COMBINATORICS AND ITS CONNECTION WITH REPRESENTATION THEORY

CHEN XU

Abstract. In this paper, we study two objects on which the symmetric group acts: posets, with particular emphasis on the Boolean algebra, and the permutation modules in representation theory. We connect the study of these two objects via understanding the degrees of irreducible modules using combinatorial methods. Along the way, we investigate how some other combinatorial constructions have interesting applications in other mathematical fields.

Contents

1. Introduction

In this introduction, we explain why we study the mathematical objects mentioned in the abstract, present a roadmap that will guide readers through the body of this paper, and briefly discuss the various applications this study has.

The purpose of this paper is two-fold:

(1): First, we want to show readers how elementary abstract algebra and linear algebra can solve problems in combinatorics and representation theory. Thus, the use of linear algebra and group actions for studying sets with various properties will be frequent. Hopefully, we will attract more readers into this beautiful mathematical world that connects various elementary and advanced topics.

(2): Second, after an elementary understanding of some combinatorial objects has been established, we motivate the discussion of the more advanced topic of representation theory. Studying representation theory will become natural once one appreciates the intimate connection between it and combinatorics. It is remarkable how simple counting can reduce the complexity and abstractness of representation theory.

Date: DEADLINES: Draft AUGUST 14 and Final version AUGUST 28, 2017.

Since this paper is devoted to understanding the symmetric group S_n, which acts on posets and modules, **the process of study follows two paths:** in the first four sections, we concentrate on understanding various properties of partially ordered sets via studying B_n, the Boolean algebra. We consider its quotient posets and a general counting formula called *Polya's theorem*. Once we have proven the theorem, which actually lies in representation theory, we make the transition in the last section into the study of irreducible representations for S_n. We will show how the Standard Young Tabloids (SYT) from combinatorics allow the study of irreducible modules.

We summarize the basic strategies used in proving the main results in each section in one or two sentences:

Section 1: To prove the Sperner property, we will define order-matchings and order-raising operators and show how the one-to-oneness of those operators produces the results we want.

Section 2: Use extensively Burnside's formula to figure out inequivalent objects (orbits) under group action by understanding how *fixed colorings* correspond to the type of a cycle π.

Section 3: Regard each walk U or D as a linear transformation and prove a similar identity we had for order-raising (or lowering) operators in Section 1.

Section 4: Define *Permutation modules* and their submodules (*Specht modules*) and prove in succession that

1. Specht modules constitute the whole set of irreducible modules for S_n.
2. Standard polytabloids form *bases* for Specht modules and connect this section to Section 3 by deducing $\dim(S^\lambda)=f^\lambda$ as a corollary.

At the end of each section on combinatorics (this paper focuses primarily on combinatorics), we will also introduce briefly how the main theorems previously proven extend to other mathematical fields; they shine new lights on solving rather difficult problems in **graph theory, number theory, and definitely, representation theory**.

Finally, it should be noted that sections in this paper follow in logic and structure Chapter 4,5,7, and 8 of Stanley's *Algebraic Combinatorics* and Chapter 2 of Sagan's *The Symmetric Group* but are explained in more intuitive and succinct ways so that the connection between the two theories is accessible.

2. The Sperner Property of the Boolean Algebra B_n and Quotient Poset B_n/G

Certain partially ordered sets (posets) have special properties. In this section and the next, we study one special property, the Sperner property, of the Boolean algebra B_n and its quotient poset B_n/G. We use Polya's theorem to count the number of elements within a set or poset under a group action. Our proof of the Sperner Property for B_n involves more lemmas and propositions than necessary (i.e. compare with the proof by Lubell, [Stanley, Chapter 4, p.**38**]) but works in greater generality; it works also for B_n/G. After presenting the proofs, some questions mentioned in the introduction will be answered. The proofs in this section follow Stanley [**1**, Chapter 4 and 5].

