducibly is $\mathbb{C}^n$, with respect to matrix multiplication.

Theorem 6 (Burnside's Theorem, [CST10], Section 7.1.2) *Let V be a finite dimensional complex vector space. Then $\mathrm{End}(V)$ acts irreducibly on V by the natural action, and up to isomorphism, V is the only irreducible $\mathrm{End}(V)$-module. Furthermore, $\mathrm{End}(V)$ is the unique subalgebra of $\mathrm{End}(V)$ which acts irreducibly on V.*

This theorem is known as Burnside's Theorem, named after William Burnside. We first note the following lemma before giving a proof of this theorem. The proof of this lemma, as well as its generalization, can be found in [CST10], sections 7.4.1 and 7.4.2.

Lemma 2 *The only two sided ideals in $M_n(\mathbb{C})$ are the zero ideal and $M_n(\mathbb{C})$ itself.*

Proof: (**Burnside's Theorem**)
The action of $\mathrm{End}(V)$ on V is irreducible because given a nonzero element $v \in V$ and any $u \in V$, there is a linear map T such that $T(v) = u$. Hence if a nonzero subspace of V is irreducible for $\mathrm{End}(V)$, it must be V.

Now suppose that $\rho : \mathrm{End}(V) \to \mathrm{End}(W)$ is an irreducible representation of $\mathrm{End}(V)$. Then the kernel of ρ, being a two-sided ideal of $\mathrm{End}(V)$ not equal to $\mathrm{End}(V)$, must be zero. Hence there is some $f \in V^*$, $v \in V, w \in W$ such that $\rho(f \otimes v)w \neq 0$. (Here, $(f \otimes v)(w) := f(w)v$, which establishes a G-module isomorphism between $\mathrm{End}(V)$ and $V^* \otimes V$) But then the map sending x to $\rho(f \otimes x)(w)$ is a nonzero $\mathrm{End}(V)$-module homomorphism between V and W, whence by Schur's lemma, the two spaces are isomorphic. We refer to [CST10], Section 7.1.2, for the proof of the last statement of the theorem.

∎

This result can be extended to algebras of the form $A = \oplus_{j \in J} \mathrm{End}(V_j)$ where J is a finite set, and each V_j is a finite dimensional vector space. Since group algebras have this form, such a theory is extremely important. For $\lambda \in J$ we get a representation of V_λ, which we will simply refer to as $\sigma_\lambda : A \to \mathrm{End}(V_\lambda)$, given by projection of A onto $\mathrm{End}(V_\lambda)$. If J is a singleton then clearly we have only one such projection, which is irreducible as a representation of A. The main result is that for an arbitrary finite collection of indices, these representations exhaust all irreducible A-modules. We will state the following extensions of the previous results without proofs, which can be found in [CST10], Section 7.4.2.

Theorem 7 *Consider an algebra A of the form $A = \oplus_\lambda \mathrm{End}(V_\lambda)$, where we index over a finite set of indices, and each V_λ is a finite dimensional complex vector space.*

1. *Any two sided ideal of A is of the form $\oplus A_\lambda$ where A_λ is either $\{0\}$ or $\mathrm{End}(V_\lambda)$.*

2. *The representations $\sigma_\lambda : A \to \mathrm{End}(V_\lambda)$ defined above constitute all irreducible A-modules, up to isomorphism.*

2.2.3 Further Structure of a G-Module

In this section we will discuss three different decompositions of a G-module V which arise from the action of G. These decompositions further help to understand how G acts on V. This section follows the manuscript [CST10], Section 7.3.3.

Let (ρ, V) be a representation of G. Recall that the decomposition $V \cong \oplus_{\lambda \in \widehat{G}} V_\lambda^{\oplus m_\lambda}$ is equivalent to the decomposition $V \cong \oplus_{\lambda \in \widehat{G}} \oplus_{i=1}^{m_\lambda} I_{\lambda,i}(V_\lambda)$ where the $I_{\lambda,i}$ are the imbedding operators defined in Section 2.2.1. Since, for a fixed λ, these operators are a basis for $\mathrm{Hom}_G(V_\lambda, V)$, we have a decomposition $V \cong \oplus_\lambda(\mathrm{Hom}_G(V_\lambda, V) \otimes V_\lambda)$. The isomorphism is given by the rule $T \otimes v \mapsto T(v)$ for $T \in \mathrm{Hom}_G(V_\lambda, V)$ and $v \in V_\lambda$, of course extended by linearity. Now to see how G acts on V with respect to this decomposition, we take T and v as above, in which case we have:

$$\begin{aligned}
\rho(g)T \otimes v &= \rho(g)T(v) \\
&= T(\lambda(g)v) \\
&= T \otimes \lambda(g)v \\
&= (\mathrm{Id} \otimes \lambda)(g)T \otimes v
\end{aligned}$$

Hence G acts on $\mathrm{Hom}_G(V_\lambda, V) \otimes V_\lambda$ as $\mathrm{Id} \otimes \lambda$.

A representation (ρ, V) induces many other important representations.

1. $\mathrm{Hom}_G(\lambda, \rho)$ becomes an irreducible $\mathrm{End}_G(\rho)$-module. We denote the action by $l(T)S = TS$ for $T \in \mathrm{End}_G(\rho)$ and $S \in \mathrm{Hom}_G(\lambda, \rho)$.

2. V becomes a $\mathrm{End}_G(\rho) \otimes \mathbb{C}[G]$ module, with respect to the action $(l \otimes \rho)(T \otimes g)v = T\rho(g)v = \rho(g)Tv$, for $T \in \mathrm{End}_G(\rho)$ and $g \in G$.

3. $\mathrm{Hom}_G(\lambda, \rho) \otimes V_\lambda$ becomes an irreducible $\mathrm{End}_G(\rho) \otimes \mathbb{C}]G]$ module , with the action $l \otimes \lambda$. By this we mean the normal action of the tensor product of algebras.

The action in (1) of $\mathrm{End}_G(\rho)$ on $\mathrm{Hom}_G(\lambda, \rho)$ is irreducible for the following reason: $\mathrm{End}_G(\rho) \cong \oplus_\lambda M_{m_\lambda}(\mathbb{C}) \cong \oplus_\lambda \mathrm{End}(W_\lambda)$ has only finitely many irreducibles, namely $W_\lambda = \mathbb{C}^{m_\lambda}$. It is clear that $\mathrm{Hom}_G(\lambda, \rho) \cong W_\lambda$ and it is a simple calculation to show that the action of $\mathrm{End}_G(\rho)$ on $\mathrm{Hom}_G(\lambda, \rho)$ is isomorphic to the representation σ_λ from the previous section.

For (2) and (3), we return briefly to the theory of finite group representations. If the groups G and H act irreducibly on the spaces V and W (respectively), then we get an action of $G \times H$ on $V \otimes W$, where $(g, h)(v \otimes w) = (gv) \otimes (hw)$. We call this the *external tensor product* of the modules V and W. A short calculation shows that this representation of $G \times H$ has unit character, and hence is irreducible. In fact, these constitute all irreducible $G \times H$ modules, which can be seen by counting conjugacy classes in $G \times H$.

We clearly have $\mathbb{C}[G \times H] \cong \mathbb{C}[G] \otimes \mathbb{C}[H]$, and so irreducible representations of $\mathbb{C}[G] \otimes \mathbb{C}[H]$ are in correspondence with those of $\mathbb{C}[G \times H]$, which are in correspondence with irreducibles V of G and W of H. Hence it is clear that the representation in (3) is irreducible.

We now summarize all this information into the following theorem.

Theorem 8 ([CST10], Theorem 7.3.4) *Let (ρ, V) be a representation of G.*

1. *$V = \oplus_\lambda(\mathrm{Hom}_G(V_\lambda, V) \otimes V_\lambda)$ is the decomposition of V into irreducible $\mathrm{End}_G(V) \otimes \mathbb{C}[G]$ modules.*

2. *With respect to this decomposition, G acts on V as $\oplus_\lambda \mathrm{Id} \otimes \lambda$, $\mathrm{End}_G(V)$ acts as $\oplus_\lambda l \otimes \mathrm{Id}$, and $\mathrm{End}_G(V) \otimes \mathbb{C}[G]$ acts as $\oplus_\lambda l \otimes \lambda$.*

2.2.4 Structure of the Centralizer $Z(G, H)$

Recall the following facts:

1. As G-modules, $\mathbb{C}[G] = \oplus_\lambda V_\lambda^{\oplus d_\lambda}$, where d_λ is the dimension of V_λ. This comes from character theory.

2. $\mathrm{End}_G(\mathbb{C}[G]) = \oplus_\lambda M_{d_\lambda}(\mathbb{C})$. This comes from the structure of the commutant.

These two facts give the following powerful result, which can be found in [CST10], Section 1.5:

Theorem 9 *As G-modules, and as $*$ algebras, $\mathbb{C}[G] \cong \mathrm{End}_G(\mathbb{C}[G])$. Hence the group algebra $\mathbb{C}[G]$ is (star) isomorphic to the direct sum of full matrix algebras.*

Proof:

We view $\mathbb{C}[G]$ as acting on itself by right multiplication, that is if $a, b \in \mathbb{C}[G]$ then $b \mapsto ba^*$ is a G-module homomorphism. It is straightforward to check that every element of $\mathrm{End}_G(\mathbb{C}[G])$ is of this form.

∎

Now for a subgroup $H \subset G$ we define the *centralizer* $Z(G, H)$ to be the subalgebra $\{a \in \mathbb{C}[G] | ab = ba \; \forall b \in \mathbb{C}[H]\}$. Note that $Z(G, G)$ is simply the center of $\mathbb{C}[G]$.

If V is a representation of the group G and $H \subset G$ is a subgroup, then V is also a representation of H, by restriction. This representation is denoted by $\mathrm{Res}^G_H(V)$, or simply $\mathrm{Res}(V)$ if there is no ambiguity as to G and H.

Recall that if (λ, V_λ) is an irreducible G-module, we denote the dimension of V_λ by d_λ.

Theorem 10 ([CST10], Section 7.5.2) *Let $\mathbb{C}[G] \cong \oplus_{\lambda \in \hat{G}} M_{d_\lambda}(\mathbb{C})$, $\mathbb{C}[H] \cong \oplus_{\rho \in \hat{H}} M_{d_\rho}(\mathbb{C})$, and $\mathrm{Res}^G_H(V_\lambda) \cong \oplus_{\rho \in \hat{H}} W_\rho^{\oplus m(\lambda,\rho)}$. Then we have $Z(G, H) \cong \oplus_{\lambda,\rho} M_{m(\lambda,\rho)}(\mathbb{C})$.*

Proof: If $x \in Z(G, H)$, then we begin by writing $x = \sum_\lambda x_\lambda$ where $x_\lambda \in M_{d_\lambda}(\mathbb{C})$. It's straightforward to show that x lies in the centralizer if and only if each x_λ lies in the centralizer $Z(\sigma_\lambda(\mathbb{C}[G]), \sigma_\lambda(\mathbb{C}[H]))$. However, $\sigma_\lambda(\mathbb{C}[G]) = M_{d_\lambda}(\mathbb{C})$ so that $Z(\sigma_\lambda(\mathbb{C}[G]), \sigma_\lambda(\mathbb{C}[H]))$ is in fact $\mathrm{Hom}_H(V_\lambda) = \oplus_\rho M_{m(\lambda,\rho)}(\mathbb{C})$. The result follows. ∎

We say the subgroup $H \subset G$ is *multiplicity free* if each $m(\lambda, \rho)$ as above is either zero or one. The previous theorem provides a useful method of verifying if a subgroup is multiplicity free:

Corollary 3 *The inclusion $\mathbb{C}[H] \subset \mathbb{C}[G]$ is multiplicity free if and only if $Z(G, H)$ is abelian.*

Proof: The only commutative full matrix algebra is $\mathbb{C}$. Hence $Z(G,H)$ is abelian if and only if each $m(\lambda,\rho) = 1$ or 0. ∎

2.2.5 The Group Algebra and the Fourier Transform

In this section we will study another variation on the problem of viewing the group algebra as the direct sum of full matrix algebras.

Recall that each irreducible representation $\lambda : G \to GL(V_\lambda)$ of a finite group can be extended uniquely to an algebra representation $\lambda : \mathbb{C}[G] \to \mathrm{End}(V_\lambda)$. Hence, we can construct a map called the *Fourier Transform* $\mathcal{F}$ between the group algebra $\mathbb{C}[G]$ and the $*$ algebra $\oplus_\lambda \mathrm{End}(V_\lambda)$, which is commonly denoted $A(G)$. Specifically we have that $\mathcal{F}(g) = \sum_\lambda \lambda(g)$, and we extend by linearity to the group algebra.

Proposition 7 ([CST10], Proposition7.4.9) *The Fourier Transform $\mathcal{F}$ is an algebra isomorphism between $\mathbb{C}[G]$ and $A(G)$, with the inverse transform given by $\mathcal{F}^{-1}(\sum_\lambda A_\lambda)(g) = \frac{1}{|G|}\sum_\lambda d_\lambda \mathrm{Tr}(\lambda(g^{-1})A_\lambda)$.*

To prove this, we recall that both the group algebra of G and $A(G)$ have the same dimension, namely $\sum_\lambda d_\lambda^2$. Hence it suffices to show $\mathcal{F}$ is injective. But if, for an element $x \in \mathbb{C}[G]$, we had that each $\lambda(x)$ was zero, then the image of x under the regular representation would be zero. However the regular representation is injective, and thus x would be zero.

In order to make the formula for the inverse transform more clear, we will introduce the following idea. With respect to the inner product on the group algebra which is linear in the first variable, and with respect to which all permutations orthogonal, norm one, and unitary, we get a tracial functional, which we will denote $\psi(x)$, given by $\psi(x) = \langle x, 1_G\rangle$. This trace has the following properties:

1. $\psi(x)$ is the coefficient of 1_G in x.

2. $x = \sum_g \psi(g^{-1}x)g$.

3. $\psi(x) = \sum_{\lambda\in\hat{G}} \frac{d_\lambda}{|G|}\chi_\lambda(x)$.

Returning to the inverse formula, since $\mathcal{F}$ is invertible, suppose that $A = \sum_\lambda A_\lambda = \mathcal{F}(x)$. Then clearly $A_\lambda = \lambda(x)$. Then calculating $\psi(g^{-1}x)$ we get $\psi(g^{-1}x) = \sum_\lambda \frac{d_\lambda}{|G|}\chi_\lambda(g^{-1}x) = \sum_\lambda \frac{d_\lambda}{|G|} \operatorname{Tr}(\lambda(g^{-1})A_\lambda)$ as required.

If we fix an orthonormal basis of each V_λ, then we have yet another way of viewing the group algebra as block matrices indexed by irreducible representations. Further, if we know the forms of all the matrices $\lambda(g)$ then we can use elementary methods of linear algebra to go between viewing elements of the group algebra as matrices (and vice versa). We will do this in the next section for S_3 and more generally for S_n later.

2.2.6 An Example: Matrix Decomposition of $\mathbb{C}[S_3]$

In this section we will compute the matrix elements $E^\lambda_{i,j}$, with respect to the Fourier Transform discussed in the previous section, for the case of $\mathbb{C}[S_3]$.

S_3 has three irreducible representations of dimensions 1, 2, and 1. The first is the trivial representation U, where each permutation acts as the identity on $\mathbb{C}$. The second is to so called *standard representation* V of S_3. This is a 2 dimensional representation, which is constructed as follows: First note that S_3 acts on $\mathbb{C}^3$ by permuting the entries of a column vector. This is not an irreducible representation, as it contains the one dimensional space spanned by $(1, 1, 1)$, which is isomorphic to the trivial representation. The orthogonal compliment of this space is V, and has a basis $\{e_1 - e_2, e_2 - e_3\}$. That this is an irreducible representation can be verified by checking that its character has length equal to 1. With respect to this basis, S_3 acts as follows:

$$1 \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad (1,2) \mapsto \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$(1,3) \mapsto \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \qquad (2,3) \mapsto \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$$

$$(1,2,3) \mapsto \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \qquad (1,3,2) \mapsto \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}$$

The final irreducible representation is the alternating representation U' where a permutation acts on $\mathbb{C}$ by multiplying by its sign (that is 1 if the permutation is the product of an even number of transpositions, and -1 if it is the product of an odd number of transpositions).

Combining these three representations lets us view S_3 as matrices acting simultaneously on $U \oplus V \oplus U'$, in which we can represent them as the following matrices:

$$1 \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$(1,2) \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$(1,3) \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$(2,3) \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$(1,2,3) \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$(1,3,2) \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Using these matrices, we have a general form for an element of $\mathbb{C}[S_3]$:

$$\begin{pmatrix} a+b+c+d+e+f & 0 & 0 & 0 \\ 0 & a-b+d-f & b-c-e+f & 0 \\ 0 & -c+d+e-f & a+b-d-e & 0 \\ 0 & 0 & 0 & a-b-c-d+e+f \end{pmatrix}$$

where $(a, b, c, d, e, f) \in \mathbb{C}^6$. Any element in $\mathbb{C}[S_3] \cong M_1(\mathbb{C}) \oplus M_2(\mathbb{C}) \oplus M_1(\mathbb{C})$ is thus uniquely determined as a solution to six linear equations in six variables, which corresponds to the six by six matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 0 & 1 & -1 & 0 & -1 & 1 \\ 0 & 0 & -1 & 1 & 1 & -1 \\ 1 & 1 & 0 & -1 & -1 & 0 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{pmatrix}$$

Let us denote this matrix by A. Then the solution to $Ax = e_i$ will determine the group algebra element corresponding to the matrix units. Specifically, we have the following:

1. The solution to $Ax = e_1$ is given by $(1/6, 1/6, 1/6, 1/6, 1/6, 1/6)$, which corresponds to the group algebra element $1/6(1 + (1,2) + (1,3) + (2,3) + (1,2,3) + (1,3,2))$.

2. The solution to $Ax = e_2$ is $(1/3, -1/3, 0, 1/3, -1/3, 0)$, which corresponds to $1/3(1 - (1,2) + (2,3) - (123))$.

3. The solution to $Ax = e_3$ is $(0, 0, -1/3, 1/3, -1/3, 1/3)$, corresponding to $1/3(-(1,3)+(2,3)-(1,2,3)+(1,3,2))$.

4. The solution to $Ax = e_4$ is $(0, 1/3, -1/3, 0, 1/3, -1/3)$, corresponding to $1/3((1,2)-(1,3)+(1,2,3)-(1,3,2))$.

5. The solution to $Ax = e_5$ is $(1/3, 1/3, 0, -1/3, 0, 1/3)$, corresponding to $1/3(1+(1,2)-(2,3)-(1,3,2))$.

6. The solution to $Ax = e_6$ is $(1/6, -1/6, -1/6, -1/6, 1/6, 1/6))$, corresponding to $1/6(1-(1,2)-(1,3)-(2,3)+(1,2,3)+(1,3,2))$.

