APPLICATIONS OF REPRESENTATION THEORY TO COMBINATORICS

ALEX GHORBANI

Abstract. This paper contains an introduction to the representation theory of finite groups, an application of this theory to the symmetric group, and an exploration into the combinatorial nature of objects discussed along the way. We begin with the basic results of representation theory and use them to develop ideas about the symmetric group and its irreducible representations. Young Tableaux are introduced, and two proofs are given for the Hook-Length Formula with an application to the Catalan numbers. We then conclude our treatment of the representation theory of the symmetric group with an overview of Young's Rule and the Kostka Numbers.

Contents

1. Introduction

Representation theory is a field of mathematics that systematically translates objects from the abstract world of group theory to the more concrete environment of vector spaces. This lets mathematicians use the extensive and powerful tools of linear algebra to answer questions about groups and their properties. In this paper, we will look at the basic ways that groups get translated into vector space automorphisms and the important results from the introduction to representation theory that follow. We will apply these results to the study of the symmetric group, which is the group of bijective functions from a set of n objects to itself. The symmetric group plays an important role in the field of combinatorics and, as we will see, some of its combinatorial properties can be understood from the perspective of representation theory.

Date: August 17, 2020.

2. Representations of Groups

Definition 2.1. Let G be a finite group, and let V be a finite dimensional vector space over the complex numbers. We say that a *representation* of G on V is a group homomorphism

$$\rho : G \to \mathrm{GL}(V),$$

where $\mathrm{GL}(V)$ is the set of all invertible linear transformations from V to itself, otherwise known as the general linear group of V.

This map is what lets us translate the elements of our group into linear transformations of a vector space. The vector space V is sometimes called a G-module, because the group acts on V in a way that preserves its abelian structure, and is said to *carry a representation* of G. It is convention to omit ρ and simply write $g\mathbf{v}$ instead of $\rho(g)\mathbf{v}$.

Definition 2.2. Let V be a G-module, and let W be a subspace of V. We say that W is a *submodule* of V if W is a vector subspace of V and is closed under the action of G:

$$g\mathbf{w} \in W \text{ for all } g \in G, \mathbf{w} \in W.$$

Given a vector space V, consider the subspaces $W = V$ and $W = \{0\}$. The reader can verify that these subspaces are indeed submodules, but we call these *trivial* submodules. This idea of the triviality of representations brings us to one of the first major ideas of representation theory.

Definition 2.3. Let V be G-module. V is said to be *irreducible* if V contains only trivial submodules. If V contains non-trivial submodules then V is *reducible.*

Now, we have some conception of "atomic units" of representations, i.e., representations that do not contain other nontrivial representations. Our goal is now to decompose representations into these atomic units, bringing us to our next important idea.

Definition 2.4. Given two G-modules, U and W, there exists a new G-module, $U \oplus W$, where $\oplus$ denotes the usual direct sum of vector spaces.

This means that we can construct new representations from old representations. Here are some examples of group representations.

Example 2.5.

(1) The *trivial representation* of a group G is a map $\rho : G \to \mathbb{C}$ such that,

$$g\mathbf{z} = \mathbf{z} \text{ for all } \mathbf{z} \in \mathbb{C}.$$

(2) Consider a set $S = \{s_1, s_2, \cdots, s_n\}$ on which G acts. Now we can look at the vector space $\mathbb{C}\mathbf{S}$ which consists of formal linear combinations

$$c_1\mathbf{s}_1 + c_2\mathbf{s}_2 + \cdots + c_n\mathbf{s}_n,$$

where $c_i \in \mathbb{C}$ for all i. The elements of S are acting as our basis. The reader can verify that this vector space is indeed a G-module.

If a group G acts on a set S then the G-module $\mathbb{C}\mathbf{S}$ is called the *permutation representation* and the elements of S are called the *natural or standard basis.*

(3) Groups act on themselves, so when we let $S = G$ itself, we can also form a G-module, which we call the *group algebra*, or the *regular representation.*

To begin our discussion of these two important results, we will discuss the maps between representations.

Definition 2.6. Let V and W be G-modules. A *G-homomorphism* or *G-linear map* is a linear transformation $\phi : V \to W$ such that

$$\phi(g\mathbf{v}) = g\phi(\mathbf{v}) \text{ for all } g \in G, \mathbf{v} \in V.$$

Definition 2.7. Let V and W be G-modules. V and W are said to be G-equivalent, written $V \cong W$, if there exists a bijective G-homomorphism $\phi : V \to W$.