First, we give a sequence of definitions:

Definition 2.1. A *poset* P is a finite set, together with a binary relation denoted $\leq$ satisfying three axioms: reflexivity ($x \leq x$ for all x in P), antisymmetry (If $x \leq y$ and $y \leq x$ then $x = y$), and transitivity (if $x \leq y$ and $y \leq z$, then $x \leq z$.) To represent posets pictorially, The *Hasse diagram* of a poset P is a planar drawing, with elements of P drawn as dots. If $x < y$ in P (i.e. $x \leq y$ and $x \neq y$), then y is drawn "above" x (i.e., with a larger vertical coordinate). An edge is drawn between x and y if y covers x, i.e., $x < y$ and no element z satisfies $x < z < y$. We then write $x \lessdot y$). Note: The Hasse Diagram determines the relations of P by the cover relations due to transitivity. A *chain* C in a poset is a totally ordered subset of P, i.e., if $x, y \in C$ then either $x < y$ or $y < x$ in P.

Definition 2.2. A finite poset is *graded of rank* n if every *maximal chain has length* n, namely, consisting of n elements. (A chain is maximal if it is contained in no larger chain.). A chain $y_0 < y_1 < ... < y_j$ is saturated if each y_{i+1} covers y_i . If P is graded of rank n and $x \in P$, then we say that x has *rank* j, denoted $p(x) = j$, if the largest saturated chain of P with top element x has length j. The collection of all x with rank i is called the *i-th level of* P, denoted P_i. (Note, if P is graded of rank n then it can be written as the disjoint union of its levels). We write p_i as the number of elements of P of rank i and say a graded poset P of rank n is *rank-symmetric* if $p_i = p_{n-i}$ for $0 \leq i \leq n$ and *rank-unimodal* if $p_0 \leq p_1 \leq ... \leq p_j \geq p_{j+1} \geq p_{j+2} \geq ... \geq p_n$ for some $0 \leq j \leq n$.

Definition 2.3. Also, an *antichain* in a poset P is a subset A of P for which no two elements are comparable. Lastly, let P be a graded poset of rank n. We say that P has the *Sperner property* if no antichain is larger than the largest level P_i (in terms of the number of elements in the antichain A).

Example 2.4. Let P be any collection of sets. If $x, y \in P$, then define $x \leq y$ in P if $x \subset y$ as sets. This definition is a partial order on P. If P consists of *all* subsets of an n-element set $\{1, 2, ..., n\}$, then P is called a (finite) *Boolean algebra of rank* n and is denoted by B_n. Also, $P_j = \{x \subset \{1, 2, ..., n\} : |x| = j\}$ and $p_j = \binom{n}{j}$. Since the binomial coefficients $\binom{n}{0}, \binom{n}{1}, ..., \binom{n}{n}$ are unimodal and symmetric, B_n is rank-symmetric and rank-unimodal. The image below shows the Hasse diagram for B_3, the Boolean algebra on the 3-element set.

$$
\begin{array}{ccc}
 & \{1,2,3\} & \\
\{2,3\} & \{1,3\} & \{1,2\} \\
\{3\} & \{2\} & \{1\} \\
 & \{\emptyset\} &
\end{array}
$$

Certain graded posets P are guaranteed to be Sperner. We define an *order-matching* from P_i to P_{i+1} to be a *one-to-one* function $f : P_i \to P_{i+1}$ satisfying $x < f(x)$ for all $x \in P_i$. Similarly, we define an *order-matching* from P_i to P_{i-1} to be a *one-to-one* function $f : P_i \to P_{i-1}$ satisfying $x > f(x)$ for all $x \in P_i$.

Proposition 2.5. *Let P be a graded poset of rank n. Suppose there exists an integer $0 \leq j \leq n$ and order-matchings*

$$P_0 \to P_1 \to P_2 \to ...P_{j-1} \to P_j \leftarrow P_{j+1} \leftarrow P_{j+2} \leftarrow ... \leftarrow P_n.$$

Then P is rank-unimodal and Sperner

Proof. We define a subgraph G of the Hasse diagram of P so that vertices of this subgraph are elements of P and the graph itself is a disjoint union of chains (represented by edges connecting these vertices) that pass through all the vertices. In that case, we have pushed G into p_j chains so that any antichain A can intersect the chains of G at most once, $|A| \leq p_j$. More specifically, we define G by the rule that its vertices are elements of P and $x, y \in G$ are connected if and only if there is an order-matching P_i (as assumed in the hypothesis) such that $P_i(x) = y$. Note that G satisfies the description above by construction. □