Note that the element in (1) is indeed projection onto U, that of (6) is projection onto U', and the sum of (2) and (5) gives $2/3(1-2((1,2,3)+(1,3,2))$, which is projection onto the isotypic component of V.

Another result which follows from this will be of greater concern later. Since, in any matrix algebra the span of the diagonal entries $E_{i,i}$ is a maximally abelian subalgebra, we have the following:

Proposition 8 *In* $\mathbb{C}[S_3]$ *the subalgebra spanned by the set* $\{1/6(1+(1,2)+(1,3)+(2,3)+(1,2,3)+(1,3,2)), 1/6(1-(1,2)-(1,3)-(2,3)+(1,2,3)+(1,3,2)), 1/3(1-(1,2)+(2,3)-(123)), 1/3(1+(1,2)-(2,3)-(1,3,2))\}$ *is a maximally abelian subalgebra.*

We can get even more from the results of this section. Recall that a full matrix algebra $M_d(\mathbb{C})$ has only one irreducible module, that is $\mathbb{C}^d$, on which $M_d(\mathbb{C})$ acts naturally. Now consider the subspaces $A_i \subset M_d(\mathbb{C})$, for $i = 1, .., d$, which consists of those matrices which are zero except on the i^{th} column. It is easy to see that $A_i \cong \mathbb{C}^d$ for each i, and that the action of left multiplication of $M_d(\mathbb{C})$ on A_i is isomorphic to the natural action of $M_d(\mathbb{C})$ on $\mathbb{C}^d$. Hence each subspace is irreducible. Now the irreducible $\mathbb{C}[S_3]$-modules are $A_1 = \text{span}\{1/6(1+(1,2)+(1,3)+(2,3)+(1,2,3)+(1,3,2))\}$ which is the trivial representation, and $A_4 = \text{span}\{1/6(1-(1,2)-(1,3)-(2,3)+(1,2,3)+(1,3,2))\}$ which is the alternating representation. However, we also obtain specific constructions of the two occurrences of V in $\mathbb{C}[S_3]$, namely span$\{1/3(1-(1,2)+(2,3)-(123)), 1/3((1,2)-(1,3)+(1,2,3)-(1,3,2))\}$, and span$\{1/3(-(1,3)+(2,3)-(1,2,3)+(1,3,2)), 1/3(1+(1,2)-(2,3)-(1,3,2))\}$.

One of the downsides of this section is that by our choice of matrices representing the permutations acting on each irreducible module, we don't have a $*$ isomorphism between $\mathbb{C}[S_3]$ and $M_1(\mathbb{C}) \oplus M_2(\mathbb{C}) \oplus M_1(\mathbb{C})$. However, in the next section we will develop a method for (theoretically) computing the matrices of S_n acting on each irreducible S_n-module, and in such a way that our isomorphism will be one of $*$ algebras. It would be possible then to repeat our methods of this section and obtain similar formulas for S_n in general.

Chapter 3

The Representation Theory of the Symmetric Groups: An Overview

In the next two chapters, we will give a review of two different approaches to classifying the irreducible representations of the symmetric groups. In particular, we are interested in determining the multiplicities in the isotypic decomposition of an irreducible S_n-module when restricted to the subgroup $\{\sigma \in S_n | \sigma(n) = n\} \cong S_{n-1}$.

The first step in classifying the irreducible S_n-modules involves paramaterizing conjugacy classes in S_n. In both of the approaches we will discuss, conjugacy classes are paramaterized by combinatorial objects called *Young diagrams*. To each Young diagram of size n, say λ, we associate an irreducible S_n-module, denoted V_λ in [VO05] and by S_λ in [Sag01]. We then deduce algebraic properties of the representation V_λ from the combinatorial properties of the diagram λ. One of the most interesting results we will discuss is the isotypic decomposition of S_n-modules when restricted to S_{n-1}. As we will see, it is possible to associate Young diagrams to irreducible S_n-modules in such a way that an S_{n-1}-module V_β is an isotypic component of an irreducible S_n-module V_λ if and only if β *sits inside of* λ in an appropriate and natural sense. This result allows us to understand almost all of the algebraic properties of V_λ in a purely combinatorial way.

The classical approach to classifying the irreducible S_n-modules is focused on the process of associating, to a *particular* diagram λ, an irreducible module S_λ. Doing so relies almost entirely on exploiting the combinatorial properties of the diagram

λ rather than the algebraic properties of S_n. We discuss two such constructions in Chapter 5. Our primary references for these constructions will be [Sag01] and [FH04]. While the constructions are quite technical, they don't require a deep understanding of the symmetric groups, and can be accomplished with only a knowledge of basic linear algebra and, of course, Young diagrams. However, the entire process is quite unintuitive, and since it relies entirely on studying Young diagrams, it does not lend itself to classifying the representations of groups which are structurally similar to S_n.

In particular, the symmetric groups S_n have the following properties which should be used, rather than the combinatorial properties of Young diagrams, in the construction of the irreducible modules and the deduction of the restriction rules:

1. S_n is a *Coxeter group* (see [BB05] for an excellent reference on the subject of Coxeter groups) with generators $s_i = (i, i+1), i = 1, \ldots, n-1$, and these generators should play a central role in the representation theory.

2. $S_1 \subset S_2 \subset S_3 \subset \ldots$ forms an inductive chain of finite groups, and hence the representations should be constructed inductively.

3. Since irreducible S_n-modules are irreducible $\mathbb{C}[S_n]$-modules, and vice-versa, the additional structure of the group algebra $\mathbb{C}[S_n]$ should be exploited in determining the irreducible S_n-modules.

This is carried out in [VO05], and elaborated upon in [CST10], which will be our primary references in Chapter 4. The primary difference between the approach in [VO05] and the classical approach is that, rather than focusing on determining the irreducible representations of a *particular* S_n and determining the branching rules for a particular n, we start by considering a directed graph, called the *Bratteli diagram of the chain* $S_1 \subset S_2 \subset \ldots$, which encodes the structure of the irreducible S_n-modules (for all n) as well as the branching rules. This graph can be defined for any chain of finite groups $G_0 = \{1\} \subset G_1 \subset G_2 \subset \ldots$, and in particular, if we know in advance that the multiplicities are all either 0 or 1 (a concept known as *simplicity* of the Bratteli diagram), even more can said about the Bratteli diagram. The simplicity of the Bratteli diagram of the symmetric groups is shown (see [VO05], Theorem 2.1), as we will see, using only the algebraic properties of $\mathbb{C}[S_n]$, and from this fact it is only a few short steps to determining completely the structure of the Bratteli diagram. Young diagrams do, in fact, reappear after knowing the structure of the Bratteli diagram. However, paraphrasing A. Vershik in the preface of [VO05], *the correspondence between Young diagrams and irreducible modules is given a far more natural explanation than in the classical approach* ([VO05], pg. 2). We refer to the preface and introduction of [VO05] for a similar and more detailed summary of the representation theory of S_n than the one we have provided in this chapter.

Chapter 4

The Vershik and Okounkov Approach

4.1 Introduction and Basic Definitions

This section will give an account of the paper [VO05] of Vershik and Okounkov. The main result is that the rules for canonically decomposing an irreducible S_n-module into irreducible S_{n-1}-modules upon restriction can be understood by studying certain combinatorial objects called *Young diagrams*.

Recall that a conjugacy class of S_n is completely described by the common disjoint cycle structure of its elements. In other words, two permutations are conjugate in S_n if and only if they have the same number of k-cycles, for each k, when decomposed into disjoint cycles. The cycle structure of a permutation in S_n is completely described by an integer vector called a *partition of n*. A partition of n is a sequence of integers $\lambda = (\lambda_1, ..., \lambda_k)$ such that $\lambda_1 \geq ... \geq \lambda_k > 0$ and $\sum_i \lambda_i = n$. We have the following definitions:

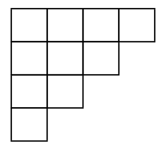
1. The values λ_i are called the *parts of* λ. The *non-trivial parts* of λ are those λ_i which are greater than 1.

2. n is called the *size of* λ.

3. The number of parts of λ is called the *length of* λ and is denoted $l(\lambda)$.

There are several ways of denoting partitions. A common form is by denoting $\lambda = (1^{m_1}2^{m_2}3^{m_3}...)$ where m_i denotes the number of parts of λ which are equal to i. We call m_i the *multiplicity of i in λ*. See [Mac79], Section 1, Chapter 1, for a more detailed treatment of partitions.

The cycle structure of a permutation is commonly represented by a combinatorial object known as a *Young diagram*:



The above object represents the cycle structure $\lambda = (4, 3, 2, 1)$. In a similar way a filling of the above Young diagram represents a specific permutation belonging to the conjugacy class of $\lambda = (4, 3, 2, 1)$. For example, the permutation $(1, 2, 3, 4)(5, 6, 7)(8, 9)$ would look as follows:

1	2	3	4
5	6	7	
8	9		
10			

If we fill the boxes of a diagram λ of size n with the numbers $1, 2, \ldots, n$ without repetition, we obtain an object called a *tableau*. The set of tableaux corresponding to a partition λ will be denoted by $\mathrm{Tab}(\lambda)$, and the set of all tableaux of partitions of size n will be denoted by $\mathrm{Tab}(n)$. Vershik and Okounkov reserve this notation for so called *standard* tableaux, which will be discussed later. However, because standard tableaux play such an important role in the theory, as do regular fillings, we will denote the set of standard tableaux by $\mathrm{Stab}(\lambda)$ and $\mathrm{Stab}(n)$. This notation should not be confused with the stabilizer subgroup of an action.

Two orderings on Young diagrams of a fixed size n are used in the theory of Vershik and Okounkov, as well as in the construction of the irreducible S_n-modules. The first is the *dominance ordering*. If $\lambda = (\lambda_1, ..., \lambda_k)$ and $\mu = (\mu_1, .., \mu_k)$ are two diagrams, we say that the diagram λ dominates the diagram μ, written $\lambda \trianglerighteq \mu$, if $\lambda_1 + \cdots + \lambda_i = \mu_1 + \cdots + \mu_i$ for each $i \geq 1$. In other words, $\lambda \trianglerighteq \mu$ if and only if row 1 of λ has more (or an equal number of) boxes than row 1 of μ, rows 1 and 2 of λ have more (or an equal number of) boxes combined that rows 1 and 2 of μ combined, and so forth for all rows. Note that this is not a total ordering. For a more detailed analysis of this ordering, as well as equivalent formulations, see [CST10], Section 3.6.1. However, we will not need them for the purposes of this thesis.

The second ordering, which is total, is known as the *lexicographic* ordering. With the same notation as above, we say that λ is greater than μ, written $\lambda \geq \mu$ exactly when the first nonzero value of $\lambda_i - \mu_i$ is strictly positive. The following proposition is straightforward:

Proposition 9 ([Sag01], Proposition 2.2.6) *The lexicographic ordering is a refinement of the dominance ordering.*

4.2 Branching Diagrams

The main object of study in [VO05] is an object called the *Bratteli diagram*, which we associate to a chain of finite groups $\{1\} = G_0 \subset G_1 \subset G_2 \subset \cdots$. The Bratteli diagram of such a chain is the the following multigraph:

1. The vertex set consist of all isomorphism classes of complex irreducible G_k modules, for $k \in \mathbb{N}$.

2. Given irreducible representations V of G_n and W of G_{n+1}, there are exactly k directed edges from V to W if V appears exactly k times in the isotypic decomposition of W when restricted to the subgroup G_n.

The Bratteli diagram is a fundamental object in the study of *almost finite dimensional algebras*, or AF algebras. An algebra A is called almost finite dimensional (AF) if there exists a chain $\mathbb{C} \cong A_1 \subset A_2 \subset A_3 \subset \cdots$ of finite dimensional subalgebras of A so that A is the inductive limit of the chain of finite dimensional subalgebras. The group algebra of the infinite symmetric group S_∞ which consists of all permutations of $\mathbb{N}$ which move only finitely many elements can be approximated by the chain of finite dimensional algebras $\mathbb{C}[S_1] \subset \mathbb{C}[S_2] \subset \cdots$, and the representation theory of

S_∞ relies heavily on the representation theory of the algebras $\mathbb{C}[S_n]$. For a detailed treatment of AF algebras, and how the theory is applied to the representation theory of S_∞ see [Ker03] and [Bra72].

The Bratteli diagram of the chain $S_1 \subset S_2 \subset \cdots$, as well as the information which is encoded in it, will be the primary object of study of this thesis.

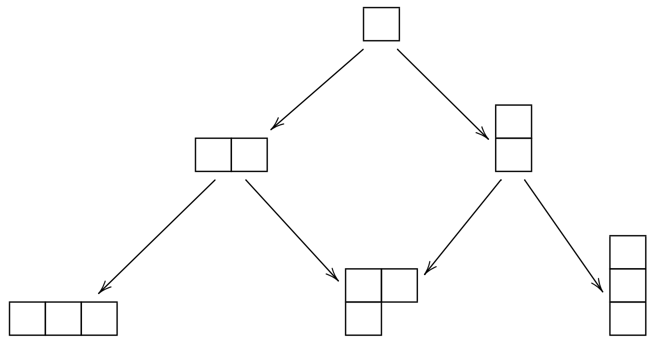
4.3 The Bratteli Diagram of the Symmetric Groups

There is a natural way of viewing S_n as a subgroup of S_{n+1}: if σ permutes the elements $\{1, ..., n\}$ then it also permutes the letters $\{1, ..., n+1\}$ by simply fixing $n+1$. It is clear that the above map is an injective group homomorphism $S_n \to S_{n+1}$. Hence we have an infinite chain of groups $\{1\} = S_1 \subset S_2 \subset S_3 \subset \cdots$, hence we can study the Bratteli diagram of S_n with respect to these embeddings.

S_1 has a single 1 dimensional irreducible representation, corresponding to its unique conjugacy class. This is the trivial representation U_1. S_2 has two, the trivial representation U_2 and the alternating representation U_2', corresponding to the two conjugacy classes of S_2. Since each is one dimensional, they are both isomorphic to U_1 when restricted to S_1. S_3 has the trivial and alternating representations, as well as the standard representation V. It is easy to check using character tables and Frobenius reciprocity (see Section 5.3, Proposition 36) that the first 3 levels of the Bratteli diagram look as follows:

$$\begin{tikzcd}
 & & U_1 \arrow[dl] \arrow[dr] & & \\
 & U_2 \arrow[dl] \arrow[dr] & & U_2' \arrow[dl] \arrow[dr] & \\
U_3 & & V & & U_3'
\end{tikzcd}$$

The goal of the paper of Vershik and Okounkov is to show (although the result was known prior to the paper) that the Bratteli diagram for the symmetric groups is the same as the following branching graph for Young diagrams, called the *Young Graph*, and denoted by $\mathbb{Y}$:



At the n^{th} level of the graph, there are all possible Young diagrams of n boxes, and there are arrows between two diagrams if one can be obtained from the other by adding a single square. By $\mathbb{Y}_n$ we mean the graph consisting of the first n levels of $\mathbb{Y}$. Thus the above graph is actually $\mathbb{Y}_3$, while $\mathbb{Y}$ is the infinite graph consisting of all levels. Since irreducible S_n modules correspond to Young diagrams (see the appendix on S_n, Proposition 34), it is clear that the objects of the above graph $\mathbb{Y}$ correspond to those of the Bratteli diagram. What isn't as clear is that the branching rules are the same.

What is important is that almost all interesting questions concerning the representation theory of the symmetric groups can be answered by studying the combinatorial structure of the Young graph. For example, we can answer the following questions:

1. What is the dimension of the irreducible representation corresponding to a partition λ?

2. What is the irreducible character corresponding to a partition λ?

3. What is the multiplicity of β in λ, if β is a Young diagram of k boxes, and λ is one of n, for $k \leq n$?

4. What are the matrices of the permutations on the irreducible representation corresponding to a partition λ?

We will not address all of these questions in this thesis. For example, the irreducible characters can be studied by the famous *Murnahgan-Nakayama Rule* (see

[CST10] or [VO05], Section 8). However, the philosophy is to replace difficult algebraic questions concerning representations to simpler combinatorial questions concerning Young diagrams. This approach is often called *combinatorial representation theory.*

4.4 Simplicity of the Bratteli Diagram

The first step, and possibly the most important, in showing that the branching graph for Young diagrams corresponds to the Bratteli diagram of the symmetric groups is to show that the Bratteli diagram is *simple*. In terms of the diagram, this means that there is either 0 or 1 arrow going between an irreducible S_n module and an S_{n+1} module. In terms of the representations themselves, simplicity of the branching diagram means that $\mathrm{Dim}(\mathrm{Hom}_{S_{n-1}}(V_\lambda, V_\beta))$ is either 1 or 0 for any S_{n-1} module V_λ and for any S_n module V_β. Again, using the character tables of S_n, as well as Frobenius Reciprocity, this can be demonstrated for the first five levels of the branching graph. However, proving it for the entire Bratteli diagram will require new techniques. Instead of working with representations of the group S_n, we will work with those of the group algebra $\mathbb{C}[S_n]$. The benefit of this is that, for one thing, there is a bijective correspondence between representations of a group and of its group algebra. Secondly, one can consider Bratteli diagrams for inductive families of semi-simple complex algebras, and the inclusion $\mathbb{C}[S_n] \subset \mathbb{C}[S_{n+1}]$ gives such a family. Because representations of G and of $\mathbb{C}[G]$ correspond, so too do the Bratteli diagrams. Further, passing to the group algebra, as we shall see for the symmetric groups, often provides additional algebraic structure with which it is possible to completely classify all irreducible representations.

Given a semi-simple algebra A, and a subalgebra B, denote by $Z(A, B)$ the centralizer algebra $\{a \in A | ab = ba \ \forall b \in B\}$. We recall the following fact, proved earlier in this thesis for a group algebra.

Proposition 10 ([CST10],Theorem 7.5.6) *For a semi-simple algebra and subalgebra, say A and B, the following are equivalent:*

1. *$Z(A, B)$ is commutative.*

2. Any irreducible A-module has simple multiplicities when restricted to B.

Hence in order to show the symmetric groups have simple multiplicities, we need to show is that $Z(\mathbb{C}[S_n], \mathbb{C}[S_{n-1}])$ is commutative. This algebra is denoted by $Z(n-1, 1)$ in [VO05] and so we will use the same notation.