Lemma 2.8. *Let V and W be G-modules and let $\phi : V \to W$ be a G-linear map. Then* $\operatorname{Ker}(\phi)$ *is a submodule of V and* $\operatorname{Im}(\phi)$ *is a submodule of W.*

Proof. Because ϕ is a linear transformation, $\operatorname{Ker}(\phi)$ is a subspace of V and $\operatorname{Im}(\phi)$ is a subspace of W. Take an arbitrary $\mathbf{v} \in \operatorname{Ker}(\phi)$ and take an arbitrary $g \in G$. Consider $g\phi(\mathbf{v})$, because our vector is in the kernel, we can say

$$g\phi(\mathbf{v}) = g\mathbf{0} = \mathbf{0} = \phi(g\mathbf{v})$$
$$\iff g\mathbf{v} \in \operatorname{Ker}(\phi).$$

Thus $\operatorname{Ker}(\phi)$ is closed with respect to G.
Similarly, take an arbitrary $\mathbf{w} \in \operatorname{Im}(\phi)$. Then,

$$\phi(g\mathbf{v}) = g\phi(\mathbf{v}) = g\mathbf{w}$$

Because $\operatorname{Im}(\phi)$ is a submodule of W, $g\mathbf{w} \in \operatorname{Im}(\phi)$, so $\operatorname{Im}(\phi)$ is closed with respect to G. □

This gives us the following important result.

Lemma 2.9 (Schur's Lemma)**.** *Let V and W be irreducible G-modules. If $\phi : V \to W$ is a G-linear map, then exactly one of the following holds.*

(1) ϕ is a G-isomorphism.
(2) ϕ is the zero map.

Proof. By Lemma 2.3, we have that $\operatorname{Ker}(\phi)$ is a submodule of V. We also know that V is irreducible, so $\operatorname{Ker}(\phi)$ is either $\{\mathbf{0}\}$ or V. Similarly, because W is also irreducible, we have that $\operatorname{Im}(\phi)$ is also either $\{\mathbf{0}\}$ or W.
If $\operatorname{Ker}(\phi) = V$ or $\operatorname{Im}(\phi) = \{\mathbf{0}\}$, then ϕ must be the zero map. If $\operatorname{Ker}(\phi) = \{\mathbf{0}\}$ and $\operatorname{Im}(\phi) = W$, then ϕ is an isomorphism. □

Lemma 2.10. *Let V be a G-module with W as a submodule. Then there exists some submodule U complementary to W so that $V = W \oplus U$.*

Before the proof of this lemma, we will introduce some technical definitions.

Definition 2.11. An inner product on V is *invariant* under the action of G if

$$\langle g\mathbf{v}, g\mathbf{w}\rangle = \langle \mathbf{v}, \mathbf{w}\rangle,$$

for all $g \in G$ and $\mathbf{v}, \mathbf{w} \in V$.

Definition 2.12. Given an inner product on a vector space V and a subspace $W \leq V$, we can form the *orthogonal complement of W*, denoted $W^\perp$, as follows:

$$W^\perp = \{\mathbf{v} \in V \mid \langle \mathbf{v}, \mathbf{w}\rangle = 0 \text{ for all } \mathbf{w} \in W\}.$$

It is always true that $V = W \oplus W^\perp$ when W is a subspace. We can construct a G-invariant inner product from an arbitrary inner product on V by the following definition. Given an arbitrary inner product $\langle \cdot, \cdot \rangle'$ on V, we can define a G-invariant inner product by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{g \in G} \langle g\mathbf{u}, g\mathbf{v} \rangle'.$$

Now, we have the tools to prove the lemma.

Proof. Let $U = W^\perp$. We want to show that, for all $g \in G$ and $\mathbf{u} \in W^\perp$, we have $g\mathbf{u} \in W^\perp$. Take an arbitrary $\mathbf{w} \in W$. Then, we can say

$$\langle g\mathbf{u}, w \rangle = \langle g^{-1}g\mathbf{u}, g^{-1}\mathbf{w} \rangle,$$

since the inner product is G-invariant. By the properties of the group action, we get

$$\langle g^{-1}g\mathbf{u}, g^{-1}\mathbf{w} \rangle = \langle \mathbf{u}, g^{-1}\mathbf{w} \rangle.$$

We know that $\mathbf{u} \in W^\perp$ and $g^{-1}\mathbf{w} \in W$ so we can say

$$\langle g\mathbf{u}, \mathbf{w} \rangle = 0.$$

Thus $g\mathbf{u} \in W^\perp$, giving us that $W^\perp$ is a submodule of V. Finally, we can write $V = W \oplus W^\perp$. □

The following theorem lets us talk about the decomposition of any representation in terms of irreducible subrepresentations.