Enlightened by this proposition, we aim to find order-matchings from P_i to P_{i+1} (or $i-1$) in the Boolean algebra B_n. Before doing so, recall that for any (finite) set S, $\mathbb{R}S$ denotes the real vector space consisting of all formal linear combinations (with real coefficients) of elements of S. Thus, S is a basis for $\mathbb{R}S$. The lemma below ensures the existence of those order-matchings and will be a crucial one to which we frequently refer:

Lemma 2.6. *Suppose there exists a linear transformation $U : \mathbb{R}P_i \to \mathbb{R}P_{i+1}$ satisfying*

1. U is one-to-one and

2. For all $x \in P_i$, $U(x)$ is a linear combination of elements $y \in P_{i+1}$ satisfying $x < y$ (based on 2, we call U an order-raising operator).

Then there exists an order-matching $p : P_i \to P_{i+1}$

Similarly, suppose there exists a linear transformation $U : \mathbb{R}P_i \to \mathbb{R}P_{i+1}$ satisfying

1. U is onto and

2. U is an order-raising operator.

Then there exists an order-matching $p : P_{i+1} \to P_i$

Proof. We consider here the case where U is one-to-one. The case when U is onto will be evident after this proof. Write the matrix $[U]$ of U in terms of basis vectors $x_1, ..., x_a \in P_i$ of $\mathbb{R}P_i$ and $y_1, ..., y_b \in P_{i+1}$ of $\mathbb{R}P_{i+1}$. Since U is one-to-one, it has p_i linearly independent columns (rows). Hence, by changing the indices, if necessary, we obtain an invertible matrix $p_i \times p_i$ A from $[U]$ by choosing the first p_i rows of $[U]$. We note

$$det(A) = \sum \pm a_{1\pi(1)} a_{2\pi(2)} ... a_{p_i \pi(p_i)},$$

where the sum is over all permutations π of p_i. At least one summand in this expansion is non-zero. Hence, since $a_{i\pi(i)} \neq 0$ $\forall i$, $U(x_{\pi(i)})$ then defines the order-matching $p : P_i \to P_{i+1}$ by $p(x_i) = y_{\pi^{-1}(i)}$.

The case when U is onto is proven similarly by considering the transpose of $[U]$ defined above. □

Now, we apply **Proposition 2.5** and **Lemma 2.6** to the boolean algebra B_n. For each $0 \leq i \leq n$, we define $U_i : \mathbb{R}B_i \to \mathbb{R}B_{i+1}$ such that $U_i(x) =$

$\sum_{y\in(B_n)_{i+1} \ and \ y>x} y$, the sum of the set of y in $(B_n)_{i+1}$ that cover x. By definition, U_i is order-raising and we want to show U_i is one-to-one if $i < n/2$ and onto if $i \geq n/2$. We introduce dual operators $D_i : \mathbb{R}B_i \to \mathbb{R}B_{i-1}$ to the U_i's by $D_i(x) = \sum_{y\in(B_n)_{i-1} \ and \ y<x} y$, the sum of the set of y in $(B_n)_{i-1}$ that x covers. Notice, by considering the indices for these matrices, with respect to the standard basis,

$$[D_{i+1}] = [U_i]^t,$$

which will be a crucial identity. For the next lemma, set $U_n = 0$ and $D_0 = 0$.

Lemma 2.7. *Let $0 \leq i \leq n$, then*

$$D_{i+1}U_i - U_{i-1}D_i = (n-2i)I_i,$$

the identity transformation on $\mathbb{R}B_i$.