Theorem 11 ([VO05], Theorem 2.1) *$Z(n-1,1)$ is a commutative subalgebra of $\mathbb{C}[S_n]$. Hence the Bratteli diagram of the symmetric groups is simple.*

Proof: $\mathbb{C}[S_n]$ becomes an involutive algebra by defining $\sigma^* = \sigma^{-1}$ and extending by conjugate linearity. Since $\mathbb{C}[S_n]$ is the star complexification (see the appendix for the definition of star complexification) of $\mathbb{R}[S_n]$ and $Z(\mathbb{C}[S_n], \mathbb{C}[S_{n-1}])$ is the star complexification of $Z(\mathbb{R}[S_n], \mathbb{R}[S_{n-1}])$ it suffices to show that any real element of $Z(n-1,1)$ is self conjugate (see the appendix on involutive algebras). Consider $f = \sum_{g \in S_n} \lambda_g g$ where $\lambda_g \in \mathbb{R}$ and $f = gfg^{-1}$ for every $g \in \mathbb{C}[S_{n-1}]$. For any $g \in S_n$ we can find $h \in S_{n-1}$ such that $g^{-1} = hgh^{-1}$. For any such h, $f = hfh^{-1}$, implying that $\lambda_g = \lambda_{g^{-1}}$ for each g. But this implies that $f^* = f$ as required.

∎

The proof of this theorem can be generalized to the Bratteli diagrams of real semi-simple $*$ algebras with a linear base which is:

1. closed under involution; and

2. composed of unitary elements.

See Theorem 2.4 in [VO05] for a precise statement of this result.

4.5 Structure of the Group Algebra $\mathbb{C}[S_n]$

In addition to classifying the Bratteli diagram of the symmetric groups, we would like to construct the representations of the symmetric groups inductively: that is, we would like to construct the representations of S_n using those of S_{n-1}. This is accomplished in [VO05] by studying the algebra structure of $\mathbb{C}[S_n]$. As we will later show, $\mathbb{C}[S_n]$ is generated, as an algebra, by $\mathbb{C}[S_{n-1}]$ and a certain subalgebra studied in an upcoming section, called the Hecke Algebra $H(2)$, with easily classifiable irreducible representations.

In this section we will look more at the algebraic structure of $\mathbb{C}[S_n]$. As we mentioned earlier, we can freely move between studying representations of the group G and of its group algebra $\mathbb{C}[G]$ with no loss of generality. In fact more is true: studying the algebraic properties of the group algebra often reveals information about possible representations of G. The first example we saw of this was the canonical decomposition of the group algebra as $\mathbb{C}[G] = \oplus V^{\oplus \operatorname{Dim}(V)}$, indexing over all irreducible G-modules. Hence not only do all irreducibles appear in $\mathbb{C}[G]$, but their dimensions are encoded in it as well.

A second way that the structure of the group algebra is connected to the representation theory of the group is in the structure of the center of the group algebra (See [FH04], Proposition 2.28):

Proposition 11 *Let G be a finite group. Then the center of $\mathbb{C}[G]$ corresponds to the space of class functions on G (ie. $\sum_{g\in G} \lambda_g g \in Z(\mathbb{C}[G])$ if and only if $g \mapsto \lambda_g$ is conjugacy class independent). Further, under the left regular representation, $Z(\mathbb{C}[G])$ consists of $\mathbb{C}[G]$-module homomorphisms.*

Now consider a G-module $V = V_1^{\oplus a_1} \oplus ... \oplus V_r^{\oplus a_r}$. A short calculation using Schur's lemma (see [FH04], section 2.4) gives that projection of V onto $V_i^{\oplus a_i}$ is given by $p_i(v) = \frac{(\operatorname{Dim} V_i)}{|G|} \sum_{g\in G} \overline{\chi_{V_i}(g)} g(v)$. Specifically we have that $p_i = \frac{(\operatorname{Dim} V_i)}{|G|} \sum_{g\in G} \overline{\chi_{V_i}(g)} g \in Z(\mathbb{C}[G])$ are central idempotents (or central projections) in $\mathbb{C}[G]$.

We will now restrict ourselves to the group algebra $\mathbb{C}[S_n]$ and its algebraic properties. We have already seen one example of how the algebra structure of $\mathbb{C}[S_n]$ determines the representations of S_n: the commutativity of the centralizer $Z(n-1,1)$ implies the simplicity of the Bratteli diagram.

Recall that the chain of finite groups $S_1 \subset S_2 \subset S_3 \subset \dots$ induces a chain of semi-simple algebras $\mathbb{C}[S_1] \subset \mathbb{C}[S_2] \subset \mathbb{C}[S_3] \subset \dots$. Rather than simply studying the algebraic properties of $\mathbb{C}[S_n]$ we will see how the structure of $\mathbb{C}[S_n]$ determines the structure of $\mathbb{C}[S_{n+1}]$. Later, we will use this inductive structure to study representations of the symmetric groups inductively.

Recall that S_n is generated by S_{n-1} and the Coxeter generator s_{n-1} (see the appendix on the symmetric groups). Hence the algebra $\mathbb{C}[S_n]$ is generated by $\mathbb{C}[S_{n-1}]$ and s_{n-1}. In more common notation (as well as the notation used in [VO05]), $\mathbb{C}[S_n] = \langle \mathbb{C}[S_{n-1}], s_{n-1} \rangle$. The fact that $s_j^2 = 1$ implies, in particular, that $\mathbb{C}[S_n] = \langle \mathbb{C}[S_{n-1}], \mathbb{C}[S_2] \rangle$, where $\mathbb{C}[S_2] \cong \langle 1, (n-1, n) \rangle$.

We previously classified the center of the group algebra $\mathbb{C}[S_n]$ as conjugacy class independent functions. However, $\mathbb{C}[S_n]$ contains the centers of all $\mathbb{C}[S_i]$ for $i \leq n$. The commutative subalgebra of $\mathbb{C}[S_n]$ generated by the centers $Z(\mathbb{C}[S_i])$, $1 \leq i \leq n$ is called the *Gelfand-Tsetlin Subalgebra*, or GZ-Algebra for short. This algebra contains the following elements which are key to the approach of Vershik and Okounkov: Define, for $i = 1, ..., n$, the *Young-Jucys-Murphy* elements $X_i = \sum_{j=1}^{i-1} (j, i) \in \mathbb{C}[S_n]$ (where $X_1 = 0$). Since X_i is the difference between the sum of all transpositions in S_i and the sum of all transpositions in S_{i-1}, each X_i lies in the GZ-algebra, hence the X_i all commute. In particular, we have the following:

Proposition 12 ([CST10], Section 3.3.2) *Let X_i, X_{i+1}, and s_i be as above. Then the following hold:*

1. $X_i X_{i+1} = X_{i+1} X_i$.
2. $s_i X_j = X_j s_i$ *for* $j \neq i, i+1$.
3. $s_i X_i + 1 = X_{i+1} s_i$.

Denote by $H_i(2)$ the subalgebra generated by X_i, X_{i+1}, and s_i. This is an example of an affine Hecke algebra. Since $\mathbb{C}[S_n] = \langle \mathbb{C}[S_{n-1}], s_{n-1} \rangle$, we have that $\mathbb{C}[S_n] = \langle \mathbb{C}[S_{n-1}], H_{n-1}(2) \rangle$. This may seem like a more complicated presentation of the group algebra than $\langle \mathbb{C}[S_{n-1}], \mathbb{C}[S_2] \rangle$, but it is this presentation which Vershik and Okounkov use to make the inductive step from representations of S_n to those of S_{n+1}.

4.6 The Gelfand-Tsetlin Basis of an S_n-Module

We will now begin considering specific representations of S_n and how the elements of $\mathbb{C}[S_n]$ we discussed in the previous section act on a given representation. It is in this section that the simplicity of the branching graph comes into play.

Suppose that V is an irreducible representation of S_n. Restricting to the subgroup S_{n-1} allows us to canonically decompose V into irreducible S_{n-1}-modules. By induction, we can decompose V into one dimensional subspaces (specifically irreducible S_1-modules). Choosing an element from each one dimensional summand and normalizing with respect to the S_n-invariant inner product on V gives a basis of V called the *Young basis*, or *Gelfand-Tsetlin Basis*, or GZ-basis for short. The dimension of V is equal to the number of paths in the Bratteli diagram from V to the unique $\mathbb{C}[S_1]$ module. Given such a path T in the Bratteli diagram, denote by v_T the normalized element of the GZ-basis corresponding to that path, and denote by V_T the one dimensional span of v_T. Thus, we have a canonical decomposition of any S_n-module V as

$$V = \oplus_T V_T \tag{4.6.1}$$

where we index over all paths T in the Bratteli diagram ending at V. The connection between the GZ-basis and the GZ-algebra is found in the following proposition, which applies to any multiplicity free chain of finite groups:

Proposition 13 ([CST10], Theorem 2.2.2) *The GZ-subalgebra is the algebra of all operators diagonal in the GZ-basis. Specifically,* $GZ(n) = \{a \in \mathbb{C}[S_n] | \rho(a)v_T \in \mathbb{C}v_T \ \forall \rho \in \hat{G}, \forall T\}$ *Hence, the GZ-algebra is maximally commutative in* $\mathbb{C}[S_n]$.

Proof: Recall that projection of a G-module onto the copies of an irreducible submodule is given by multiplication by a central idempotent in $Z(\mathbb{C}[G])$. Hence, by simplicity of the Bratteli diagram, if V is a $\mathbb{C}[S_n]$-module then projection of V onto V_T is given by consecutive multiplication of idempotents lying in $Z(\mathbb{C}[S_i])$ for $i \leq n$. Since each of these by definition lies in the GZ-algebra, we conclude that the algebra of operators diagonal in the GZ-basis lies in the GZ-algebra. Since such an algebra is maximally commutative, and since the GZ-algebra is commutative, the proposition follows. ∎

One important implication of this, which we will use later, is that an element of the Young basis, say v_T, is determined, up to a scalar factor, by its eigenvalues with respect to GZ(n).

Proposition 14 ([CST10], Corollary 2.2.3) *Let (ρ, V_ρ) be an irreducible S_n-module. Suppose that the eigenvalues of $GZ(n)$ acting on $v \in V_\rho$ are the same as those of a Young element v_T. Then v is a multiple of v_T.*

Proof: Let $F_T \in GZ(n)$ be orthogonal projection of V_ρ onto $\mathbb{C}v_T$. Write $v = \sum \lambda_T v_T$. Then, by linearity we have $\rho(F_T)v = \lambda_T v_T$. However, since $\rho(F_S)v_T = \delta_{T,S}v_S$, we also have that $\rho(F_S)v = \delta_{T,S}v$. In particular we have that $v = \lambda_T v_T$ as required. ∎

We will soon show that the same result holds for $v \in V_\lambda$ if we only look at the eigenvalues of the Jucys Murphy elements. This will prove useful later, when we will show that a certain element in V_λ is a Young basis element by showing that each X_i acts diagonally on it.

4.7 A Generating Family for $GZ(n)$

The goal of this section is to show that the GZ-algebra is generated by the YJM-elements. Since consecutive YJM-elements together with an involution generate the Hecke algebra $H(2)$, we will be able to classify the spectrum of the YJM-elements on the GZ-basis by studying representations of $H(2)$, which will rely on basic linear algebra. This section and the proofs in it were largely inspired by the excellent summary article [Sri07], Section 4.

As already stated, the main result of this section is the following:

Theorem 12 ([CST10], Corollary 3.2.7) *The GZ-algebra $GZ(n)$ is generated by the YJM elements $X_1, X_2, ..., X_n$.*

The theorem will be proved by induction on n, the case of $n = 1, 2$ being trivial. The induction step, as we shall see, will follow from the following sequence of facts. We denote, for λ a Young diagram, $P_\lambda = \sum \sigma$ where $\sigma \in S_{n-1}$ has shape λ. We also denote $P_{\lambda,i} = \sum \sigma$ where $\sigma \in S_n$ has shape λ and contains n in a i-cycle. For example $P_{(i)}$ is the sum of all i-cycles in S_{n-1}, while $P_{(2),2} = X_n$.

Lemma 3 ([Sri07], Section 4) *The centralizer $Z(n-1,1)$ is generated by the elements $P_{(2)},...,P_{(n-1)} \in Z(n-1)$ and $P_{(2),2},..,P_{(n),n} \in Z(n-1,1)$.*

Proof:

Since $Z(n-1,1)$ has as a linear basis all $P_{\lambda,i}$ where λ is a Young diagram of n boxes and i is a part of λ, it suffices to show that each of these elements lies in the algebra $\langle P_{(2)},...,P_{(n-1)},P_{(2),2},..,P_{(n),n}\rangle$. This follows by induction on the number k of elements moved by a permutation with shape λ. If $k=0,1$ then we get the identity permutation, which is clearly in $\langle P_{(2)},...,P_{(n-1)},P_{(2),2},..,P_{(n),n}\rangle$. Assuming the result holds for for k parts moved, we let $\lambda=(\lambda_1,...,\lambda_r)$ have $k+1$ moved parts. Let $\lambda_1,..,\lambda_s$ be the nontrivial parts of λ. First we consider $P_{\lambda,1}$. Consider the product $P_{(\lambda_1)}...P_{(\lambda_s)}$. As an element of $Z(n-1,1)$ we can write this product as $\alpha P_{\lambda,1}+\sum \alpha_\beta P_{\beta,1}$, where we sum over β which move strictly less parts than λ, and where α is nonzero. By induction, $P_{\lambda,1}$ lies in $\langle P_{(2)},...,P_{(n-1)},P_{(2),2},..,P_{(n),n}\rangle$. For the general $P_{\lambda,i}$ and we assume that $i=\lambda_1$, we repeat the procedure with $P_{(\lambda_1),\lambda_1}P_{(\lambda_2)}...P_{(\lambda_s)}$. ∎

Using a similar proof, we also have the following:

Lemma 4 ([Sri07], Section 4) *The center $Z(n)$ of $\mathbb{C}[S_n]$ is generated by all P_λ, where λ is a cycle in S_n.*

Proposition 15 ([Sri07], Section 4) $Z(n-1,1)=\langle Z(n-1),X_n\rangle$

Proof: By the previous lemmas it suffices to show that each of $P_{(k),k}$ lies in $\langle Z(n-1),X_n\rangle$. If $k=2$, $P_{(k),k}=X_n$. Assuming that the result holds for $1,2,...,k$, we consider the product $P_{(k),k}X_n$. A summand in this product $(i_1,..,i_{k-1},n)(i,n)$ is either of the form $(i,i_1,...,i_{k-1},n)$ or $(i_1,...,i_j)(i_{j+1},...,n)$, depending whether or not $i=i_j$ for some j. As an element of $Z(n-1,1)$ we can write this as $P_{(k+1),k+1}+\sum \alpha_{\mu,i}P_{\mu,i}$ where we sum over μ with at most k moved elements. By induction, $P_{(k+1),k+1}$ lies in $\langle Z(n-1),X_n\rangle$ and the result follows. Indeed, a closer look at the proof of lemma 3 shows that $P_{\lambda,k}$ lies in $\langle P_{(2)},...,P_{(r)},P_{(2),2},..,P_{(r),r}\rangle$ where r is the number of elements moved by a permutation of shape λ.

∎

Returning to the proof of the main theorem of this section, we want to show that

$GZ(n) = \langle X_1, ..., X_n \rangle$. If $n = 1, 2$, this is immediately clear, so we can assume that $\langle X_1, \cdots X_n \rangle = \langle GZ(n-1), X_n \rangle$ by induction. The previous proposition shows that $Z(n) \subset Z(n-1, 1) = \langle Z(n-1), X_n \rangle \subset \langle GZ(n-1), X_n \rangle$, whence $\mathrm{GZ}(n) = \langle X_1, ..., X_n \rangle$ as required, since $X_n \in GZ(n)$.

4.8 Spectrum of the YJM Elements

Recall that the YJM elements X_i act diagonally on the GZ-basis of a given representation. If v_T lies in the GZ-basis of a given representation, denote by $\alpha(T) = (a_1, ..., a_n)$ the eigenvalues of the YJM elements on v_T (ie. $X_j v_T = a_j v_T \; \forall j$). We call this the *spectrum of the YJM elements on v_T*. Since the YJM elements generate the GZ algebra, v_T is characterized up to a scalar factor by $\alpha(T)$. Denote by $\mathrm{Spec}(n)$ the set of all $\alpha(T)$ where v_T is a GZ-basis element of an irreducible representation of S_n. In this section we will study the structure of such vectors in $\mathrm{Spec}(n)$ in order to understand when two elements of $\mathrm{Spec}(n)$ arise from basis elements lying in the same irreducible representation of S_n. Another interpretation is that elements of $\mathrm{Spec}(n)$ arise from paths in the Bratteli diagram, and we would like to understand the connection between elements of $\mathrm{Spec}(n)$ induced from paths with the same endpoint. Later, we will show how elements of $\mathrm{Spec}(n)$ can be interpreted in terms of the Young graph $\mathbb{Y}$, which, as we have already hinted at, is the same as the Bratteli diagram.

Proposition 16 ([VO05], Proposition 4.1) *Let $\alpha(T) = (a_1, ..., a_n)$ be the spectrum of the GZ-element v_T. Then we have the following:*

1. *$a_1 = 0$, and $\forall i, a_i \neq a_{i+1}$.*

2. *$a_{i+1} = a_i \pm 1$ if and only if $s_i v_T = \pm v_T$.*

3. *If $a_{i+1} \neq a_i \pm 1$ then $(a_1, ..., a_{i+1}, a_i, ..., a_n) \in \mathrm{Spec}(n)$. Further, this is the spectrum of the Young basis element $(s_i - \frac{1}{a_{i+1} - a_i}) v_T$, which lies in the same irreducible S_n-module as v_T.*

These properties of vectors in $\mathrm{Spec}(n)$ are fundamental to the theory developed in [VO05], as are the methods by which they are proven.

To begin with, we take an element $\alpha = (a_1, ..., a_n)$ in $\mathrm{Spec}(n)$ corresponding to the element v_T in the GZ-basis. By the definition of α, we have that $X_i v_T = a_i v_T$. Now consider the subalgebra $H_i(2)$ spanned by s_i, X_i, and X_{i+1}, subject to the relations in Section 3.4. The subspace V spanned by v_T and $s_i v_T$ is invariant under the action of $H_i(2)$. If these vectors are independent, then with respect to this basis, the generators of $H_i(2)$ have the following form:

$$X_i = \begin{pmatrix} a_i & -1 \\ 0 & a_{i+1} \end{pmatrix}$$

$$X_{i+1} = \begin{pmatrix} a_{i+1} & 1 \\ 0 & a_i \end{pmatrix}$$

$$s_i = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Since X_i acts diagonally on V^λ, it must be diagonalizable over all invariant subspaces, whence $a_i \neq a_{i+1}$. In fact, a matrix of the form

$$\begin{pmatrix} a & \pm 1 \\ 0 & b \end{pmatrix}$$

is diagonalizable if and only if $a \neq b$. Hence, if $a_i = a_{i+1}$ then neither X_i nor X_{i+1} could be diagonalizable.