Theorem 2.13 (Maschke's Theorem)**.** *Let G be a finite group and let V be a non-zero G-module. Then,*

$$V = W^{(1)} \oplus W^{(2)} \oplus \cdots \oplus W^{(k)},$$

where each (not necessarily distinct) $W^{(i)}$ is an irreducible submodule of V.

Proof. We will use induction on the dimension of V. If $\dim(V) = 1$, then V is irreducible, and we are done. Now assume this is true for $\dim(V) = n$, and we will study the case where $\dim(V) = n+1$. If V is irreducible, we are done. Otherwise, V contains at least one nontrivial submodule, W. Then, $V = W \oplus W^\perp$ where $\dim(W)$, $\dim(W^\perp)$ are each less than n. This means we can use induction, decomposing each piece into irreducibles, and we are done. □

Definition 2.14. We say that a representation V is *completely reducible* if it can be written as the direct sum of irreducible submodules.

Maschke's Theorem tells us that every representation of a finite group over the complex numbers is completely reducible. This lets us study group representations in terms of its irreducibles, which is often a simpler task.

3. Representations of the Symmetric Group

In this section, we will use the tools of representation theory that we developed in the last section to talk about the symmetric group.

Definition 3.1. Suppose $\lambda = (\lambda_1, \lambda_2, \lambda_3, \cdots, \lambda_l)$ partitions n, written $\lambda \vdash n$. This means that $n = \lambda_1 + \lambda_2 + \cdots + \lambda_l$, where $n \geq \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l > 0$. The *shape* of λ is an array of n dots with l left-justified rows with row i containing λ_i dots for $1 \leq i \leq l$.

Example 3.2. Let $\lambda = (4, 2, 2, 1)$ partition 10. The corresponding shape is

$$\mathrm{sh}(\lambda) = \begin{array}{cccc} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & & \\ \bullet & \bullet & & \\ \bullet & & & \end{array} .$$

Now we can associate a given $\lambda \vdash n$ with a subgroup of S_n.

Definition 3.3. Let $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_l)$ partition n. The corresponding *Young subgroup* of S_n is

$$S_\lambda = S_{\{1,2,\cdots,\lambda_1\}} \times S_{\{\lambda_1+1,\lambda_1+2,\cdots,\lambda_1+\lambda_2\}} \times \cdots \times S_{\{n-\lambda_l+1,n-\lambda_l+2,\cdots,n\}}$$

Definition 3.4. Let λ partition n. A *Young diagram, or tableau, of shape* λ is an array t obtained by replacing the dots of shape λ with the numbers $1, 2, \cdots, n$ bijectively.

Example 3.5. Let $\lambda = (4, 3, 2, 1)$ be a partition of 10. A Young Tableau t of shape λ might look like

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

We will now introduce an equivalence relation on the tableaux.

Definition 3.6. Two Young-tableaux t_1 and t_2 are *row equivalent*, $t_1 \sim t_2$, if corresponding rows of the two tableaux contain the same elements. A *tabloid of shape* λ is then

$$\{t\} = \{t_1 \mid t_1 \sim t\}$$

where the shape of t is λ.

Example 3.7. Suppose we have the following two tableaux, t_1 and t_2 respectively

$$t_1 = \begin{array}{|c|c|c|} \hline 3 & 1 & 4 \\ \hline 2 & & \\ \cline{1-1} \end{array}, \quad t_2 = \begin{array}{|c|c|c|} \hline 4 & 3 & 1 \\ \hline 2 & & \\ \cline{1-1} \end{array}.$$

These two tableaux are row equivalent, and the corresponding tabloid is given by

$$\{t\} = \begin{array}{ccc} \hline 1 & 3 & 4 \\ \hline 2 & & \\ \cline{1-1} \end{array}.$$

Definition 3.8. Let λ partition n. Define

$$M^\lambda = \mathbb{C}\{\{\mathbf{t_1}\}, \cdots, \{\mathbf{t_k}\}\},$$

where $\{t_1\}, \cdots, \{t_k\}$ is a complete list of λ-tabloids. M^λ is, in fact, naturally an S_n-module. Then, M^λ is called the *permutation module corresponding to* λ.

Definition 3.9. A G-module M is *cyclic* if there is a $\mathbf{v} \in M$ such that

$$M = \mathbb{C}\left[G\mathbf{v}\right],$$

where $G\mathbf{v} = \{g\mathbf{v} \mid g \in G\}$. In this case, we say that M is *generated* by $\mathbf{v}$.