Proof. We have

$$D_{i+1}U_i(x) = D_{i+1}(\sum_{|y|=i+1 \ and \ x\subset y} y) = \sum_{|y|=i+1 \ and \ x\subset y} \sum_{|z|=i \ and \ z\subset y} z. \tag{2.8}$$

If $x, z \in (B_n)_i$ satisfy $|x \cap z| < i-1$, where $|\cdot|$ denotes the rank of the input, then there is no $y \in (B_n)_{i+1}$ such that $x \subset y$ and $z \subset y$ (hence coefficient of the expansion is 0). If $|x \cap z| = i-1$, then $y = x \cup z$. Finally if $x = z$ then y can be any element of $(B_n)_{i+1}$ containing x, so there are $n-i$ such y in all. Therefore, we can rewrite $D_{i+1}U_i(x)$ as

$$D_{i+1}U_i(x) = (n-i)x + \sum_{|z|=i \ and \ |z\cap x|=i-1} z. \tag{2.9}$$

Similarly, $U_{i-1}D_i(x) = ix + \sum_{|z|=i \ and \ |z\cap x|=i-1} z$. Subtracting these two yields $D_{i+1}U_i - U_{i-1}D_i(x) = (n-2i)I_i(x)$. □

Theorem 2.10. *The operator U_i defined above is one-to-one if $i < n/2$ and is onto if $i \geq n/2$.*

Proof. By the previous lemma, $D_{i+1}U_i = (n-2i)I_i + U_{i-1}D_i$. By the observation that $[D_i] = [U_{i-1}]^t$ and using that in general AA^T is positive semi-definite with nonnegative eigenvalues, we see that eigenvalues of $D_{i+1}U_i$ are strictly positive when we add each eigenvalue of $U_{i-1}D_i$ by $(n-2i)$, $i < \frac{n}{2}$. Thus, $D_{i+1}U_i$ is invertible as its determinant is nonvanishing and so U_i is one-to-one. Similarly, U_iD_{i+1} is invertible based on $U_iD_{i+1} = D_{i+2}U_{i+1} + (2i+2-n)I_{i+1}$ so U_i is onto analogously. □

Finally, since we have those one-to-one and onto order-raising operators U_i, we have the order-matchings satisfying **Proposition 2.5**. Therefore, given the presentation up to this point, one might regard the following corollary as the theorem:

Corollary 2.11. *The boolean algebra B_n has the Sperner Property.*

Next, we consider when G is subgroup of S_n, the symmetric group on $1, ..., n$, and G acts on the boolean algebra B_n by permuting entries of the elements of B_n, which are subsets of the n-element set $[n]$, $[n]=\{1, 2, ..., n\}$. Quotienting by the action we then obtain B_n/G which also has the Sperner Property. More precisely, we have the definitions below:

Definition 2.12. An *automorphism* π of a poset P to P is an order-preserving bijection, where $x \leq y$ in P if and only if $\pi(x) \leq \pi(y)$ in P.

For the case where $P = B_n$, any permutation π of $\{1, 2, ..., n\}$ acts on B_n as follows: if $x = \{i_1, i_2, ..., i_k\} \in B_n$, then $\pi(x) = \{\pi(i_1), \pi(i_2), ..., \pi(i_k)\}$. In particular, if $|x| = i$, then $|\pi(x)| = i$. This action of S_n on B_n thus defines a *group action* of $G = S_n$ on B_n. Note that π is also an automorphism. The following definition thus naturally arises:

Definition 2.13. The elements in the *quotient poset* B_n/G where G is a subgroup of S_n are orbits of G. If o and o' are two orbits, then define $o \leq o'$ in B_n/G if there exist $x \in o$ and $y \in o'$ such that $x \leq y$ in B_n.

Note, this relation $\leq$ on the quotient poset is a partial order. Also, we have the following proposition:

Proposition 2.14. *The quotient poset B_n/G defined above is graded of rank n and rank-symmetric.*

Proof. By the definition above, the rank of an element $o \in B_n/G$ inherits the rank of any one of the elements $x \in o$. Hence, B_n/G is graded of rank n. By passing to the complement of $x \in o$, denoted as $\bar{x}$, we realize that $\{x_1, ..., x_m\}$ is an orbit of x if and only if $\{\bar{x}_1, ..., \bar{x}_m\}$ is an orbit of $\bar{x}$ so the number of orbits of rank i in B_n/G equals to the number of complementary orbits of rank $(n - i)$ in B_n/G. So B_n/G is rank-symmetric.