Now if v_T and $s_i v_T$ were dependent, then s_i being an involution, we must have that $s_i v_T = \pm v_T$. Then the relations in Section 4.5, Proposition 12, imply that $a_{i+1} = a_i \pm 1$. Conversely, if $a_{i+1} = a_i \pm 1$ then the span of v_T and $s_i v_T$ has exactly one 1 dimensional invariant subspace, the span of $s_i v_T \pm v_T$. Since the Hecke algebra is semi-simple, this implies that v_T and $s_i v_T$ must be proportional, whence $s_i v_T = \pm v_T$.

Assuming that $a_{i+1} \neq a_i \pm 1$, then with respect to the basis $\{v_T, s_i v_T - \frac{1}{a_{i+1} - a_i} v_T\}$, we have the following matrix forms for X_i and X_{i+1}:

$$X_i = \begin{pmatrix} a_i & 0 \\ 0 & a_{i+1} \end{pmatrix}$$

$$X_{i+1} = \begin{pmatrix} a_{i+1} & 0 \\ 0 & a_i \end{pmatrix}$$

Hence $s_i v_T - \frac{1}{a_{i+1}-a_i} v_T$ is a common eigenvector for X_i and X_{i+1} with eigenvalues a_{i+1} and a_i respectively. Moreover, since for $j \neq i, i+1$ we have $X_i s_j = s_j X_i$, we have that $X_j(s_i v_T - \frac{1}{a_{i+1}-a_i} v_T) = a_j(s_i v_T - \frac{1}{a_{i+1}-a_i} v_T)$ for $j \neq i, i+1$. Hence, as a common eigenvector for each X_i, $s_i v_T - \frac{1}{a_{i+1}-a_i} v_T$ is an element of the Young basis corresponding to $(a_1, ..., a_{i+1}, a_i, ...a_n) \in \text{Spec}(n)$ as required.

4.9 Young Diagrams and Content Vectors

In the previous section we defined the set $\text{Spec}(n)$ in terms of the Bratteli diagram of the symmetric group. Specifically, elements of $\text{Spec}(n)$ correspond to paths in the Bratteli diagram. In this section we will define a set $\text{Cont}(n)$, which will correspond similarly to paths in the Young graph. This set is the analogue of $\text{Spec}(n)$, and as we will show in this section, the two sets are in fact the same, as are the equivalence relations on the sets which determines which object they correspond to (irreducible modules in the first case, and Young diagrams in the second).

Every path in $\mathbb{Y}_n$ ending at a given Young diagram gives a unique numbering of that diagram with the property that the numbering is increasing along every row and column. To each path in the Young graph $\mathbb{Y}_n$, we associate to it its *content vector* $(a_1, ..., a_n) \in \mathbb{Z}^n$ as follows: We first consider the following infinite array of numbers, called the *content grid*:

0	1	2	3	4	5
-1	0	1	2	3	4
-2	-1	0	1	2	3
-3	-2	-1	0	1	2
-4	-3	-2	-1	0	1
-5	-4	-3	-2	-1	0

Next, to define the contents of a diagram we place the diagram in the top left hand corner of the content grid as follows:

0	1	2	3	4	5
-1	0	1	2	3	4
-2	-1	0	1	2	3
-3	-2	-1	0	1	2
-4	-3	-2	-1	0	1
-5	-4	-3	-2	-1	0

The numbers inside of the diagram are called the *contents* of the diagram. A path in the Young graph ending at λ corresponds to both a standard filling of λ, and to a specific order in which the boxes of λ were added. Now given a standard filling of a given diagram, we order its contents into a vector called the *content vector* of the filling by keeping track of the order in which the contents are added, with respect to the path T corresponding to the standard filling in question. The set of all such content vectors for diagrams of size n is denoted Cont(n). Clearly, the first element of any content vector is 0, the second ± 1, and so on. There are several possible contents which define the same diagram. However, as we will see, the equivalence relation on the set of content vectors which doesn't distinguish between content vectors of the same diagram is the same as the relation on the spectrum of GZ-elements lying in the same irreducible component. From this we will be able to deduce that the Bratteli diagram of the symmetric groups is isomorphic to the Young graph $\mathbb{Y}$.

First, we will explain in more detail what a content vector is. We know from the definition that the first two elements of any content vector are $(0, 1)$ or $(0, -1)$. We have the following proposition:

Proposition 17 ([CST10], Theorem 3.1.10) *Let* $(a_1, ..., a_n)$ *be a content vector for a Young diagram. Then:*

1. $a_1 = 0$.

2. *$a_k = a_i \pm 1$ for some $1 \leq i < k$ and $\forall k$.*

3. *If $a_j = a_k = a$ for $j < k$ then both of $a+1$ and $a-1$ will occur between the two occurrences of a.*

Further, these properties uniquely define all possible content vectors of Young diagrams.

In [VO05] and [CST10] a content vector in $\mathrm{Cont}(n)$ is defined to be a vector in $\mathbb{C}^n$ satisfying the three properties above, with no initial reference to Young diagrams. It is then shown that taking the contents of a standard tableau provides a bijection between standard tableaux and content vectors in the sense above. The proof relies on the fact that if $(a_1, ..., a_n)$ is a content vector then so is (a_1), (a_1, a_2) and so on, which allows for induction on the size of the vector, or equivalently, the size of the Young diagram. The above conditions on content vectors provides an intuitive set of numerical axioms governing the growth of Young diagrams, which is an extremely large subject (see, for example, [Ker03]).

A natural question is that of determining when two content vectors define the same diagram. We have, for example, that $(0, 1, -1, 0)$ and $(0, -1, 1, 0)$ define the same diagram, and that the only difference between the two is that the second and third entries have been transposed. However, it is not true that any transposition can be applied to a content vector and still define a content vector. For example, while $(0, 1, -1)$ defines a Young diagram, $(1, 0, -1)$ does not. In [VO05] a permutation is called *admissible* with respect to a standard tableau if the application of the permutation to the entries of the tableau still gives a standard tableau.

Proposition 18 ([CST10], Section 3.1.4) *Let $\alpha = (a_1, ..., a_n), \beta = (b_1, ..., b_n)$ be content vectors.*

1. *The entries of β can be obtained from those of λ by consecutive admissible transposition of entries (that is, consecutive applications of permutations of the form $(j, j+1)$) if and only if they define the same diagram.*

2. *A Coxeter generator s_i is admissible for α if and only if $a_{i+1} \neq a_i \pm 1$.*

We have defined the sets Spec(n) in terms of the Bratteli diagram of S_n, and Cont(n) in terms of $\mathbb{Y}_n$, as well as equivalence relations on both sets, equating vectors which define the same object. In Spec(n), this means that two spectrums belong to basis elements lying in the same irreducible S_n-module. In Cont(n), this means that the two content vectors define the same diagram. Our goal is to show that $\text{Cont}(n) = \text{Spec}(n)$, and that the equivalence relations defined above are in fact the same. From this we will be able to deduce that the Bratteli graph of S_n is in fact isomorphic to $\mathbb{Y}$.

Proposition 19 ([VO05], Theorem 5.1) $\text{Spec}(n) \subset \text{Cont}(n)$.

Proof:

Since $X_1 = 0$, it is clear that if $(a_1, ..., a_n) \in \text{Spec}(n)$ then $a_1 = 0$. We will show, by induction on n, that the vector $(a_1, ..., a_n) \in \text{Spec}(n)$ satisfies (2) and (3) of Proposition 17, the case of $n = 2$ being trivial. If a_n is not equal to $a_i \pm 1$ for any $i < n$ then $(a_1, ..., a_{n-2}, a_n, a_{n-1}) \in \text{Spec}(n)$, and hence $(a_1, ..., a_{n-2}, a_n) \in \text{Spec}(n-1)$. But this contradicts (2) and the induction hypothesis. Hence (2) holds.

Now assume that $a_m = a_n = a$ and that both $a+1$ and $a-1$ do not occur amongst $a_{m+1}, ..., a_{n-1}$. Further, assume that m is the largest value with this property. Then $a+1$, for example, can occur at most once in $a_{m+1}, ..., a_{n-1}$, since if it occurred twice, a would occur between them, by the induction hypothesis. Thus, either we have $(a, ..., a)$ or $(a, ..., a+1, ..., a)$. In the first case we could apply admissible transpositions to have an element of Spec(n) with consecutive a's, a contradiction. In the second case we could again apply admissible transposition to have $(a, a+1, a)$ in the vector. This, however, can be shown to contradict the Coxeter relation $s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1}$.

∎

We need one simple lemma in order to prove the main result of this section. We formulate it set theoretically, although we will need it in relation to equivalence relations. This proposition is used intrinsically in [VO05] although it is neither stated nor proven, so we do both for completeness.

Proposition 20 *Let X and Y be nonempty sets with $X \subset Y$. Suppose that $X = \bigcup X_i$ and $Y = \bigcup Y_i$, where both unions are disjoint and have the same number of* *components. Suppose further that if $X_i \bigcap Y_j$ is nonempty then $Y_j \subset X_i$. Then $X = Y$ and up to reordering $X_i = Y_i$.*

Proof: Since X is a subset of Y, each component X_i intersects, and hence contains, at least one of the Y_j's. Since there are the same number of components, and a single Y_j cannot be contained in two of the X_i's, we have that Y is contained in X, so that $X = Y$.

∎

Lemma 5 ([VO05], Corollary 5.5) *If $\alpha \in \mathrm{Spec}(n) \subset \mathrm{Cont}(n)$, and if α, β are equivalent in $\mathrm{Cont}(n)$, then $\beta \in \mathrm{Spec}(n)$ and they are equivalent in $\mathrm{Spec}(n)$.*

Proof: Since α and β being equivalent in $\mathrm{Cont}(n)$ implies that the two vectors differ by a sequence of admissible Coxeter generators, we have by Section 4.8, Proposition 16(iii), that α and β are equivalent in $\mathrm{Spec}(n)$, as required.

∎

We finally proceed to the main result in [VO05]:

Theorem 13 ([VO05], Theorem 5.8) *The Bratteli graph of S_n and the Young graph $\mathbb{Y}$ are isomorphic. Specifically, $\mathrm{Spec}(n) = \mathrm{Cont}(n)$, and the equivalence relations on both are equal.*

Proof: $\mathrm{Cont}(n)$ has the same number of conjugacy classes as partitions of n, say $p(n)$, which is also equal to the number of irreducible S_n-modules, i.e. the number of conjugacy classes of $\mathrm{Spec}(n)$. However, if an element α in $\mathrm{Spec}(n)$ is equivalent to an element β of $\mathrm{Cont}(n)$, then necessarily β belongs to $\mathrm{Spec}(n)$ and is equivalent to α (in $\mathrm{Spec}(n)$). Hence a conjugacy class A of $\mathrm{Cont}(n)$ and B of $\mathrm{Spec}(n)$ are either disjoint, or $A \subset B$. Then the equivalence classes of $\mathrm{Spec}(n)$ and $\mathrm{Cont}(n)$ satisfy the above proposition, and hence the $\mathrm{Spec}(n) = \mathrm{Cont}(n)$ and the equivalence classes are the same.

∎

4.10 Action of the Coxeter Generators

While it may not be immediately apparent, the results in [VO05] actually contain complete information about how S_n acts on its irreducible modules. Specifically, we have explicit formulas for how the Coxeter generators of S_n, namely the transpositions $(1,2),(2,3),...,(n-1,n)$ act on a Young basis. Since any permutation can be decomposed into a product of Coxeter generators, it is possible (at least hypothetically) to write down the matrices of S_n acting on a given irreducible module.

Let v_T be a Young basis element for an irreducible S_n-module, and let $\alpha(T) = (a_1,...,a_n)$ be the spectrum of the GZ algebra on v_T. Let $s_i = (i,i+1)$ be a Coxeter generator of S_n. Recall that if s_i is inadmissible for T, that is if $a_{i+1} = a_i \pm 1$, then $s_i v_T = \pm v_T$, with the sign depending on whether i and $i+1$ lie in the same row or the same column of the element of $\mathrm{Stab}(n)$ corresponding to T. Note in this case that switching a_i and a_{i+1} does not yield a valid content vector.

In the case where s_i is admissible for T, we have that $s_i v_T$ is a linear combination of v_T and v_{s_iT} where s_iT denotes the Young basis element corresponding to the content vector with a_i and a_{i+1} switched. We know that v_{s_iT} is $(s_i - r_i^{-1})v_T$ where r_i is the difference between a_{i+1} and a_i. In the literature this value is known as the *axial distance* between i and $i+1$ in the standard tableau corresponding to T. To verify that the above vector is indeed a Young basis element with content $(a_1,..,a_{i+1},a_i,...,a_n)$, one simply needs to check the action of the YJM elements, which was done in a previous section.

Since Young basis elements are defined up to multiplication by a complex number, we now choose scalars so that the matrices of the Coxeter generators are of a particularly nice form. Supposing that v_T and v_{s_iT} are as above, we know that $s_i v_T = v_{s_iT} + r_i^{-1} v_T$, and hence $s_i v_{s_iT} = (1 - r_i^{-2})v_T - r_i^{-1} v_{s_iT}$. However, is is possible to chose a Young basis vector v_T for each T so that all Coxeter generators act as is Proposition 16, that is:

1. If i and $i+1$ lie in the same row of the tableau associated to T, then $s_i v_T = v_T$.

2. If they lie in the same column, then $s_i v_T = -v_T$

3. If they lie in different rows and columns then s_i acts on $\mathrm{span}\{v_T, v_{s_iT}\}$ as the matrix
$$\begin{pmatrix} r_i^{-1} & 1 - r_i^{-2} \\ 1 & -r_i^{-1} \end{pmatrix}$$

If we normalize the Young vectors with respect to the canonical inner product on V_λ making all permutations unitary, and denote by $w_T = \dfrac{v_T}{\|v_T\|_{S_n}}$, we get the so called *Young's orthogonal form* of V_λ, we have that s_i acts on the span of w_T and w_{s_iT} as

$$\begin{pmatrix} r_i^{-1} & (1 - r_i^{-2})^{1/2} \\ (1 - r_i^{-2})^{1/2} & -r_i^{-1} \end{pmatrix}$$

Hence the action of the Coxeter generator s_i on w_T can be summarized as follows:

1. If i and $i+1$ lie in the same row of the tableau associated to T, then $s_i w_T = w_T$.

2. If they lie in the same column, then $s_i w_T = -w_T$.

3. If they lie in different rows and columns then s_i acts as the above orthogonal matrix on $\mathrm{span}\{w_T, w_{s_iT}\}$.

For $T, S \in \mathrm{Stab}(\lambda)$ define $d(T, S)$ as the minimal number of admissible Coxeter generators which send T to S. This is a well defined metric on the space of paths $\mathrm{Stab}(\lambda)$. From the discussion above, and by a simple induction argument, we get the following.

Corollary 4 *Let $\sigma \in S_n$ and let w_T be a Young vector for the path T. Then $\sigma w_T \in \mathrm{span}\{w_S | d(T, S) \leq k\}$, where k is the minimal number of Coxeter generators in a decomposition of σ.*

It is important to note that the action of S_n on a given irreducible is completely encoded in the structure of the Young graph. Most constructions of the irreducible S_n-modules, specifically the Specht modules in the next chapter, use the combinatorial structure of Young diagrams as an intrinsic part of their construction.

4.11 Matrix Decomposition of $\mathbb{C}[S_n]$

One of the important implications of results in [VO05] concerns the problem of decomposing the group algebra $\mathbb{C}[S_n]$ into block matrices. This can be done, theoretically, for the group algebra of any finite group G, assuming that we take for granted the decomposition $\mathbb{C}[G] \cong \oplus_\lambda V_\lambda^{\oplus d_\lambda}$. This decomposition, however, relies essentially on abstract representation and character theory, and does not give any concrete methods on how go between sums of group elements and block matrices. This problem, however, is solved for the symmetric group algebra, because of the results in [VO05].

We saw a simple example of this in Section 2.2.6, for $\mathbb{C}[S_3]$. We were able to reduce the problem of moving between group algebra elements and block matrices to solving a six by six system of linear equations. In order to do this, we used the fact that we knew what the matrices of all the permutations were acting on the three irreducible representations U, U', and V. We could then write out the matrices of S_3 acting on the space $U \oplus V \oplus U'$, which constitute a basis of the group algebra, and then reduce the problem stated above to solving a simple system of linear equations.

In general, we now know the matrices of each of the Coxeter generators $s_i = (i, i{+}1)$ acting on each irreducible S_n module V_λ. Since these elements generate the full symmetric group, one could, in theory, compute all the matrices for all permutations acting on the space $\oplus_\lambda V_\lambda$. These matrices would result in a system of equations represented by a matrix of size $(\sum_\lambda d_\lambda)^2$. Then given a block matrix, we could recover which group algebra element the matrix represents simply by solving a matrix equation.

In the example of $\mathbb{C}[S_3]$, the six by six matrix we used to determine which matrix corresponded to which element of the group algebra was

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 0 & 1 & -1 & 0 & -1 & 1 \\ 0 & 0 & -1 & 1 & 1 & -1 \\ 1 & 1 & 0 & -1 & -1 & 0 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{pmatrix}$$

However, if instead of using the two by two matrices which we chose to represent S_3 acting on V, we use the two by two orthogonal matrices from the paper of Vershik and Okounkov, that is the matrix

$$s_i = \begin{pmatrix} r_i^{-1} & (1-r_i^{-2})^{1/2} \\ (1-r_i^{-2})^{1/2} & -r_i^{-1} \end{pmatrix}$$

where $s_i = (i, i+1)$ and $r_i = a_{i+1} - a_i$, then the matrix we arrive at is

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1/2 & -1/2 & -1/2 & -1/2 \\ 0 & 0 & -\sqrt{3}/2 & \sqrt{3}/2 & \sqrt{3}/2 & -\sqrt{3}/2 \\ 0 & 0 & -\sqrt{3}/2 & \sqrt{3}/2 & -\sqrt{3}/2 & \sqrt{3}/2 \\ 1 & -1 & 1/2 & 1/2 & -1/2 & -1/2 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{pmatrix}$$

From this matrix we can freely move between matrix elements and group algebra elements.