Proposition 3.10. *If λ partitions n, then M^λ is cyclic, generated by any given λ-tabloid.*

Proof. We know that any tabloid of shape λ can be taken to another tabloid of shape λ by some permutation in S_n, so M^λ is cyclic. □

Theorem 3.11. *Let λ partition n. Consider the Young subgroup S_λ and an arbitrary tabloid $\{t^\lambda\}$. Then the representations $V^\lambda = \mathbb{C}\left[S_n/S_\lambda\right]$ and $M^\lambda = \mathbb{C}\left[S_n\{\mathbf{t}^\lambda\}\right]$ are S_n-isomorphic.*

Proof. We can map between the Young subgroup

$$S_\lambda = S_{\{1,2,\cdots,\lambda_1\}} \times S_{\{\lambda_1+1,\lambda_1+2,\cdots,\lambda_1+\lambda_2\}} \times \cdots \times S_{\{n-\lambda_l+1,n-\lambda_l+2,\cdots,n\}}$$

and the tabloid

$$\{t^\lambda\} = \begin{array}{l} \hline 1 \quad 2 \quad \cdots \quad \lambda_1 \\ \hline \lambda_1+1 \quad \lambda_1+2 \quad \cdots \quad \lambda_2 \\ \hline \cdots \\ \hline n-\lambda_l+1 \quad \cdots \quad n \\ \hline \end{array}.$$

From this, we can think of the coset πS_λ corresponding with the tabloid $\{\pi t^\lambda\}$. So, let $\pi_1, \pi_2, \cdots, \pi_k$ be a transversal (a set containing exactly one element from each coset) for S_λ. Consider the map

$$\theta : V^\lambda \to M^\lambda$$

such that $\theta(\pi_\mathbf{i}\mathbf{S}_\lambda) = \{\pi_\mathbf{i}\mathbf{t}^\lambda\}$ for $1 \leq i \leq k$ and vice-versa. Because θ preserves the action of permutation, our map is indeed an S_n-homomorphism. The correspondence above is bijective, so our map is also an S_n-isomorphism. □

In our examples of representations in the first section, we have already seen examples of these modules.

Example 3.12.

a) If $\lambda = (n)$, then

$$M^{(n)} = \mathbb{C}\left\{ \begin{array}{l} \hline 1 \quad 2 \quad \cdots \quad n \\ \hline \end{array} \right\} \cong \mathbb{C}$$

which gives us the trivial representation.

b) If $\lambda = (1^n)$, then each equivalence class $\{t\}$ consists of only one tableau. We can think of this tableau as a permutation written in one-line notation, but transposed. The action of S_n is preserved by this correspondence, so we can say

$$M^{(1^n)} \cong \mathbb{C}\left[S_n\right].$$

This gives us the regular representation.

c) If $\lambda = (n-1, 1)$, then each λ-tabloid is uniquely determined by the element in the second row of the shape. Thus, each tabloid can just be thought of as a number $1, 2, \cdots, n$. The action of S_n is preserved by this correspondence, so we can say

$$M^{(n-1,1)} \cong \mathbb{C}\{\mathbf{1}, \mathbf{2}, \cdots, \mathbf{n}\}.$$

This gives us the defining representation.

4. Hook-Length Formula

In this section, we will discuss the notion of standard tableaux and use the Hook Length Formula to count the number of standard tableaux with a given partition $\lambda \vdash n$.

Definition 4.1. A tableau t is *standard* if the rows and columns of t are increasing sequences. We say f^λ is the number of standard tableau of shape λ.

Definition 4.2. If (i, j) is a node in the diagram of λ, then it has *hook*

$$H_{i,j} = \{(i, j') \mid j' \geq j\} \cup \{(i', j) \mid i' \geq i\}$$

with corresponding *hook length*

$$h_{i,j} = |H_{i,j}|.$$

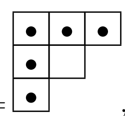
The *arm* and *leg* of $H_{i,j}$ are

$$A_{i,j} = \{(i, j') \mid j' \geq j\} \text{ and } L_{i,j} = \{(i', j) \mid i' \geq i\},$$

respectively, where *arm length* and *leg length* are given as

$$\mathrm{al}_{i,j} = |A_{i,j}| \text{ and } \mathrm{ll}_{i,j} = |L_{i,j}|.$$

Example 4.3. Let $\lambda = (3, 2, 1)$ partition 6. The hook of cell $(1, 1)$ is given by,

$H_{1,1} =$  ,

giving $h_{1,1} = 5$.

The Hook Formula is motivated by the fact that the number of standard Young tableaux of a given shape, f^λ, must be less than the number of total Young tableaux of a given shape, $n!$. After experimentation, the following expression was produced.