□

Let $\pi \in S_n$. We associate with π a linear transformation $\pi : \mathbb{R}(B_n)_i \to \mathbb{R}(B_n)_i$ by the rule

$$\pi(\sum_{x \in (B_n)_i} c_x x) = \sum_{x \in (B_n)_i} c_x \pi(x).$$

where each c_x is a real number. Note that the two π's are different but have the same notation. This defines an action of S_n, or any subgroup G of S_n, on the vector space $\mathbb{R}(B_n)_i$. We want to understand the elements of $\mathbb{R}(B_n)_i$ which are fixed by every element of a subgroup G of S_n. We denote the set of such elements as $\mathbb{R}(B_n)_i^G$, which by definition equals to $\{v \in \mathbb{R}(B_n)_i : \pi(v) = v \text{ for all } \pi \in G\}$ (notice we are working with linear transformations here).

We have the lemma below:

Lemma 2.15. *A basis for $\mathbb{R}(B_n)_i^G$ consists of the elements*

$$v_o := \sum_{x \in o} x,$$

where $o \in (B_n)_i/G$, the set of $G-$orbits for the action of G on $(B_n)_i$.

Proof. First, $\pi(v_o) = v_o$ for all $\pi \in G$ since π only permutes the order of elements in v_o. Thus, $v_o \in \mathbb{R}(B_n)_i^G$ and are linearly independent because different orbits are disjoint.

Given $v = \sum_{x \in (B_n)_i} c_x x \in \mathbb{R}(B_n)_i^G$, we want to show v is in the span of v_o. Apply $\pi \in G$ to x, denote G_x as the set of stabilizers of x and observe the fact that

$\sum_{\pi\in G}\pi(x) = |G_x|\cdot v_{Gx}$ for $x \in (B_n)_i$ since each y in the orbit of x appears $|G_x|$ times. Hence, since $\pi(v) = v$ by assumption,

$$|G|\cdot v = \sum_{\pi\in G}\pi(v) = \sum_{\pi\in G}\Big(\sum_{x\in(B_n)_i} c_x\pi(x)\Big) = \sum_{x\in(B_n)_i} c_x\Big(\sum_{\pi\in G}\pi(x)\Big) = \sum_{x\in(B_n)_i} c_x\cdot|G_x|\cdot v_{Gx}.$$

Dividing by $|G|$ expresses v as a linear combination of the elements v_{Gx}, as desired. $\square$

We also show that applying the order-raising operator U_i to an element v in $\mathbb{R}(B_n)_i^G$ generates an element in $\mathbb{R}(B_n)_{i+1}^G$.

Lemma 2.16. *If $v \in \mathbb{R}(B_n)_i^G$, then $U_i(v) \in \mathbb{R}(B_n)_{i+1}^G$.*

Proof. Note that since $\pi \in G$ is an automorphism of B_n, we have $x < y$ in B_n if and only if $\pi(x) < \pi(y)$ in B_n. It follows that if $x \in (B_n)_i$, then $U_i(\pi(x)) = \pi(U_i(x))$ (the set of y that cover $\pi(x)$ is the same as the set of $\pi(y)$ for each y covering x). Therefore, by linearity, $U_i(\pi(v)) = \pi(U_i(v))$ for all $v \in \mathbb{R}(B_n)_i$. In other words, $U_i\pi = \pi U_i$. Therefore, if $u \in \mathbb{R}(B_n)_i^G$, then $\pi(U_i(u)) = U_i(\pi(u)) = U_i(u)$. So $U_i(u) \in \mathbb{R}(B_n)_{i+1}^G$. $\square$

We now prove the main result that B_n/G has the Sperner property.

Theorem 2.17. *Let G be a subgroup of S_n. Then the quotient poset B_n/G is graded of rank n, rank-symmetric, rank-unimodal, and Sperner.*

Proof. The key to this proof is the following diagram (where the isomorphisms follow from **Lemma 2.15**):

$$\begin{tikzcd}
\mathbb{R}(B_n)_i^G \arrow[r, "U_i"] \arrow[d, "\cong"'] & \mathbb{R}(B_n)_{i+1}^G \arrow[d, "\cong"] \\
\mathbb{R}(B_n/G)_i \arrow[r, "\bar{U}_i"'] & \mathbb{R}(B_n/G)_{i+1}
\end{tikzcd}$$

Notice that if the diagram commutes, we can define $\bar{U}_i$ properly by considering U_i so that $\bar{U}_i$'s give order-raising operators.