As in our first decomposition of $\mathbb{C}[S_3]$, we can look at the maximally abelian subalgebra spanned by the matrix elements $e_{11}, e_{22}, e_{33}, e_{44}$. These correspond, respectively, to the group algebra elements $1/6(1+(12)+(13)+(23)+(123)+(132)), 1/6(2+2(12)-(13)-(23)-(123)-(132)), 1/6(2-2(12)+(13)+(23)-(123)-(132)), 1/6(1-(12)-(13)-(23)+(123)+(132))$. However, because we have used the matrices of Young's orthogonal form, we know that this MASA is in fact the GZ subgalgebra which is generated by the YJM elements X_2, X_3. As block matrices, these correspond to

$$X_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and

$$X_3 = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

as prescribed by the theory of Vershik and Okounkov. We also know that these two matrices generate the GZ algebra, and hence the matrix elements e_{ii} can be written as polynomials in X_2 and X_3. If we denote the polynomial as $e_{ii} = p_i(X_2, X_3)$ then we have that

1. $p_1(x, y) = 1/6(1 + x + y + xy)$.
2. $p_3(x, y) = 1/6(2 + 2x - y - xy)$.
3. $p_4(x, y) = 1/6(2 - 2x + y - xy)$.
4. $p_4(x, y) = 1/6(1 - x - y + xy)$.

Because $X_2^2 = 1$ we also have the following polynomials representing the orthogonal projections:

1. $p_1(x, y) = 1/6(1 + x)(1 + y)$.
2. $p_3(x, y) = 1/6(x + 1)(2x - y)$.
3. $p_4(x, y) = 1/6(x - 1)(2x - y)$.
4. $p_4(x, y) = 1/6(1 - x)(1 - y)$.

While not equal as polynomials, both sets define the same group algebra elements when evaluated on the YJM elements. The first and fourth polynomial indeed corresponds to the projection formulas onto irreducible components, and $p_3 + p_4 = 1/3(2 - xy)$, which is projection $\mathbb{C}[S_3] \to V^{\oplus 2}$ as required. Both p_1 and p_4 can be extended to general formulas for any n. Specifically, we have that if $\lambda = (n)$ then the associated polynomial for p_λ is $1/6!(1 + x_2) \cdots (1 + x_n)$, while the polynomial associated to the partition $\lambda = (1^n)$ is $p_\lambda = 1/6!(1 - x_2) \cdots (1 - x_n)$. We will later discuss an inductive formula for the polynomials in the YJM elements corresponding to any path in the Young diagram.

Knowing how to view $\mathbb{C}[S_n]$ as matrix algebras allows us to ask other interesting questions. For example, according to our matrix model above, the canonical inclusion $\mathbb{C}[S_2] \subset \mathbb{C}[S_3]$ is given by all matrices of the form

$$\begin{pmatrix} a+b & 0 & 0 & 0 \\ 0 & a+b & 0 & 0 \\ 0 & 0 & a-b & 0 \\ 0 & 0 & 0 & a-b \end{pmatrix}$$

where a and b are complex numbers. This forms a two dimensional abelian subalgebra of $\mathbb{C}[S_3]$. However, $\mathbb{C}[S_3]$ contains two other copies of $\mathbb{C}[S_2]$, namely the algebras spanned by $\{1, (1,3)\}$ and by $\{1, (2,3)\}$. The first consists of all matrices of the form

$$\begin{pmatrix} a+b & 0 & 0 & 0 \\ 0 & a-b/2 & -\sqrt{3}/2b & 0 \\ 0 & -\sqrt{3}/2b & a+b/2 & 0 \\ 0 & 0 & 0 & a-b \end{pmatrix}$$

while the second consists of all matrices of the form

$$\begin{pmatrix} a+b & 0 & 0 & 0 \\ 0 & a-b/2 & \sqrt{3}/2b & 0 \\ 0 & \sqrt{3}/2b & a+b/2 & 0 \\ 0 & 0 & 0 & a-b \end{pmatrix}$$

If we take an arbitrary matrix of the second form and diagonalize it, we see that the change of basis matrix is independent of the choice of a and b. The same is true of matrices of the third sort. In summary, each of the three isomorphic copies of $\mathbb{C}[S_2] \subset \mathbb{C}[S_3]$ are conjugate. That being said, we can ask what is the geometric relationship between the three bases which diagonalize each of these subalgebras. We will focus on the center two by two matrices, since the outer corners are the same in each of the algebras. Assuming that the two dimensional algebra consisting of the matrices of the form

$$\begin{pmatrix} a+b & 0 \\ 0 & a-b \end{pmatrix}$$

is diagonal in the basis $\{x, y\}$ of $\mathbb{C}^2$, then the algebra of matrices

$$\begin{pmatrix} a-b/2 & -\sqrt{3}/2b \\ -\sqrt{3}/2b & a+b/2 \end{pmatrix}$$

is diagonal in the basis $\{-1/\sqrt{3}x+y, \sqrt{(}3)x+y\}$. Indeed, a matrix of the first form has eigenvalues $a \pm b$ and a change of basis matrix given by

$$\begin{pmatrix} -1/\sqrt{3} & \sqrt{3} \\ 1 & 1 \end{pmatrix}$$

Similarly, a matrix of the form

$$\begin{pmatrix} a-b/2 & \sqrt{3}/2b \\ \sqrt{3}/2b & a+b/2 \end{pmatrix}$$

is diagonalized in the basis $\{\sqrt{3}x+3y, -3x+\sqrt{3}y\}$, with change of basis matrix

$$\begin{pmatrix} \sqrt{3} & -3 \\ 3 & \sqrt{3} \end{pmatrix}$$

The angle of rotation between consecutive bases is $\arctan(1/\sqrt{3})$, or 30 degrees. However, the sub-algebra generated by 1 and $(1,3)$ is superfluous since the other two generate $\mathbb{C}[S_3]$. This follows from the fact that they contain all the Coxeter generators of the symmetric groups.

4.12 Central Idempotents in $GZ(n)$

There is an inductive formula for the polynomials $p_\lambda^T(x_2, ..., x_n)$ which satisfy the relation $p_\lambda^T(X_2, ..., X_n) = E_\lambda^T$ where E_λ^T is projection onto the Young basis element v_λ^T, and $X_2, ..., X_n$ are the YJM elements. Such polynomials exist since $GZ(n)$ is generated by the YJM elements.

Before we begin, if T is a path in the Young graph with end point λ, a partition of n, or equivalently T is a standard filled diagram for λ, let T' denote the standard filling of the diagram λ' of $n-1$ boxes obtained by removing the box in T containing n.

Proposition 21 ([CST10], Theorem 3.4.11) *Let T, λ, and p_λ^T be as above. Then we have the following:*

$$p_\lambda^T = p_{\lambda'}^{T'} \left(\prod_{S \neq T, T' \to S} \frac{(a_n(S) - X_n)}{(a_n(S) - a_n(T))} \right)$$

The proof of this fact relies on two things. First, since $X_i v_T = a_T(i) v_T$ for each i and any T, we have that $p(X_2, ..., X_n) v_T = p(a_T(2), ..., a_T(n)) v_T$ for any T and any polynomial $p(x_2, .., x_n)$ in $n-1$ variables. Secondly, the central projections are uniquely determined by the relations

1. $E_\lambda^T v_\lambda^T = v_\lambda^T$.

2. $E_\lambda^T v_\beta^S = 0$ if $\lambda \neq \beta$ or $T \neq S$.

One then checks that the right hand side of the equation in the above proposition satisfies (2) and scales v_λ^T. Appropriately normalizing gives the desired polynomials.

Chapter 5

The Irreducible Specht Modules S_λ

5.1 Introduction

The goal of this section is to give another construction, found in [Sag01], Chapter 2.3, of the irreducible S_n-modules, known as *Specht modules*, and to give a brief outline of another construction, found in [FH04], Chapter 4.2, of the irreducibles as cyclic submodules of $\mathbb{C}[S_n]$.

Having classified conjugacy classes as Young diagrams, we will associate to each diagram an irreducible S_n-module in a way that respects the branching rules of the Young graph. The general method for constructing the irreducible S_n-modules is to associate to each Young diagram of n boxes, and hence for each conjugacy class of S_n, an idempotent in the group algebra $\mathbb{C}[S_n]$ such that different diagrams give orthogonal idempotents. The left ideals in $\mathbb{C}[S_n]$ generated by each idempotent are pairwise non-isomorphic irreducible modules of S_n with respect to left multiplication by S_n. Thus, by exhaustion, we will have constructed all of the irreducible S_n-modules. We will do this first concretely, in the form of the well known Specht Modules (see [Sag01], Section 2.3), and then abstractly in terms of cyclic submodules of $\mathbb{C}[S_n]$ (see [FH04], Section 4.2).

5.2 Specht Modules

Given a cycle structure $\lambda = (\lambda_1, \lambda_2, ..., \lambda_r)$ which is understood to represent a conjugacy class in S_n, we will construct an irreducible module, denoted S_λ, and called the *Specht module corresponding to* λ. S_λ will, as discussed above, be a submodule of a reducible S_n-module, here denoted by M_λ.

To begin with, recall that $\mathrm{Tab}(\lambda)$ is the set of all filled Young diagrams of a fixed shape λ, that is, diagrams of shape λ in which we bijectively place the numbers $\{1, 2, ..., n\}$. Denote by t_λ the so called *canonical standard filling* of the diagram λ, where the upper left hand box is filled with a 1, the next rightmost with a 2, and so on until the bottom most righthand box is filled with an n. The filling of $\lambda = (4, 3, 2, 1)$ given by

$$\begin{array}{|c|c|c|c|}\hline 1 & 2 & 3 & 4 \\ \hline 5 & 6 & 7 \\ \cline{1-3} 8 & 9 \\ \cline{1-2} 10 \\ \cline{1-1}\end{array}$$

is an example of a standard filling. In general an element t of $\mathrm{Tab}(\lambda)$ is called *standard* if the values of t are increasing along each row and column. The subset of $\mathrm{Tab}(\lambda)$ consisting of standard fillings will be denoted $\mathrm{Stab}(\lambda)$.

When we are discussing t_λ for a fixed diagram λ, we will simply write t, assuming there will be no confusion. Next define, for $f \in S_n$, t_f to be the filling obtained by replacing each value i in t by the value $f(i)$. For example, with the standard filling of 10 boxes above, t_f would be

$$\begin{array}{|c|c|c|c|}\hline f1 & f2 & f3 & f4 \\ \hline f5 & f6 & f7 \\ \cline{1-3} f8 & f9 \\ \cline{1-2} f10 \\ \cline{1-1}\end{array}$$

With this notation in mind, we have a transitive action of S_n on $\mathrm{Tab}(\lambda)$, given by $\sigma t_\beta = t_{\sigma\beta}$. This action extends to a representation of S_n called the *permutation representation*, or the *permutation module*. This module is the complex vector space with $\mathrm{Tab}(\lambda)$ as a basis, and the action of S_n on this module is defined first over the basis (in terms of the transitive action defined above), and then extended by linearity.

This module is not irreducible for $n \geq 2$. Indeed, the one dimensional span of the element $\sum_{\sigma \in S_n} t_\sigma$ is isomorphic to the trivial representation, while the span of $\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) t_\sigma$ is isomorphic to the alternating representation. However, it is from this permutation module that we will construct all irreducible S_n-modules.

Let $\sim$ be the equivalence relation on $\mathrm{Tab}(\lambda)$ where two fillings are equivalent exactly when their respective rows consist of the same elements. M_λ is defined to be the free complex vector space over the set $\mathrm{Tab}(\lambda)/\sim$. We will denote the equivalence class of t_σ by T_σ. These equivalence classes are called *tabloids*.

We define an action of S_n on the vector space M_λ by letting, for $\beta, \sigma \in S_n$, βT_σ to be $T_{\beta\sigma}$. To verify that this is a well defined action, we suppose that $t_{\sigma_1} \sim t_{\sigma_2}$. Let $\{\alpha_1, ..., \alpha_k\}$ be the elements of a given row of t_{σ_1}, and hence also of the same row of t_{σ_2}. Then that row of both $t_{\beta\sigma_1}$ and $t_{\beta\sigma_2}$ consists of the elements $\{\beta\alpha_1, ..., \beta\alpha_k\}$. Hence $\beta t_{\sigma_1} = \beta t_{\sigma_2}$ as required.

Hence the vector space M_λ is a well defined S_n-module. In fact, M_λ is a cyclic $\mathbb{C}[S_n]$-module, in that any element of M_λ can be written as

$$(\sum_{\sigma \in S_n} \lambda_\sigma \sigma) T_{id} = \sum_{\sigma \in S_n} \lambda_\sigma T_\sigma$$

for complex numbers λ_σ. For simplicity sake, we will refer to t_{id} as simply t, and T_{id} as T. Then M_λ is the cyclic module $\mathbb{C}[S_n]T$.

If λ is a diagram, we define a diagram called λ^t, called the *transpose* of the diagram λ, as follows: λ^t is simply the diagram whose i^{th} row is the i^{th} column of λ. For example, according to our definition, the diagram $\lambda = (4, 3, 2, 1)$ is symmetric, in that it is equal to its transpose. More often, however, we will consider the transpose of a filled diagram. Thus for example, the transpose of the diagram considered at the beginning of page 63 would be

1	5	8	10
2	6	9	
3	7		
4			

With this definition, no filled diagram can be equal to its transpose, provided we are considering permutations of two or more elements.

5.3 Row and Column Stabilizers

Given a tableau t_σ we define the subgroup R_σ of S_n to be the subgroup of all elements which preserve each row of t_σ. We call R_σ the *Row Stabilizer* of t_σ. Analogously, we define C_σ, the *Column Stabilizer* of t_σ, to be all permutations preserving the columns of t_σ.

Note 3 *Because all elements of a given tabloid t have the same row elements, the subgroup $R_{\sigma T_\lambda}$ in independent of the particular representative, and hence we can refer to the row stabilizer of the tabloid itself, R_t. However, the column stabilizing subgroup does depend on the particular tableau we use.*

Next, for a tableau t_σ we define an element of the group algebra $\mathbb{C}[S_n]$, denoted b_σ, which is equal to

$$\sum_{\beta \in C_\sigma} \operatorname{sgn}(\beta)\beta.$$

Denote by e_σ the element $b_\sigma T_\sigma \in M_\lambda$ and define the *Specht Module* S_λ corresponding to the conjugacy class λ to be the submodule of M_λ spanned by all e_σ, where $\sigma \in S_n$. We of course need to verify that the span of the e_σ is invariant under the action of S_n.

For simplicity sake, we will denote R_{Id} by simply R, C_{Id} by C, b_{Id} by b, and e_{Id} by e. The reason for this will appear in the next proposition.

Since M_λ is spanned by all T_σ we need to understand the connection, for any σ, between R, C, e, and R_σ, C_σ, and e_σ respectively. The following lemma summarizes the necessary information:

Proposition 22 ([Sag01], Sections 2.3, 2.4) *Given $\sigma \in S_n$ the following hold:*

1. $R_\sigma = \sigma R \sigma^{-1}$ *and* $C_\sigma = \sigma C \sigma^{-1}$.

2. $b_\sigma = \sigma b \sigma^{-1}$.

3. $e_\sigma = \sigma e$.

4. *for* $\sigma \in C$, $\sigma b = b\sigma = sgn(\sigma) b$.

5. *b is self adjoint.*

6. *if a transposition (s,t) lies in C then $(1-(s,t))$ is a multiplicative factor of b.*

Proof:

1. Supposing that $\beta\sigma t = \sigma t$, we have $\sigma^{-1}\beta\sigma t = t$. Hence $\sigma^{-1}\beta\sigma$ lies in R_t, and so β lies in $\sigma R_t \sigma^{-1}$ as required. The opposite inclusion is similar.

2. We have that

$$b_\beta = \sum_{\sigma \in C_\beta} \operatorname{sgn}(\sigma)\sigma = \sum_{\sigma \in C} \operatorname{sgn}(\beta\sigma\beta^{-1})\beta\sigma\beta^{-1} = \beta(\sum_{\sigma\in C}\operatorname{sgn}(\sigma)\sigma)\beta^{-1} = \beta b \beta^{-1}.$$

3. Using (2), we have that $e_\sigma = b_\sigma T_\sigma = \sigma b \sigma^{-1}(\sigma T) = \sigma b T = \sigma e$ as required. Note that this proves the previous assertion that S_λ is indeed invariant under the action of S_n.

4. For $\beta \in C$ we have

$$\beta b = \beta(\sum_{\sigma\in C}\operatorname{sgn}(\sigma)\sigma) = \sum_{\sigma\in C}\operatorname{sgn}(\sigma)\beta\sigma = \sum_{\sigma\in C}\operatorname{sgn}(\beta^{-1}\sigma)\sigma = \operatorname{sgn}(\beta)b.$$

   The proof of the other assertion is almost identical.

5. First, because M_λ is an S_n-module, there is an inner product on M_λ such that each $\sigma \in S_n$ is a unitary operator. Letting $\langle , \rangle$ denote that inner product, we have that

$$\langle bu, v\rangle = \langle \sum_{\sigma\in C}\operatorname{sgn}(\sigma)\sigma u, v\rangle = \sum_{\sigma\in C}\operatorname{sgn}(\sigma)\langle \sigma u, v\rangle = \sum_{\sigma\in C}\operatorname{sgn}(\sigma^{-1})\langle u, \sigma^{-1}v\rangle = \langle u, bv\rangle$$

6. Choosing a transposition $(s,t) \in C$, we have that $\{\mathrm{Id}, (s,t)\}$ is a subgroup of C, so we can write $C = \bigcup_{i=1}^m \{\sigma_i, \sigma_i(s,t)\}$ for some $\sigma_i \in S_n$. Hence $b = k(\mathrm{Id} - (s,t))$ where $k = \sum_i \operatorname{sgn}(\sigma_i)\sigma_i \in \mathbb{C}S_n$.

∎

Note the following consequences of this lemma:

1. For any α, β, we have that $R_{\alpha\beta} = \alpha R_\beta \alpha^{-1}$, and similarly for $C_{\alpha\beta}$.

2. If a transposition (s, r) lies in C then any basis element $x \in M_\lambda$ containing s and r in the same row lies in the kernel of b. To see why this is, if $(s, r) \in C$ we can write $b = k(\mathrm{Id} - (s, r))$, and so $bx = k(\mathrm{Id} - (s, r))x$. Assuming that s and r lie in the same row of x, we have that $(s, r)x = x$ and thus $bx = 0$.

Since S_λ has as a generating set $\{e_\sigma | \sigma \in S_n\}$, S_λ is a cyclic $\mathbb{C}[S_n]$-submodule of M_λ generated by e, by (3) of the previous lemma.

Proposition 23 (Dominance Lemma, [Sag01], Lemma 2.2.4) *Let $t_\lambda \in \mathrm{Tab}(\lambda)$ and $t_\mu \in \mathrm{Tab}(\mu)$. If, for each i, the elements of row i of t_μ are in different columns of t_λ then λ dominates μ.*

Proof: By hypothesis we can reorder the columns of t_λ so that each of its rows contains the corresponding entries of the corresponding row of t_μ. The proposition follows. ∎

We are almost in a position to prove that each S_λ is irreducible. In order to do so, we note that if $t \in \mathrm{Tab}(\lambda)$, then b_t can be viewed as an operator on any M_μ, for μ a partition of n.