Theorem 4.4 (Hook Formula)**.** *If $\lambda \vdash n$, then*

$$f^\lambda = \frac{n!}{\prod_{(i,j)\in\lambda} h_{i,j}}.$$

Before proving this formula, we will rewrite the formula so that is easier to create a bijective proof:

$$n! = f^\lambda \prod_{(i,j)\in\lambda} h_{i,j}.$$

We know that $n!$ is the number of Young tableau of a given shape, and we know that f^λ is the number of standard tableau of a given shape. So we want our bijection to map from a Young tableau, T, to a pair (P, J) where P is a standard tableau and J is a tableau of shape λ such that the number of choices for the cell $J_{i,j}$ is $h_{i,j}$, where the shapes of T, P, J are all λ.

Now we will give a total order to the cells in a tableau. We say that

$$(i, j) \leq (i', j') \text{ if and only if } j > j' \text{ or } j \leq j' \text{ and } i \leq i'.$$

Then label the cells as $c_1 < c_2 < c_3 < \cdots < c_n$. Given a tableau T, we then define the tableau $T^{\leq c}$ to be the tableau with all the cells c' of T such that $c' \leq c$.

Our algorithm will construct a sequence of pairs (T_i, J_i) with the first pair being $(T, 0)$[1] and the final pair being is (P, J). Each iteration of the algorithm produces a new tableau, $j^c(T)$, as follows:

1) Pick c such that $T^{<c}$ is a standard tableau.
2) **While** $T^{\leq c}$ is not a standard tableau **do**
   a) If $c = (i, j)$, then let c' be the cell of $\min\{T_{i+1,j}, T_{i,j+1}\}$.
   b) Exchange T_c and T'_c and let $c := c'$.

The process terminates when a standard tableau, P, is reached.

Definition 4.5. The sequence of cells that c passes through is called the *path of* c.

If the path of j^{c_k} starts at (i, j) and ends at (i', j') after one iteration, then we have $J_k = J_{k-1}$, except for the positions

$$(J_k)_{h,j} = \begin{matrix} (J_{k-1})_{h+1,j} - 1 & \text{for } i \leq h < i' \\ j' - j & \text{for } h = i' \end{matrix} .$$

To prove that the mapping is a bijection, we will create an inverse. To see that this inverse algorithm is well-defined or to see examples, see [2] or [1].

We first consider the set of candidate cells for the end of the path of $j^{c_k}(T_{k-1})$. We define this set to be

$$C_k = \{(i', j') \mid i' \geq i_0, j' = j_0 + (J_k)_{i',j_0}, (J_k)_{i',j_0} \geq 0\}.$$

For ease of notation, let $(T', J') = (T_k, J_k)$. Then our algorithm is as follows

1) Pick $c \in C_k$ where $c_k = (i_0, j_0)$.
2) **While** $c \neq c_k$ **do**
   a) If $c = (i, j)$, then let c' be the cell of $\max\{T'_{i-1,j}, T'_{i,j-1}\}$ where $T'_{k,l} = 0$ if $k < 0$ or $l < j_0$.
   b) Exchange T'_c and $T'_{c'}$ and let $c := c'$.

5. Probabilistic Proof of the Hook-Length Formula

Besides the combinatorial proof given above, there are many other proofs of the hook-length formula. In this section, we will look at a short probabilistic proof of the formula.

Let λ be a partition of n, and let t be a λ-shaped standard tableau. Because of the condition that rows and columns are strictly increasing, the maximal element n in t is in one of the corner cells of the tableau. For our proof, we will remove a corner from t to get a new standard tableau which we will say has shape μ, written as $\mu \prec \lambda$.

Now, we introduce the notation

$$e_\lambda = \frac{n!}{\prod h_{i,j}}$$

to denote the ratio between the total number of permutations and the multiplied hook lengths of the partition λ.

If we can establish that the following identity is true, we can use induction over the size of each of the smaller tableaux until we reach a base case:

$$e^\lambda = \sum_{\mu \prec \lambda} e^\mu,$$

[1] Here, 0 is an empty hook tableau.

where μ ranges over shapes obtained by removing a corner from the shape λ. We can rewrite the above as

$$1 = \sum_{\mu \prec \lambda} \frac{e^{\mu}}{e^{\lambda}}.$$

Now we will look at the corners of t in detail. A random cell (i, j) in t has a $\frac{1}{n}$ chance of being chosen. Now consider the hook, $H_{i,j}$, and a random cell, $(i', j') \neq (i, j)$, within it. That cell has a $\frac{1}{h_{i,j}-1}$ chance of being chosen. Repeat the process with a new cell chosen from $H_{i',j'}$. Eventually, a corner will be chosen and the process will stop. This cell, (x, y), is called the terminal cell; any corner can be a terminal cell. Let $p(x, y)$ be the probability that (x, y) is a terminal cell for a given trial of the process.

Consider the path, P, consisting of cells $(a, b) = (a_1, b_1), (a_2, b_2), \cdots, (a_m, b_m) = (x, y)$, which records a trial starting at (a, b) and ending at the terminal cell (x, y).

Definition 5.1. The *vertical* and *horizontal projections of* P are the sequences $A = (a_1, a_2, \cdots, a_m)$ and $B = (b_1, b_2, \cdots, b_m)$ respectively.

Denote the probability that a random trial beginning at cell (a, b) and having the projections (A, B) by $p(A, B \mid a, b)$.

Lemma 5.2.

$$p(A, B \mid a, b) = \prod_{i \in A \setminus \{a\}} \frac{1}{h_{ib} - 1} \cdot \prod_{j \in B \setminus \{b\}} \frac{1}{h_{aj} - 1}.$$

Proof. We can rewrite the righthand side as

$$p(A, B \mid a, b) = \frac{1}{h_{ab} - 1} \left(p(A \setminus \{a_1\}, B \mid a_2, b) + p(A, B \setminus \{b_1\} \mid a, b_2) \right).$$

Using induction over m gives us that

$$p(A \setminus \{a_1\}, B \mid a_2, b) = (h_{a_1 y} - 1) \prod_{i \in A \setminus \{x\}} \frac{1}{h_{iy} - 1},$$

and

$$p(A, B \setminus \{b_1\} \mid a, b_2) = (h_{x b_1} - 1) \prod_{j \in B \setminus \{y\}} \frac{1}{h_{xj} - 1}.$$

Therefore,

$$p(A, B \mid a, b) = \frac{1}{h_{ab} - 1} \left((h_{ay} - 1) \prod_{i \in A \setminus \{x\}} \frac{1}{h_{iy} - 1} + (h_{x b_1} - 1) \prod_{j \in B \setminus \{y\}} \frac{1}{h_{xj} - 1} \right).$$

From the identity $h_{ab} - 1 = (h_{a_1 y} - 1) + (h_{x b_1} - 1)$, we get finally get that

$$p(A, B \mid a, b) = \prod_{i \in A \setminus \{x\}} \frac{1}{h_{iy} - 1}, \prod_{j \in B \setminus \{y\}} \frac{1}{h_{xj} - 1},$$

and we are done. □

Theorem 5.3. *Let* (x, y) *be a terminal cell, then* $p(x, y) = \frac{e^{\mu}}{e^{\lambda}}$.

Proof. We can expand the righthand side of the equation to

$$\frac{e^\mu}{e^\lambda} = \frac{1}{n} \prod_{1\le i<x} \frac{h_{iy}}{h_{iy}-1} \prod_{1\le j<y} \frac{h_{xj}}{h_{xj}-1}$$
$$= \frac{1}{n} \prod_{1\le i\le x} \left(1+\frac{1}{h_{iy}-1}\right) \prod_{1\le j\le y} \left(1+\frac{1}{h_{xj}-1}\right)$$

The probability $p(x,y)$ can be computed by summing over the probability of the first cell chosen and then summing over all possible vertical and horizontal projections from that cell, giving us

$$p(x,y) = \frac{1}{n} \sum_{(A,B)} p(A,B \mid a,b).$$

By the above lemma, expanding this sum results in the expanded righthand expression for $\frac{e^\mu}{e^\lambda}$. □

Corollary 5.4. $\sum_\mu \frac{e^\mu}{e^\lambda} = 1$.

Proof. A given trial always terminates, so the sum of the probabilities, $p(x,y)$, adds up to 1. □

This completes our proof.

6. Catalan Numbers

The Catalan numbers are an important sequence of numbers in combinatorics; for a list of combinatorial interpretations of the Catalan numbers, see [3]. We can use the hook-length formula to very quickly derive a formula for the Catalan numbers.

Consider a standard tableau of shape (n,n), which partitions $2n$. The hook-length of a given cell, i, in the top row of the tableau is $i+1$. Similarly, the hook-length for a given cell, j, on the bottom row is j. Letting i,j range from 1 to n, we have

$$f^{(n,n)} = \frac{(2n)!}{(n+1)!n!} = \frac{1}{n+1}\binom{2n}{n},$$

which is the formula for the n-th Catalan number. This suggests that we can find bijections between objects counted by the Catalan numbers and standard tableaux of shape (n,n).