Proposition 24 ([Sag01], Section 2.4) *Let μ and λ be partitions of n, and consider two fillings t and s of λ and μ respectively:*

1. *If b_t is not trivial on M_μ then λ dominates μ.*

2. *If $\lambda = \mu$ and $b_t S \neq 0$, then $b_t S = \pm e_t$.*

3. *b_t maps M_λ onto the one dimensional subspace spanned by e_t.*

Proof: For (1), we must have that every two elements of a given row of s are in different columns of t or else $b_t[s] = 0$ by (6) of Proposition 30. Thus (1) follows from the Dominance Lemma. If $\lambda = \mu$ then $[s] = \sigma[t]$ for some $\sigma \in C_t$. Hence $b_t[s] = b_t\sigma[t] = \mathrm{sgn}(\sigma) b_t[t] = \pm e_t$ as required. Finally, writing $u = \sum_i a_i[s_i]$ where $a_i \in \mathbb{C}$ and s_i are λ tableaux, we have that $b_t u = a e_t$ where $a = \sum_i \pm a_i$ by (2).

∎

Theorem 14 ([Sag01], Theorem 2.4.4) *If U is a submodule of M_λ then either U is contained in $S_\lambda^\perp$ or S_λ is contained in U. Hence all S_λ are irreducible.*

Proof: Take $u \in U$. By (3) of the previous proposition, $b_t u = ae_t, a \in \mathbb{C}$. If for some t we have $a \neq 0$ then $ae_t = b_t u \in U$, hence $e_t \in U$ and so S_λ in contained in U. Conversely, suppose that $b_t u = 0$ for every t. Then we have, for any t, that $\langle u, e_t \rangle = \langle u, b_t[t] \rangle = \langle b_t u, [t] \rangle = 0$. Hence U lies entirely in $S_\lambda^\perp$.

∎

Theorem 14 implies that the S_λ are indeed irreducible. In order to show that isomorphic Specht modules cannot arise from different partitions, we need to know more about the set $\mathrm{Hom}_{S_n}(S_\lambda, M_\mu)$ where λ, μ are partitions of n. However, the existence of non-trivial maps in the above Hom-space implies that λ dominates μ. To show this, begin with a nonzero homomorphism $\psi : S_\lambda \to M_\mu$, and extend to a homomorphism $\psi : M_\lambda \to M_\mu$. Since this is nonzero, we must have that $\psi(e_t)$ is nonzero for some $t \in \mathrm{Tab}(\lambda)$. Then since ψ is a homomorphism and by the definition of e_t we must have that $b_t s$ is nonzero for some $s \in \mathrm{Tab}(\mu)$, whence λ dominates μ by Proposition 24.

Further, and by a similar argument, the only possible nonzero elements of $\mathrm{Hom}_{S_n}(S_\lambda, M_\lambda)$ are scalar multiplies of the identity. Putting this together, if S_λ and S_μ are isomorphic S_n-modules, then both partitions dominate one another, and hence are equal.

Since M_λ is an S_n module, we can write $M_\lambda = \oplus_\mu S_\mu^{\oplus a_\mu(\lambda)}$ where μ are partitions of n, and $a_\mu(\lambda)$ is the multiplicity of S_μ in M_λ, in other words, $a_\mu(\lambda) = \mathrm{Dim}(\mathrm{Hom}_{S_n}(S_\lambda, M_\mu))$. Since a multiplicity $a_\mu(\lambda)$ is nonzero implies that λ dominates μ, we can write $M_\lambda = \oplus_{\mu \trianglerighteq \lambda} S_\mu^{\oplus a_\mu(\lambda)}$, with $a_\lambda(\lambda) = 1$.

5.4 Further Structure of the Specht Modules

Recall the following from the structure of the branching graph of the symmetric groups:

1. The decomposition of irreducible S_{n-1}-modules into irreducible S_n-modules is given by the inclusion rules of Young diagrams.

2. Each irreducible S_n-module has dimension equal to the number of paths in the branching graph ending at that representation.

3. Hence, each irreducible corresponding to the diagram λ has dimension equal to the number of standard tableaux of shape λ, that is, the cardinality of $\mathrm{Stab}(\lambda)$.

Having constructed a correspondence $\lambda \mapsto S_\lambda$ between Young diagrams and irreducible S_n-modules, we would like to know how these facts are reflected in the structure of the Specht modules. The proofs can be found in [Sag01] (Theorem 2.8.3 and Theorem 2.6.5).

Proposition 25

1. *Let λ be a Young diagram, and S_λ the associated Specht module. Then $S_\lambda = \oplus S_\mu$ is a multiplicity free decomposition of the S_{n-1} modules S_λ, where we index over all diagrams μ of size $n-1$ which sit inside of λ.*

2. *S_λ has as a basis $\{e_t | t \in \mathrm{Stab}(\lambda)\}$.*

Since the inclusion and restriction rules of the Specht modules are also given by the Young graph, a natural question which we can ask is the following: Are the two constructions we have given, namely the association of λ to V_λ in the case of Vershik and Okounkov, and to S_λ in the case of Specht modules, equivalent? The answer is positive, and can be seen from the following theorem. Essentially, the irreducible representations of S_n are determined completely by the first two levels of the Bratteli diagram, and by the structure of the Young graph.

Theorem 15 ([CST10], Corollary 3.3.12) *Suppose that we have a family of S_n-modules U_λ, indexed by the vertices of the Young graph, such that the following hold:*

1. *$U_{(1)}$ and $U_{(2)}$ are the trivial representations, while $U_{(1,1)}$ is the alternating representation.*

2. *The branching rules of the U_λ are given by the Young graph $\mathbb{Y}$.*

Then each $U_\lambda \cong V_\lambda$, where the V_λ are the irreducible representations constructed in [VO05].

5.5 Young Symmetrizers

In this section we will discuss an alternative method of constructing the irreducible S_n-modules. For a detailed analysis, including all proofs, see [FH04], Section 4.2. All of the content of this section is taken from this reference.

Many of the objects of interest in the method of Specht Modules are still used here. We will use the same terminology when possible. Recall that given a filled tableau t, we have an element of the group algebra $\mathbb{C}[S_n]$, denoted b_t, given by

$$\sum_{\sigma \in C_t} \operatorname{sgn}(\sigma)\sigma$$

where C_t is the column stabilizer of t. Now we introduce a similar object, denoted by a_t, which is given by

$$\sum_{\sigma \in R_t} \sigma$$

where R_t is the Row Stabilizing subgroup of t. Finally, we let $c_t = a_t b_t$, which we call the *Young Symmetrizer of t*. When indexed over all tableaux, the set of cyclic submodules $(\mathbb{C}[S_n])c_t$ forms a complete set of pairwise non-isomorphic submodules of $\mathbb{C}[S_n]$, with the action of S_n given by left multiplication.

Note 4 *The actual filling of the diagram we choose is inconsequential, and hence we can assume that a given tableau is filled canonically, with a 1 in the first block, 2 in the second to the right, and so forth. Indeed, any filling of a tableau can be obtained from any other by permuting the elements of the tableau, and a simple calculation shows* *that $c_{\sigma t} = \sigma^{-1} c_t \sigma$. Another simple calculation shows that $\mathbb{C}[S_n]c_t$ is isomorphic, as an S_n-module, to $\mathbb{C}[S_n](\sigma^{-1}c_t\sigma)$.*

From now on we will assume that t is the canonically filled tableau for the diagram λ. As in the case with Specht modules, C_t and R_t denote the column and row stabilizing subgroups of S_n.

Proposition 26 ([FH04], Section 4.2) *Let $\sigma \in S_n$. Then:*

1. *$C_{\sigma t} = \sigma C_t \sigma^{-1}$ and $R_{\sigma t} = \sigma R_t \sigma^{-1}$.*
2. *$a_{\sigma t} = \sigma a_t \sigma^{-1}$, $b_{\sigma t} = \sigma b_t \sigma^{-1}$, and $c_{\sigma t} = \sigma c_t \sigma^{-1}$.*
3. *If $p \in R_t$ and $q \in C_t$, then $pa = ap = a, qb = bq = sgn(q)b$, and $pcq = \mathrm{sgn}(q)c$. Further, c is the unique (up to rescaling) element of $\mathbb{C}[S_n]$ with this property.*

Theorem 16 ([FH04], Section 4.2) *The cyclic submodules $(\mathbb{C}[S_n])c_\lambda$, indexed over all different Young diagrams, form a complete set of irreducible representations of the symmetric group S_n. In particular, each $(\mathbb{C}[S_n])c_\lambda$ is irreducible, and if $\lambda \neq \mu$ then $(\mathbb{C}[S_n])c_\lambda$ and $(\mathbb{C}[S_n])c_\mu$ are not isomorphic as S_n- modules.*

This method of constructing the irreducible modules has the following advantage: we know, from the general theory of representations of finite groups, that all irreducible representations of a finite group G appear as subrepresentations of the group algebra with the action of left multiplication, or equivalently, as a minimal left ideal in the group algebra, with multiplicities equal to the dimension of the representation. The summands in the decomposition into simple left ideals are generated by pairwise orthogonal idempotents in the group algebra, which sum to unity. This is, of course, the content of Weddurburn's Theorem for unital rings (see [DF04], Section 18.2, Theorem 4). Hence, from a purely algebraic point of view, this construction of the irreducible S_n-modules is quite natural. The disadvantage of this method is that inherent in the construction is the combinatorial structure of the Young diagrams, and not on the internal structure of the symmetric group, as in the paper of Vershik and Okounkov.

Chapter 6

Extension Problems and Tensor Products

6.1 Introduction and Motivations

A common problem in mathematics is that of extending a function to a larger set while preserving the important properties of function. Consider the following examples:

1. (Functional Analysis) Can a continuous functional on a normed vector subspace be extended to the entire space, while preserving both continuity and norm?

2. (Topology) Can a continuous function on a closed subspace of a topological space be extended to a continuous function on the entire space?

3. (Representation Theory) Can a representation of a subgroup of a finite or compact group be extended to one of the whole group?

4. (Linear Algebra) Can a real vector space be made into a complex vector space, that is, can we extend the field of scalars from a subfield to a larger field?

The answer to these questions are, given certain constraints, all positive. For example, the fact that (1) is true for arbitrary complex normed spaces is the content of the Hahn Banach Theorem. (2) Holds when the subspace is closed, the ambient space is normal (that is, the topology separates disjoint closed sets), and the function takes values in a compact interval. This is the well known Tietze-Urysohn Extension theorem. (3) and (4) are very closely related, and will be the central topic of this section.

6.2 Extension of Scalars

We will discuss (3) in a slightly more general setting, following [DF04]. In the first section we defined the notion of a left G-module, and showed how a left G-module structure on a vector space V induces a left $\mathbb{C}[G]$-module structure on V by linear extension. We will now make this notion more clear. Let R be a unital ring, and V an abelian group. We call V a *left R-module* if there is an operation $A : R \times V \to V$ such that the following holds:

1. $A(r + s, v) = A(r, v) + A(s, v)\ \forall r, s \in R, v \in V$.
2. $A(rs, v) = A(r, A(s, v))\ \forall r, s \in R, v \in V$.
3. $A(r, v + w) = A(r, v) + A(r, w)\ \forall v, w \in V, r \in R$.
4. $A(1_R, v) = v\ \forall v \in V$.

Right modules are defined analogously. There is a bijective correspondence between G-modules in the sense of section 1 and left $\mathbb{C}[G]$-modules in the above sense. Hence the question of extending R-modules to S-modules (where R is a subring of the ring S) can be applied to the problem of extending representations of H to those of G, where G is a finite group, and H a subgroup of G.

Suppose that N is a left R-module, and $R \subset S$ is an inclusion of unital rings. In order to extend the action of R on N to one of S on N, take the free abelian group on the set $S \times N$, modulo the following constraints:

1. $(s + t, n) = (s, n) + (t, n)$ for $s, t \in S$ and $n \in N$.
2. $(s, n + m) = (s, n) + (s, m)$ for $s \in S$ and $n, m \in N$.
3. $(sr, n) = (s, rn)$ for $s \in S, r \in R$, and $n \in N$.

The resulting quotient group is denoted $S \otimes_R N$, and is called the *tensor product of* S *and* N *over* R. The canonical projection is denoted by $(s, n) \mapsto s \otimes n$.

$S \otimes_R N$ becomes a left S-module by letting $s(t \otimes n) = (st) \otimes n$. The map of N into $S \otimes_R N$ defined by $n \mapsto 1 \otimes n$ is one of R-modules.

Proposition 27 ([DF04], Section 10.4, Theorem 8) *Suppose that* $f : N \to P$ *is an* R*-module homomorphism, where* P *is an* S*-module. Then there is a unique* S*-module homomorphism* $F : S \otimes_R N \to P$ *such that* $f(n) = F(1 \otimes n)$ *for every* $n \in N$*. Conversely,* $f(n) = F(1 \otimes n)$ *is an* R*-module homomorphism for every* S*-module homomorphism* $F : S \otimes_R N \to P$*. Hence there is a bijection* $Hom_R(N, P) \cong Hom_S(S \otimes_R N, P)$*.*

6.3 Induced Representations

We now come to the problem of extending a representation of a subgroup H to one of the ambient group G. Recall that there is a correspondence between representations of a finite group G and representations of the group algebra $\mathbb{C}[G]$. Hence we want to extend a representation $\mathbb{C}[H] \to \mathrm{End}(V)$ to a representation $\mathbb{C}[G] \to \mathrm{End}(V)$, where H is a subgroup of the finite group G. Viewing V as a $\mathbb{C}[G]$-module (in the sense of the previous section) and $\mathbb{C}[H]$ as a subring of $\mathbb{C}[G]$, we can use the method of extension of scalars and form the $\mathbb{C}[G]$ -module $\mathbb{C}[G] \otimes_{\mathbb{C}[H]} V$, where $a(s \otimes v) = (as) \otimes v$ for $a \in \mathbb{C}[G]$. Hence we have turned a $\mathbb{C}[H]$-module into a $\mathbb{C}[G]$-module as required. In finite group theory, the G-module $\mathbb{C}[G] \otimes_{\mathbb{C}[H]} V$ is called the *induced representation* of V to G, and is denoted by $\mathrm{Ind}(V)$, omitting any reference to G and H when possible. If it is necessary to specify that we are inducing V from H to G, we denote the induced representation by $\mathrm{Ind}_H^G(V)$. Induction has the following properties:

Proposition 28 ([FH04], Section 3.3) *Let* $\mathrm{Res}(V)$ *be the restriction of the representation* V *from* G *to the subgroup* H*:*

1. *If* V *is a* G*-module, then* $\mathrm{Ind}(\mathrm{Res}(V)) \cong V \otimes P$*, where* P *is the permutation representation of* G *on* G/H*.*

2. *(Transitivity of Induction) If $G \supset H \supset K$ is a chain of subgroups, then* $Ind_K^G(V) = \mathrm{Ind}_H^G(\mathrm{Ind}_K^H(V))$.

3. $\mathrm{Hom}_H(W, \mathrm{Res}\, V) = \mathrm{Hom}_G(\mathrm{Ind}\, W, U)$, *where V is a G-module and U is an H-module.*

4. *(Frobenius Reciprocity) Given V and W as above, we have that* $\langle \chi_{\mathrm{Ind}\, U}, \chi_V \rangle_G = \langle \chi_U, \chi_{\mathrm{Res}\, V} \rangle_H$.

(3) is called the universal property of $\mathrm{Ind}(V)$ and characterizes it up to isomorphism. (4) asserts that if U is an irreducible H-module, then in order to decompose $\mathrm{Ind}(U)$ into its irreducible components, it suffices to decompose $\mathrm{Res}(V)$ into irreducible H-modules for all irreducible G-modules V.

6.4 Alternating and Symmetric Tensors

Here we will discuss various constructions, similar to the tensor product defined above, and the universal properties which define them. Tensor products of modules allow us to turn bilinear maps into linear ones, and in that sense, the tensor product characterizes bilinear maps. In this section we will see how to turn multilinear maps, alternating multilinear maps, and symmetric multilinear maps into linear ones, and discuss the spaces which characterize them. We will assume that V and W are complex vector spaces in what follows.

A *k-multilinear map* $f : (V \times V \times \cdots \times V) \to W$ is one which is $\mathbb{C}$-linear in each variable (we take k-factors of V in the above definition). We call f *alternating* if $f(v_{\sigma(1)}, ..., v_{\sigma(k)}) = \mathrm{sgn}(\sigma) f(v_1, \cdots, v_k)$ $\forall \sigma \in S_k$ and $v_i \in V$, or equivalently if $f(v_1, \cdots, v_k) = 0$ whenever $v_i = v_j$ for some $i \neq j$. We call f *symmetric* if $f(v_{\sigma(1)}, ..., v_{\sigma(k)}) = f(v_1, \cdots, v_k)$ $\forall \sigma \in S_k$ and $v_i \in V$.

Proposition 29 ([DF04], Section 11.5, Theorems 34, 36) *1. There exists a unique (up to isomorphism) vector space, denoted $\wedge^k V$, together with an alternating map $(V \times V \times \cdots \times V) \to \wedge^k V$, with notation $(v_1, \cdots, v_k) \mapsto v_1 \wedge \cdots \wedge v_k$, that satisfies the following universal property: If $f : (V \times V \times \cdots \times V) \to W$*

is an alternating multilinear map, there is a unique linear map $F : \wedge^k V \to W$ *such that* $f(v_1, \cdots, v_k) = F(v_1 \wedge \cdots \wedge v_k)$.

2. *There exists a unique (up to isomorphism) vector space, denoted* $\operatorname{Sym}^k V$*, together with a symmetric map* $(V \times V \times \cdots \times V) \to \operatorname{Sym}^k V$*, with notation* $(v_1, \cdots, v_k) \mapsto v_1 \cdots v_k$*, that satisfies the following universal property: If* $f : (V \times V \times \cdots \times V) \to W$ *is a symmetric multilinear map, there is a unique linear map* $F : \operatorname{Sym}^k V \to W$ *such that* $f(v_1, \cdots, v_k) = F(v_1 \cdots v_k)$.