7. Specht Modules and the Decomposition of M^λ

Definition 7.1. Given a tableau t with rows $R_1, R_2, \cdots, R_l$ and columns $C_1, C_2, \cdots, C_k$, define

$$R_t = S_{R_1} \times S_{R_2} \times \cdots \times S_{R_l}, \quad C_t = S_{C_1} \times S_{C_2} \times \cdots \times S_{C_k}$$

to be the *row-stabilizer* and *column-stabilizer* of t, respectively, in S_n.

We can see that a tabloid can be written as $R_t t$, because each row can be permutated internally with the elements in each row remaining constant.

Now we define a new object, called a *polytabloid* as a formal sum of tabloids. We introduce the element

$$k_t = \sum_{\pi \in C_t} \operatorname{sgn}(\pi)\pi \in \mathbb{C}\left[S_n\right].$$

Given a tableau t, we define the associated *polytabloid* to be

$$e_t = k_t\{t\}.$$

Definition 7.2. Given a partition λ, the associated *Specht module*, S^λ, is the submodule of M^λ spanned by the polytabloids e_t, where the shape of t is λ.

The Specht modules have the nice property of being cyclic, just like M^λ. We can see that from

$$e_{\pi t} = k_{\pi t}\{\pi t\} = \pi k \pi^{-1}\{\pi t\} = \pi k_t\{t\} = \pi e_t.$$

Now we will cite some important theorems about Specht modules with the goal of showing that they are the full set of irreducible submodules of M^λ. For a detailed look at the proofs of these results, see [2].

First we will introduce an order on partitions.

Definition 7.3. Let $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_l)$ and $\mu = (\mu_1, \mu_2, \cdots, \mu_m)$ be two partitions of n. We say λ *dominates* μ, $\lambda \trianglerighteq \mu$, if

$$\lambda_1 + \lambda_2 + \cdots + \lambda_i \geq \mu_1 + \mu_2 + \cdots + \mu_i$$

for all $i \geq 1$, where $\lambda_i = 0$ if $i > l$, similarly for μ.

The following two results are used to show that we have a complete list of irreducibles of M^λ.

Theorem 7.4 (Submodule Theorem)**.** *Let U be a submodule of M^λ. Then*

$$U \supseteq S^\lambda \text{ or } U \subseteq S^{\lambda\perp}.$$

Moreover, the S^λ are irreducible.

Proposition 7.5. *Let θ be a nonzero G-homomorphism from S^λ to M^μ. Then $\lambda \trianglerighteq \mu$. If $\lambda = \mu$, then θ is multiplication by a scalar.*

Now, we can prove the following theorem.

Theorem 7.6. *Let λ partition n. The S^λ form a complete list of irreducible S_n-modules.*

Proof. By the Submodule Theorem, the S^λ are irreducible. We also know that there is a one-to-one correspondence between the Specht Modules S^λ and conjugacy classes of S_n, as desired, by counting the irreduicible representations. Now, we must show that the Specht modules are pairwise inequivalent. If $S^\lambda \cong S^\mu$, then there is a G-homomorphism between S^λ and M^μ, by the Submodule Theorem. Then by the above proposition, we can say $\lambda \trianglerighteq \mu$. Conversely, we can conclude that $\mu \trianglerighteq \lambda$. Thus $\lambda = \mu$, and the modules are the same. □

Theorem 7.7. *(Decomposition of M^μ) A permutation module M^μ decomposes as*

$$M^\mu = \bigoplus_{\lambda \trianglerighteq \mu} m_{\lambda\mu} S^\lambda,$$

for some $m_{\lambda\mu} \in \mathbb{N}_0$.

Proof. This decomposition follows from Proposition 4.6 because if S^λ appears inside M^μ with a nonzero coefficient, then $\lambda \trianglerighteq \mu$. □

We will see that the coefficients, $m_{\lambda\mu}$, have a combinatorial interpretation later.

8. Basis for S^λ and Kostka Numbers

If we want to select a natural basis for the Specht modules, we need to find a spanning set of linearly independent polytabloids. There is a clever way of choosing such a collection, namely to choose the set of polytabloids created from standard tableaux. We have already seen that standard tableaux are interesting objects and this fact furthers that assumption.

Theorem 8.1. *Given a partition λ of n, the set*

$$\{e_t \mid t \text{ is a standard tableau of shape } \lambda\}$$

forms a basis of S^λ.

Proof. A complete treatment of this theorem can be found in [2]. Here, we will just look at a sketch of the proof.

First, to prove that the set is linearly independent, a partial ordering is given for tabloids. Then, we consider a set of m vectors in M^λ, $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m\}$. If, for each $\mathbf{v}_i$, we can choose a maximal tabloid based on our partial order, $\{\mathbf{t_i}\}$, such that all the $\{\mathbf{t_i}\}$ are distinct, then those m vectors are linearly independent.