These spaces are called the k^{th} *alternating and symmetric powers of* V. There construction is similar in idea to that of the tensor product of two modules. For $\wedge^k V$, we begin by constructing the complex vector space $V^{\otimes k}$. We then quotient out by the subspace generated by all tensors of the form $(v_1 \otimes \cdots \otimes v_k - \operatorname{sgn}(\sigma) v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)})$. We denote the composition of projections $V \times \cdots \times V \to V^{\otimes k} \to \wedge^k V$ by $(v_1, \cdots, v_k) \mapsto v_1 \wedge \cdots \wedge v_k$. For $\operatorname{Sym}^k V$ we quotient by the subspace generated by all elements of the form $(v_1 \otimes \cdots \otimes v_k - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)})$ and denote the composition of projections $V \times ... \times V \to V^{\otimes k} \to \operatorname{Sym}^k V$ by $(v_1, \cdots, v_k) \mapsto v_1 \cdots v_k$.

As it turns out, the symmetric and alternating powers have useful characterizations in terms of the group algebra $\mathbb{C}[S_k]$. Recall the following facts from the representation theory of finite groups:

1. Every irreducible representation of G appears in $\mathbb{C}[G]$ exactly $\operatorname{Dim}(V)$ times.

2. $\mathbb{C}$ has (at least) two irreducible S_k-module structures: The trivial structure, denoted U, and the alternating structure, denoted U'.

3. $V^{\otimes k}$ is a right module over $\mathbb{C}[S_k]$, with the action given by $(v_1 \otimes \cdots \otimes v_k)\sigma = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}$.

Hence we could ask how U and U' appear in $V^{\otimes k}$ as S_k-modules. It turns out that the isotypic component of U in $V^{\otimes k}$ is isomorphic to $\operatorname{Sym}^k V$, and the component of U' in $V^{\otimes k}$ is isomorphic to $\wedge^k V$. By the projection formulas from Chapter 1, we have the following characterization of the alternating and symmetric powers (see [Bum04], Proposition 36.1):

Proposition 30 *We have the following:*

1. $v_1 \otimes \cdots \otimes v_k \mapsto \frac{1}{n!}\sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(k)}$ *is projection from* $V^{\otimes k}$ *to the isotypic component of* U'.

2. $v_1 \otimes \cdots \otimes v_k \mapsto \frac{1}{n!}\sum_{\sigma \in S_k} v_{\sigma(1)} \otimes \ldots \otimes v_{\sigma(k)}$ *is projection from* $V^{\otimes k}$ *to the isotypic component of* U.

Proposition 31 *We have the following isomorphisms of* $\mathbb{C}[S_k]$*-modules:*

1. $\operatorname{Sym}^k V \cong (V^{\otimes k}) \otimes_{\mathbb{C}[S_k]} U$.

2. $\wedge^k V \cong (V^{\otimes k}) \otimes_{\mathbb{C}[S_k]} U'$.

All of the constructions in this section, namely $V^{\otimes k}$, $\wedge^k V$ and $\operatorname{Sym}^k V$, are *functorial* in the following sense: A linear map $T : V \to W$ induces linear maps $T^{\otimes k} : V^{\otimes k} \to W^{\otimes k}, \wedge^k T : \wedge^k V \to \wedge^k W$, and $\operatorname{Sym}^k T : \operatorname{Sym}^k V \to \operatorname{Sym}^k W$, such that:

1. $\operatorname{Id}_V^{\otimes k} = \operatorname{Id}_{V^{\otimes k}}$, $\wedge^k \operatorname{Id}_V = \operatorname{Id}_{\wedge^k V}$, and $\operatorname{Sym}^k \operatorname{Id}_V = \operatorname{Id}_{\operatorname{Sym}^k V}$.

2. $(T \circ S)^{\otimes k} = (T^{\otimes k}) \circ (S^{\otimes k})$, $\wedge^k(T \circ S) = (\wedge^k T) \circ (\wedge^k S)$, and $\operatorname{Sym}^k(T \circ S) = (\operatorname{Sym}^k T) \circ (\operatorname{Sym}^k S)$.

These maps all constitute actions of $GL(V)$ on the vector spaces $V^{\otimes k}, \wedge^k V$, and $\operatorname{Sym}^k V$.

6.5 Problems in Entanglement

In this section we will discuss the following ongoing research problem. Given an irreducible S_n-module V_λ where λ is a partition of n, we can write $\operatorname{Ind}_{S_n}^{S_{n+1}}(V_\lambda)$ as $\oplus_\beta V_\beta$ where we index over all partitions β of $n+1$ containing λ in the sense of the Young graph. We can also construct the induced representation $\operatorname{Ind}_{S_n}^{S_{n+1}}(V_\lambda) = \mathbb{C}[S_{n+1}] \otimes_{\mathbb{C}[S_n]} V_\lambda$ using the method of extension of scalars. However, we can also construct a (vector space) isomorphism $\operatorname{Ind}_{S_n}^{S_{n+1}}(V_\lambda) \cong V_\lambda \otimes \mathbb{C}[S_{n+1}/S_n]$ by choosing a representative g_α for each coset α and mapping $v \otimes \alpha \mapsto g_\alpha \otimes v$. It is important to keep in mind that this is merely a map of vector spaces, and not one of S_n-modules.

Now any elements of $\mathrm{Ind}_{S_n}^{S_{n+1}}(V_\lambda)$ can be written as the sum of simple tensors in $V_\lambda \otimes \mathbb{C}[S_{n+1}/S_n]$, and we call the minimum number of simple tensors needed to decompose an element the *rank* of the elements, with respect to the decomposition $V_\lambda \otimes \mathbb{C}[S_{n+1}/S_n]$. Our goal is to study the rank of V_β in $V_\lambda \otimes \mathbb{C}[S_{n+1}/S_n]$, if $\lambda \to \beta$ in the sense of the Young graph. Thanks to the following formulation of a well known result, we have an upper bound on the possible ranks:

Theorem 17 *Let V and W be finite dimensional complex vector spaces, and let $x \in V \otimes W$. Then there are non-negative real numbers $\lambda_1, \dots, \lambda_k$, orthogonal vectors $u_1, \dots, u_k$ in V, and orthogonal vectors $w_1, \dots, w_k$ in W such that $x = \sum_{i=1}^{k} \lambda_i (u_i \otimes w_i)$.*

This theorem is essentially a result concerning complex matrices. If $\mathrm{Dim}(V) = m$ and $\mathrm{Dim}(W) = n$, then $V \otimes W$ is isomorphic to $M_{m,n}(\mathbb{C})$ via the map $v \otimes w \mapsto v^* w$. Hence we state the result for matrices:

Theorem 18 ([Nic01], Section 4.11, Theorem 1) *(Singular Value Decomposition) Let $M \in M_{m,n}(\mathbb{C})$. Then there exists unitary matrices U and V, and a diagonal matrix D, such that $M = UDV^*$. The diagonal entries of D are non-negative real numbers called the singular values of M, or the Schmidt coefficients of M. They are (up to reordering) uniquely determined by M.*

One example of this is easier to attack, namely the rank of the alternating representation of S_n after inducing the alternating representation of S_{n-1}.

If V is a finite dimensional complex vector space then there is a vector space inclusion $\wedge^{k-1}(V) \otimes V \supset \wedge^k(V)$. We first define the inclusion of $\wedge^2(V)$ in $V \otimes V$ given by $a \wedge b \mapsto a \otimes b - b \otimes a$. We then define the inclusion of $\wedge^3(V)$ in $\wedge^2(V) \otimes V$ by $a \wedge b \wedge c \mapsto (a \wedge b) \otimes c + (b \wedge c) \otimes a + (c \wedge a) \otimes b$. Moving now to the general inclusion, we map $v_1 \wedge \ldots \wedge v_k \mapsto \sum_{j=1,\cdots,k} (v_{j+1} \wedge \cdots \wedge v_{j-1}) \otimes v_j$, where all indices are taken modulo $(k+1)$.

We now introduce some terminology pertinent to quantum information theory:

Definition 2 *An element $x \in V \otimes W$ is called entangled if it cannot be represented as $v \otimes w$ for some $v \in V, w \in W$. We call a subspace X of $V \otimes W$ entangled if all of its (non zero) elements are entangled.*

It is a straight forward computation to show that $\wedge^2(V)$ is entangled in $V \otimes V$. One simply fixes a basis of V, and shows that if we can write $x \in \wedge^2(V)$ as $x = y \otimes z$ then necessarily $x = 0$. More generally, we have the following new result:

Proposition 32 *Let V be a real or complex vector space of dimension n. Then for all $1 \leq k \leq n$ we have that $\wedge^k(V)$ is entangled in $\wedge^{k-1}(V) \otimes V$.*

Proof:

As above we begin by fixing a basis $\{v_1, \cdots, v_n\}$ of V. If we have $x = y \otimes z$ where $y \in \wedge^{k-1}(V)$ and $z \in V$ then x can be written as

$$\sum_{1 \leq i_1 < \cdots < i_{k-1} \leq n, k=1,\cdots,n} a_{i_1 \ldots i_{k-1}} b_k \Big(\sum_{\sigma \in S_{k-1}} \mathrm{sgn}(\sigma) v_{i\sigma(1)} \otimes \cdots \otimes v_{i\sigma(k-1)} \otimes v_k \Big) \tag{6.5.1}$$

Now if this lies in $\wedge^k(V)$ it can also be written as

$$\sum_{1 \leq i_1 < \cdots < i_k \leq n} c_{i_1 \cdots i_k} \Big(\sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) v_{i\sigma(1)} \otimes \cdots \otimes v_{i\sigma(k)} \Big) \tag{6.5.2}$$

By equating coefficients we have two observations:

1. $a_{i_1 i_2 \ldots i_{k-1}} b_{i_1} = \ldots = a_{i_1 i_2 \ldots i_{k-1}} b_{i_{k-1}} = 0$ for all $1 \leq i_1 < \ldots < i_{k-1} \leq n$.
2. For $1 \leq i_1 < \cdots < i_k \leq n$ we have $a_{i_1 i_2 \ldots i_{k-1}} b_{i_k} = \pm a_{i_{\sigma(1)} \cdots i_{\sigma(k-1)}} b_{i_{\sigma(k)}}$ whenever $i_{\sigma(1)} < \cdots < i_{\sigma(k-1)}$.

We can now show that $a_{i_1 \ldots i_{k-1}} b_k = 0$ for all $1 \leq i_1 < \ldots < i_{k-1} \leq n$. Indeed, by our first observation, if $a_{i_1 \ldots i_{k-1}} \neq 0$ then $b_{i_1} = \ldots = b_{i_{k-1}} = 0$. The proposition then follows from the second observation. ∎

As an example, and to make the proof more transparent, we will show that $\wedge^3(V)$ is entangled in $\wedge^2(V) \otimes V$. If we fix an orthonormal basis $\{e_j | j \in J\}$ of V, then any element of $\wedge^3(V)$ can be written uniquely as $\sum_{i<j<k} \alpha_{i,j,k} e_i \wedge e_j \wedge e_k = \sum_{i<j<k} \alpha_{i,j,k} [e_i \wedge e_j \otimes e_k + e_j \wedge e_k \otimes e_i + e_k \wedge e_i \otimes e_j]$ for $\alpha_{i,j,k} \in \mathbb{C}$.

Assuming this element were entangled, we could also write it as $\sum_{i<j,k} a_{i,j} b_k e_i \wedge e_j \otimes e_k = \sum_{i<j<k} (a_{i,j} b_k e_i \wedge e_j \otimes e_k + a_{i,k} b_j e_i \wedge e_k \otimes e_j + a_{j,k} b_i e_j \wedge e_k \otimes e_i) + \sum_{i<j} [a_{i,j} b_i e_i \wedge e_j \otimes e_t + a_{i,j} b_j e_i \wedge e_j \otimes e_j]$

By equating coefficients we have

1. For $i < j$ we have $a_{i,j} b_i = a_{i,j} b_j = 0$.
2. For $i < j < k$ we have $a_{i,j} b_k = a_{j,k} b_i = -a_{i,k} b_j$ (since $e_i \wedge e_j = -e_j \wedge e_i$).

Hence if $a_{i,j}$ (for $i < j$) is not zero then by (1) we have that $b_i = b_j = 0$. But then for any $k \neq i, j$ we have $a_{i,j} b_k = a_{j,k} b_i = 0$. Hence the whole sum is zero, showing that the only entangled element is 0.

Having determined the entanglement of $\wedge^k(V)$ in $\wedge^{k-1}(V) \otimes V$, we ask the following question: Given an element $x \in \wedge^k(V)$, we can write it as $x = \sum_i \lambda_i (x_i \otimes y_i)$ where the x_i are orthogonal in $\wedge^{k-1}(V)$, the y_i are orthogonal in V, and the λ_i are the singular values. What, then, are the constraints on the values of the λ_i? The minimum number of Schmidt coefficients of an entangled tensor is an widely used indicator of entanglement. Hence our problem is that of determining the minimum possible rank of an element of $\wedge^k(V)$, viewed as an element of $\wedge^{k-1}(V) \otimes V$.

There is a similar but more general question we can ask: Suppose that G is a finite group, and V and W are two representations of G, and hence $V \otimes W$ is a representation of G. Supposing that we have a subrepresentation U of $V \otimes W$, we want to understand the possible singular values which can arise from elements of U. This will likely depend on V, W, and U, but the question is exactly what is the dependence? We will continue to study these questions in the future.

Chapter 7

Appendices

7.1 Appendix A: The Symmetric Group

This section is a summary of the necessary facts concerning the symmetric groups S_n. The main source for this section is [Sag01], Section 1.1, although the results and proofs can be found in almost any textbook on abstract algebra.

The symmetric group S_n is defined as all bijections of the set $\{1, ..., n\}$. This forms a non-abelian (for $n \geq 3$) group of order $n!$ with respect to composition of functions. There are many accepted ways of denoting an element $\sigma \in S_n$. In order to stress the fact that S_n consists of functions, there is the matrix notation

$$\begin{pmatrix} 1 & 2 & ... & n \\ \sigma(1) & \sigma(2) & ... & \sigma(n) \end{pmatrix}$$

so that $i \mapsto \sigma(i)$ for each i. A permutation σ is called a *k-cycle* if there are k elements $1 \leq i_0, ..., i_{k-1} \leq n$ such that $\sigma(i_j) = i_{(j+1) mod(k)}$, and σ fixes all other elements. Such a permutation is written is cycle notation $\sigma = (i_1, i_2, ..., i_{k-1})$. The following is an important fact:

Proposition 33 *Every permutation can be written uniquely as the product of cycles with disjoint support.*

Any cycle can be written as the product of transpositions (a permutation whose support consists of two elements). By the previous proposition, every permutation can be written as the product of transpositions. One can easily show that if (i, j) is a transposition with $i < j$ then (i, j) can be written as the product of transpositions of the form $(m, m+1)$. We call these elements the *Coxeter generators of* S_n, and S_n is an example of what is called a *Coxeter group*. Specifically, S_n is generated by the elements $s_i = (i, i+1)$ for $i = 1, .., n-1$ which satisfy the following relations:

1. $(s_i s_j)^2 = 1$ if $i = j$ and 2 else.

2. $s_i s_j = s_j s_i$ if and only if $|i - j| \geq 2$.

3. $s_i s_{i+1} = (i, i+1, i+2)$ and $s_{i+1} s_i = (i, i+2, i+1)$.

As a corollary, we have that S_n is generated by the subgroup S_{n-1} and the element s_{n-1}. This fact will be very important for studying the representation theory of S_n inductively.

We define the *sign* of a permutation σ, denoted by $sgn(\sigma)$, to be 1 if σ is the product of an even number of transpositions, and -1 if σ is the product of an odd number of transpositions. This is a well defined group homomorphism $S_n \to \{1, -1\}$.

By the *cycle structure* of a permutation σ we mean a vector of natural numbers $\lambda = (\lambda_1, ..., \lambda_r)$ such that $\lambda_i \geq \lambda_{i+1}$ for every i, and where σ can be written as the product of disjoint $\lambda_1, ..., \lambda_r$ cycles. By the above proposition, this λ is well defined. The following proposition is fundamental for our analysis of representations of the symmetric groups:

Proposition 34 *Two permutations are conjugate in S_n if and only if they have the same disjoint cycle structure.*

Hence conjugacy classes of the symmetric group S_n can be parameterized by vectors of integers $\lambda = (\lambda_1, ..., \lambda_r)$ such that

1. $\lambda_i \geq \lambda_{i+1}$ $\forall i$.

2. $\lambda_i \geq 1$ for each i.

3. $\sum_{i=1}^{r} \lambda_i = n$.

The numbers λ_j indicate the lengths of the cycles in any decomposition of a permutation into disjoint cycles. Such vectors are called *partitions of* n, and are fundamental to the representation theory of the symmetric groups.

Another important property of the symmetric groups is the existence of a chain of injective homomorphisms $S_1 \to S_2 \to S_3 \to ...$ given by inclusions. Namely, if $\sigma \in S_n$, we view σ an an element of S_{n+1} which fixes $n + 1$.

Proposition 35 *Every $\sigma \in S_n$ is conjugate to its inverse. Further, there exists $\alpha \in S_{n-1}$ so that $\sigma^{-1} = \alpha^{-1}\sigma\alpha$.*

7.2 Appendix B: Results from Functional Analysis

This section contains the necessary results concerning vector spaces and linear maps. An excellent reference, which the information in this section is taken from, is [Bol90].

7.2.1 Banach Spaces

A normed space is a vector space V over $\mathbb{R}$ or $\mathbb{C}$, together with a function $\|\| : V \to \mathbb{R}$ such that:

1. $\|v\| \geq 0 \; \forall v \in V$, and $\|v\| = 0$ if and only if $v = 0$.

2. $\|\lambda v\| = |\lambda| \|v\|$ for any scalar λ.

3. $\|v + w\| \leq \|v\| + \|w\| \; \forall v, w \in V$.

If V is complete with respect to the norm, we call V a Banach space. Every normed space we deal with in this thesis is a Banach space.

7.2.2 Continuous Linear Maps and the Hahn Banach Theorem

By a linear map between the normed spaces V and W we mean a map $T : V \to W$ satisfying $T(x + y) = T(x) + T(y)$ and $T(\lambda x) = \lambda T(x)$ for vectors x, y, and scalars λ. Denote the space of linear maps between V and W by $L(V, W)$. We occasionally need to consider maps $T : V \to W$ such that $T(x+y) = T(x) + T(y)$, and $T(\lambda x) = \overline{\lambda} T(x)$ for $x, y \in V$ and $\lambda \in \mathbb{C}$. Such a map is called *conjugate linear*.