Paired with some technical lemmas meant to bridge the gap between tabloids and polytabloids, we can conclude that the set of polytabloids constructed from the standard tableaux of a given shape is linearly independent.

To show that the set spans S^λ, we introduce a process called the straightening algorithm. This process lets us take an arbitrary polytabloid as a linear combination of standard polytabloids.

We start with an arbitrary tableau t. We can assume that t has increasing columns, because we can always find a permutation within the column stablizer that will give us increasing columns on t.

Then we want to find a set of permutations, $\{\pi\}$, such that

(1) in each tableau πt, pairs of adjacent, out of order elements in a row are eliminated, and
(2) there is a special element, called the Garnir element,

$$g = e + \sum_{\pi \in \{\pi\}} \operatorname{sgn}(\pi)\pi$$

which satisfies $g\mathbf{e_t} = 0$.

Then, we can write

$$\mathbf{e_t} = -\sum_{\pi} \mathbf{e}_{\pi\mathbf{t}}.$$

This process is repeated to give us the desired linear combination. Each iteration gives a linear combination consisting of polytabloids that are closer to being standard than those from the last iteration. Once they are all standard, we can write any polytabloid as the linear combination of standard polytabloids. □

Because the standard tableaux form a basis for a given Specht module, the dimension can be given by the following corollary.

Corollary 8.2. *For any partition of n, λ,* $\dim(S^\lambda) = f^\lambda$.

Definition 8.3. A *semistandard tableau of shape λ* is like a Young tableau with slightly different conditions. The first condition is that the numbers in the boxes can repeat and the second condition is that the rows weakly increase while the columns strictly increase. The *content* of a semistandard tableau is represented by $\mu = (\mu_1, \mu_2, \cdots, \mu_m)$ where μ_i is the number of i's in the tableau.

Definition 8.4. Given two partitions of n, λ and μ, the *Kostka number* $K_{\lambda\mu}$ is the number of semistandard tableaux of shape λ and content μ.

As a consequence of the semistandard basis theorem (see [2]), we can give the following theorem, known as Young's Rule.

Theorem 8.5. *(Young's Rule) The multiplicity of S^λ in M^μ is equal to the number of semistandard tableaux of shape λ and content μ:*

$$M^\mu \cong \bigoplus_{\lambda \trianglerighteq \mu} K_{\lambda\mu} S^\lambda.$$

Now we give some examples of $K_{\lambda\mu}$ for familiar partitions λ, μ.

Proposition 8.6. *Given a partition μ of n, $K_{\mu\mu} = 1$.*

Proof. The only way to construct a semistandard tableau of shape μ and content μ would be to fill the first row with all 1's, the second row with all the 2's, and so on. □

Proposition 8.7. *Given a partition μ of n, $K_{(n)\mu} = 1$.*

Proof. There is only one way to arrange a collection of numbers in weakly increasing order, so there is only 1 semistandard tableau of this specific configuration. □

The above proposition, in conjunction with Young's rule, implies that M^μ contains exactly one copy of $S^{(n)}$, which is the trivial representation.

Proposition 8.8. *Given a partition λ of n, $K_{\lambda(1^n)} = f^\lambda$.*

Proof. A semistandard tableau with content (1^n) is a standard tableau because there will be no repetitions of numbers in the boxes, and the rows and columns will both increase strictly. So, the number of semistandard tableaux of shape λ and content (1^n) is equal to the number of standard tableaux of shape λ. □

Young's rule then says that

$$M^{(1^n)} \cong \bigoplus_{\lambda} f^\lambda S^\lambda.$$

We know that $M^{(1^n)}$ is just the regular representation. If we take the dimension of both sides of this expression, we find that

$$n! = \sum_{\lambda} (f^\lambda)^2.$$

9. Acknowledgments

First and foremost, I would like to thank my mentor Calder Sheagren for helping me learn the material in this paper and for helping me edit my writing. And I would like to thank Pallav Goyal and Peter Huxford for looking over my paper. I would also like to thank Daniil Rudenko for running the apprentice program and writing fun and challenging problems sets, as well as Peter Huxford and Alicia Xiao for hosting the apprentice problem sessions. Finally, I would like to thank Peter May for organizing this year's REU.

References

[1] J.-C. Novelli, I. Pak, and A. V. Stoyanovskii. A direct bijective proof of the hook-length formula.

[2] B. E. Sagan. *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*. Springer, 2000.

[3] R. P. Stanley. Catalan addendum. 2013.