Proposition 36 ([Bol90], Chapter 2, Theorem 2) *For a linear map $T : V \to W$ between normed spaces, the following are equivalent:*

1. *T is continuous.*
2. *T is continuous at zero.*
3. *T is continuous any point.*
4. *T is bounded.*

The infimum of all bounds on a map T constitutes a norm on the space of bounded linear maps between V and W, denoted $B(V, W)$. We can equivalently use the supremum of the map on the unit ball, and denote it by $\|T\|$, called the operator norm of T. If W is a Banach space, then with respect to this norm, $B(V, W)$ is a Banach space. In fact we have the following stronger result:

Proposition 37 ([Bol90], Chapter 2, Theorem 4) *$B(V, W)$ is a Banach space if and only if W is a Banach space.*

Continuous linear maps can be classified by their image in the following way:

Theorem 19 ([Bol90], Chapter 5, Theorem 8) *Let X and Y be Banach spaces. Then a linear map $T : X \to Y$ is continuous if and only if its graph is closed in $X \times Y$ with the product topology.*

If X and Y are normed spaces, a linear map $T : X \to Y$ is called invertible if it is a continuous bijection with a continuous inverse T^{-1}. For finite dimensional normed spaces, an invertible linear map is given by an invertible matrix, hence the inverse is also continuous. This hold, in fact, for arbitrary dimensions, assuming we work in Banach spaces:

Theorem 20 ([Bol90], Chapter 5, Theorem 7) *Let T be a continuous linear bijection of Banach spaces. Then T^{-1} is also continuous.*

In the special case when W is the base field, then $L(V, W)$ is denoted by V', and is called the algebraic dual space of V. We call $B(V, W)$ the analytic dual space of V, and denote it V^*. One of the most important result in functional analysis is the *Hahn Banach Theorem*, which asserts the following:

Theorem 21 ([Bol90], Chapter 3, Theorem 5) *1. Let V be a real vector space, Y a subspace, and suppose $f \in Y'$ is dominated by some convex function p, in the sense that $f(y) \leq p(y)\ \forall y \in Y$. Then f can be extended to all of V in such a way that the extension is still dominated by p.*

2. Let V be a complex vector space, Y a subspace, and suppose $p : X \to \mathbb{R}$ satisfies $p(\lambda x + \beta y) \leq |\lambda| p(x) + |\beta| p(y)\ \forall |\lambda| + |\beta| = 1$. Suppose further that $f \in Y'$, and $|f|$ is dominated by p. Then we can extend f to all of V in such a way that $|f|$ is still dominated by p.

In fact the corollaries of the Hahn Banach theorem are more important than the result itself:

Proposition 38 ([Bol90], Chapter 3) *(Corollaries of the Hahn Banach Theorem)*

1. *Let Y be a subspace of the normed space V. Then any continuous functional on Y can be extended to a continuous functional on V while preserving the norm.*

2. *For every $v \in V$, there is some $f \in V^*$ such that $f(v) = \|v\|$ and $\|f\| = 1$.*

3. *If $f(v) = f(w)$ for every continuous functional then $v = w$.*

4. *V^* separates points, in that for every pair of vectors $v \neq w$ there is a continuous functional f such that $f(v) \neq f(w)$.*

7.2.3 Duals, Double Duals, and Reflexive Banach Spaces

An isometric imbedding of a normed space V into a normed space W is linear map $T : V \to W$ which preserves the norm, in the sense that $\|T(v)\| = \|v\|\ \forall v \in V$. We often refer to an isometric imbedding as simply an imbedding. Note that any imbedding is automatically both injective and continuous.

Proposition 39 ([Bol90], Chapter 3, Theorem 10) *Let X be a normed space. Then X canonically imbeds into X^{**}.*

Proof: For any $x \in X$ denote by $\hat{x} : X^* \to \mathbb{C}$ the map $\hat{x}(f) = f(x)$. One checks that the map is a linear isometry, and that $\|\hat{x}\| = \|x\|$. ∎

A space for which this map is also surjective is called a *reflexive normed space*. We will use the notion of reflexivity when discussing vector valued integrals in relation to the representation theory of compact groups.

In this thesis we will work exclusively with Hilbert spaces (treated in the next section) which constitutes an important class of reflexive spaces.

7.2.4 Hilbert Spaces

In this section we will discuss a particularly nice example of a Banach space, called a Hilbert space. A Hilbert space, because of how its norm arrises, has much more geometric structure than an arbitrary Banach space, and it is this geometric structure which we will study.

Definition 3 *By a scalar product on a complex vector space V we mean a mapping $V \times V \to \mathbb{C}$, denoted $(v, w) \mapsto \langle v, w \rangle$, subject to the following constraints:*

1. *The scalar product is complex linear in the first variable.*
2. $\langle v, w \rangle = \overline{\langle w, v \rangle}$.
3. $\langle v, v \rangle \geq 0$, *with equality only when v is the zero vector.*

The map $v \mapsto (\langle v, v \rangle)^{1/2}$ defines a norm on V. When V is complete with this norm, we call V a Hilbert Space. In this paper we often take an inner product to be linear in the second variable, and conjugate linear in the first. However this does not affect the theory of Hilbert spaces.

Two vectors whose inner product is zero are called orthogonal. Similarly, two subspaces are called orthogonal if all pairs of vectors chosen from them are orthogonal. If F is a subspace, we denote by $F^{\perp}$ the subspace of elements orthogonal to all those in F. One of the fundamental results concerning Hilbert spaces is the following:

Theorem 22 ([Bol90], Chapter 9, Theorem 6) *Let F be a closed subspace of a Hilbert space H. Then $H = F \oplus F^{\perp}$.*

Given a closed subspace F of a Hilbert space H, we denote by P_F the operator $P_F : H \to H$ which maps a vector to its F-component. We call P_F the projection operator onto the subspace F. Projections have the following properties:

Proposition 40 ([Bol90], Chapter 9, Corollary 7) *(Properties of Projection Operators)*

1. $P_F \in B(H)$
2. $\|P_F\| = \|Id_H - P_F\| = 1$ *when F is not the zero subspace.*
3. $P_{F^{\perp}} = Id_H - P_F$.
4. $P_F^2 = P_F$.
5. $\langle P_F x, y\rangle = \langle P_F x, P_F y\rangle = \langle x, P_F y\rangle$

(4) says that projections are idempotents, and (5) says that projections are self adjoint. In fact, it is an easy exercise to show that these two properties are equivalent to being a projection operator.

Given $v \in H$, the map $\psi_v(w) = \langle w, v\rangle$ is a bounded linear functional on H. The following theorem asserts that all bounded functionals on a Hilbert space are of this form for a unique v.

Theorem 23 ([Bol90], Chapter 9, Theorem 9) *(Riesz Representation Theorem) The map $\psi : H \to H^*$ is an conjugate linear isometric isomorphism of H and H^**

7.2.5 Adjoint Operators

Given $T \in B(H)$ where H is a Hilbert space, and given $v \in H$, the map $w \mapsto \langle Tw, v\rangle$ is a continuous linear form on H. Hence by the Riesz Representation Theorem, there is a unique element of H, which we will denote T^*v, such that $\langle Tw, v\rangle = \langle w, T^*v\rangle$. The map T^* constitutes a bounded linear map on H, called the *adjoint of* T. The map $T \mapsto T^*$ is an isometric conjugate linear isomorphism of $B(H)$.

Definition 4 *Let $T \in B(H)$. We call T:*

1. *Self adjoint if $T = T^*$.*

2. *Unitary if T is invertible and $T^{-1} = T^*$.*

3. *Normal if $TT^* = T^*T$.*

7.2.6 Banach Space Valued Integrals

In order to extend the representation theory of finite groups to the representation theory of compact groups, we will need the notion of the integral, with respect to the normalized Haar measure, of a continuous function $x : G \to V$ where V is a reflexive Banach space, and G is a compact topological group. We assume reflexivity for two reasons. First, all Banach spaces considered in this thesis are reflexive. Secondly, assuming reflexivity makes the definition of the integral rather straightforward. The integral will be an element of V, denoted $\int_G x(g)d\mu(g)$ and will satisfy the necessary properties of the integral, namely:

1. $\int_G (x(g) + y(g))d\mu(g) = \int_G x(g)dg + \int_G y(g)d\mu(g)$.

2. $\int_G \lambda x(g)d\mu(g) = \lambda \int_G x(g)d\mu(g)$.

3. $|\int_G x(g)d\mu(g)| \leq \int_G |x(g)|d\mu(g)$.

4. If $v \in V$ then $\int_G v d\mu(g) = v$.

5. If $T \in B(V)$ then $T(\int_G x(g)d\mu(g)) = \int_G T(x(g))d\mu(g)$.

In order to define the integral, we note that if $x : G \to V$ is continuous and $f \in V^*$, then $\int_G f(x(g))d\mu(g)$ exists, and lies in $\mathbb{C}$. The map $f \mapsto \int_G f(x(g))d\mu(g)$ constitutes a continuous functional on V^*, and hence by reflexivity of V, there is a unique $v \in V$ such that $f(v) = \int_G f(x(g))d\mu(g)$ for every continuous functional f. We call this unique v the integral of x over G, and denote it $v = \int_G x(g)d\mu(g)$. Using this definition, the properties above are easy to prove, and they all rely on two facts: The Hahn Banach theorem, and the properties of the normal integral of functions $G \to \mathbb{C}$. For example, we prove (1) by showing that for any functional

$f \in V^*$, $f(\int_G (x(g)+y(g))d\mu(g)) = f(\int_G x(g)d\mu(g) + \int_G y(g)d\mu(g))$. But this follows by linearity of f and by the definition of the integral. For (3), we choose a functional f such that $\|f\| = 1$ and $\|f(\int_G x(g)d\mu(g))\| = |\int_G x(g)d\mu(g)|$, by the Hahn Banach Theorem. (3) then follows easily.

As an important example, suppose that V is also a Hilbert space. Then V is reflexive, so the above definition of the integral is applicable, and by the Riesz Representation theorem, $\int_G x(g)d\mu(g)$ is the unique element of V such that for every $w \in V$, we have that $\langle \int_G x(g)d\mu(g), w\rangle = \int_G \langle x(g), w\rangle d\mu(g)$. This property will prove very useful in the representation theory of compact groups.

7.2.7 Involutive Banach Algebras

In this section we collect the information about $*$ algebras which is necessary for understanding the Vershik and Okounkov approach to the representation theory of S_n. We begin with the following definition:

Definition 5

1. *A Banach Algebra is a Banach space B with an algebra structure such that $\|ab\| \leq \|a\|\|b\|$ for all $a, b \in B$. If B has a unit e, we require that $\|e\| = 1$.*

2. *A Banach $*$ algebra is a Banach algebra B together with an involution operator $x \mapsto x^*$ which is conjugate linear in x, satisfies $x^{**} = x$, $(xy)^* = y^*x^*$, and $e^* = e$ if B has a unit e.*

A Banach Algebra over the real numbers is defined analogously, with conjugate linear replaced by linear in the definition above. If, in addition, a Banach algebra A satisfies the relation $\|aa^*\| = \|a\|^2$ for all a, we call A a C^**-algebra*.

Since $*$ algebras have additional structure coming from the $*$ operation, it is natural to consider maps between $*$ algebras which preserve this structure. Such a map is called an $*$ *homomorphism*. Specifically, if $\psi : A \to B$ is a linear map between $*$ algebras such that $\psi(ab) = \psi(a)\psi(b)$, and $\psi(a^*) = \psi(a)^*$, we call ψ a $*$ *homomorphism between A and B*.

The simplest example of a Banach $*$ algebra is the complex numbers with complex conjugation. The next most important example is $B(H)$, where H is a Hilbert space, with the adjoint operation. We have similar definitions for a Banach $*$ algebra as for $B(H)$. Namely, we call $x \in B$ *normal* if $xx^* = x^*x$, *self adjoint* if $x = x^*$, and *unitary* if x is invertible with $x^* = x^{-1}$. We will be less concerned in this thesis with the analytic structure of Banach algebras, that is the norm, and we will often simply assume we are working with a $*$ algebra, that is an algebra with a conjugate linear involution.

A natural question is that of embedding a real $*$ algebra inside of a complex $*$ algebra. Specifically, if A is a $*$ algebra over the real numbers, is their a natural way of extending the real $*$ algebra structure of A to a complex $*$ algebra structure, in accordance with the extension of the real numbers to the complex numbers? There is, and the resulting complex $*$ algebra is called the *complexification* of A and is denoted by $A_{\mathbb{C}}$. Mimicking the extension of the reals to the complex numbers, we construct $A_{\mathbb{C}}$ as follows:

1. As a vector space, $A_{\mathbb{C}}$ is simply $A \times A$. Complex multiplication is given by the injective $*$ homomorphism from $\mathbb{C}$ to $M_2(\mathbb{R})$, $a + ib \mapsto$

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

   together with the natural action of $M_2(\mathbb{R})$ on $A \times A$.

2. Involution in $A_{\mathbb{C}}$ is defined by $(a, b)^* = (a^*, -b^*)$.

3. The map $a \mapsto (a, 0)$ is an injective $*$ homomorphism of A into $A_{\mathbb{C}}$.

It is clear that any element of $A_{\mathbb{C}}$ can be written as $a + ib$ where $a, b \in A$. Hence we define multiplication in $A_{\mathbb{C}}$ by $(a + ib)(c + id) = (ac - bd) + i(ad + bc)$.

An equivalent way of constructing $A_{\mathbb{C}}$ is to use the method of extension of scalars and to form the left $\mathbb{C}$-module $\mathbb{C} \otimes_{\mathbb{R}} A$. Multiplication is given by $(a \otimes x)(b \otimes y) = (ab) \otimes (xy)$, and we define the star structure on $\mathbb{C} \otimes_{\mathbb{R}} A$ by $(a \otimes x)^* = \overline{a} \otimes x^*$ where the star on the right is taken with respect to the real star structure on A. For the proofs of the following results, see [Sri07], Section 3.

Proposition 41 *Let A be a complex $*$ algebra. Then the following holds:*

1. *An element of A is normal if and only if it can be written as $a + ib$ for a and b which are self adjoint and commute.*

2. *A is abelian if and only if every element of A is normal.*

Proposition 42 *Let A be a real $*$ algebra. Then $A_{\mathbb{C}}$ is abelian if and only if every real element of A (that is, every element of the form $(a, 0)$ for $a \in A$) is self adjoint.*

The last proposition will be used to prove the simplicity of the branching of the Bratteli diagram of the symmetric groups, a fundamental result in the work of Vershik and Okounkov.

7.2.8 Finite Dimensional Normed Spaces

The study of finite dimensional normed spaces is relatively simple due to the following collection of facts:

Proposition 43 ([Bol90], Chapter 4) *Let V and W be a finite dimensional normed spaces over the field $\mathbb{K}$. Then :*

1. *All norms on V are equivalent.*

2. *V is isometrically isomorphic to $\mathbb{K}^n$, where n=$\mathrm{Dim}(V)$.*

3. *Every linear map between V and W is continuous.*

4. $\mathrm{End}_{\mathbb{K}}(V, W) \cong M_{m,n}(\mathbb{K})$ *where m and n are the dimensions of V and W respectively.*

Hence the classification of the space $\mathrm{End}(V)$ where V is a finite dimensional normed space is reduced to the classification of various sorts of matrices. In particular, we have the following, which can be found in almost any undergraduate linear algebra textbook (see, for example, [Lan87], Chapter 7):

Proposition 44 *Let $T : V \to V$ be a linear map, where V is finite dimensional. Let A_T be the matrix of T relative to a given base. Then the matrix of T^* is the matrix ${A_T}^*$, where the star on the right denotes the conjugate transpose.*

An important class of subspaces are called *invariant subspaces.* If T is an endomorphism of V, we call a subspace $U \subset V$ *T-invariant* if the image of U under T is contained in U. The restriction of T to such a subspace can clearly be considered as an endomorphism of U.

Proposition 45 ([EW06], Corollary 16.5) *Suppose $T \in \mathrm{End}(V)$ is diagonalizable, and let $U \subset V$ be a T-invariant subspace. Then the restriction of T to U, as an element of $\mathrm{End}(U)$, is also diagonalizable.*

Bibliography

[BB05] A. Bjorner and F. Brenti. *Combinatorics of Coxeter Groups.* Springer Graduate Texts in Mathematics, New York, N.Y, 2005.

[Bol90] Bela Bollobas. *Linear Analysis.* Cambridge University Press, Cambridge, U.K., 1990.

[Bra72] Ola Bratteli. *Inductive Limits of Finite Dimensional C^* Algebras.* Transactions of the American Mathematical Society, Vol 171 (1972), pp.195-234, 1972.

[Bum04] Daniel Bump. *Lie Groups.* Springer - Verlag, New York, New York, 2004.

[Coh80] Donald L. Cohn. *Measure Theory.* Birkhauser, Boston, MA., 1980.

[CST10] Scarabotti F. Ceccherini-Silberstein, T. and F. Tolli. *Representation Theory of the Symmetric Groups: The Okounkov-Vershik Approach, Character Formulas, and Partition Algebras.* Cambridge University Press, New York, New York, 2010.

[DF04] D. Dummit and R. Foote. *Abstract Algebra.* John Wiley and Sons Inc, Vermont, US, 2004.

[EW06] K. Erdmann and M. Wildon. *Introduction to Lie Algebras.* Springer Undergraduate Mathematics Series, London, U.K., 2006.

[FH04] W. Fulton and J. Harris. *Representation Theory: A First Course*. Springer - Verlag, New York, New York, 2004.

[Ful97] William Fulton. *Young Tableaux*. Cambridge University Press, Cambridge, UK, 1997.

[Ker03] S. V. Kerov. *Asymptotic Representation Theory of the Symmetric Group and its Applications in Analysis*. American Mathematical Society, Providence, Rhode Island, 2003.

[Lan87] Serge Lang. *Linear Algebra*. Springer Undergraduate Texts in Mathematics, New York, N.Y., 1987.

[Mac79] I. G. Macdonald. *Symmetric Functions and Hall Polynomials*. Clarendon Press, New York, New York, 1979.

[Nic01] W. K. Nicholson. *Elementary Linear Algebra, Second Edition*. Mc-Graw-Hill Ryerson, New York, N.Y., 2001.

[Sag01] Bruce E. Sagan. *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*. Springer - Verlag, New York, New York, 2001.

[Sri07] Murali K. Srinivasan. Notes on the Vershik-Okounkov Approach to the Representation Theory of the Symmetric Groups. `http://www.bprim.org/vershik_okounkov-murli-new-aug07.pdf`, 2007.

[VO96] A. M. Vershik and A. Y. Okounkov. A New Approach to the Representation Theory of the Symmetric Groups. *Selecta Math. (N.S.)*, 2(4):581–605, 1996.

[VO05] A. M. Vershik and A. Yu. Okounkov. A New Approach to the Representation Theory of the Symmetric Groups II. (Russian). *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 307 (2004), (Teor. Predst.

Din. Sist. Komb. i Algoritm. Metody. 10):5798, 281; translation in J. Math Sci. (N.Y.) 131, no. 2, 54715494., 2005.