Precisely, let $P = B_n/G$ and define the maps for $\bar{U}_i$ and U_i as follows:

$$U_i : \mathbb{R}(B_n)_i \to \mathbb{R}(B_n)_{i+1} \ s.t.\ U_i(v_o) = \sum_{o'\in(B_n)_{i+1}/G} c_{o,o'}v_{o'}.$$

$$\bar{U}_i : \mathbb{R}P_i \to \mathbb{R}P_{i+1} \ s.t.\ \bar{U}_i(o) = \sum_{o'\in(B_n)_{i+1}/G} c_{o,o'}o',$$

so $\bar{U}_i(o)$ inherits the coefficients. Here, U_i is the usual order-raising operator defined after **Lemma 2.6** and we extend $U_i(v_o)$ by linearity. Notice that $\bar{U}_i$ satisfies conditions of **Lemma 2.6** by the following two observations:

$\bar{U}_i$ is order-raising since $v'_o = \sum_{x'\in o'} x'$ and the only way that $c_{o,o'}$ is nonzero is for some $x' \in o'$ to satisfy $x' > x$ for some $x \in o$. Then, $o' > o$ when the coefficient $c_{o,o'}$ is nonzero.

Also, since U_i is one-to-one for $i < n/2$, its restriction to the subspace $\mathbb{R}(B_n)_i^G$ is one-to-one. Hence, due to the isomorphism between $\mathbb{R}(B_n)_i^G$ and $\mathbb{R}(B_n/G)_i$, we know that $\bar{U}_i$ is also one-to-one when $i < n/2$.

We omit an analogous proof for $\bar{D}_i$, which guarantees the existence of an order-matching $p : P_{i+1} \to P_i$. $\square$

We reiterate here for clarity that throughout the proofs of the Sperner property for B_n and B_n/G, our main efforts lay in defining those order-raising operators and proving their one-to-oneness. Indeed, this strategy works in a wide range of cases, for instance, in the proof of the *weak Erdös-Moser conjecture* whose details can be found in Stanley [**1**, Chapter 6, p.70].

Let us consider the application of the above theorem to graph theory. Let $M = \{1, .., m\}$. Set $X = \binom{M}{2}$, the set of all two-element subsets of M (think of X as (possible) edges of a simple graph with vertex set M). If B_X is the boolean algebra of all subsets of X, then an element $x \in B_X$ is a collection of edges on the vertex set M (or a simple graph on M). Define a subgroup G of S_X as follows: if $\pi \in S_M$, then define $\bar{\pi} \in S_X$ by $\bar{\pi}(\{i, j\}) = \{\pi(i), \pi(j)\}$ (we permute the edges by permuting vertices it connects). So G consists of all permutations of the $\binom{M}{2}$ edges that are induced from permutations of the vertices M.

The elements of B_X/G are isomorphism classes of simple graphs on the vertex set M. In particular, $|(B_X/G)_i|$ is the number of nonisomorphic such graphs with i edges. Thus, based on the previous theorem, we have the highly nontrivial theorem below as an easy application:

Theorem 2.18. *(a): Fix $m \geq 1$. Let p_i be the number of nonisomorphic simple graphs with m vertices and i edges. Then the sequence $p_0, p_1, ..., p_{\binom{m}{2}}$ is symmetric and unimodal.*

(b): Let T be a collection of simple graphs with m vertices such that no element of T is isomorphic to a spanning subgraph of another element of T. Then the number of T is maximized by taking T to consist of all nonisomorphic simple graphs with $\left\lfloor \frac{1}{2}\binom{m}{2} \right\rfloor$ edges.

3. Polya's Theorem for Enumeration Under Group Action

For this section, we present a general theory for *enumerating objects under a group of symmetries*, which will allow us to calculate the numbers p_i of the number of elements of rank i in the quotient poset B_n/G mentioned in the previous section. The machinery to do this is called *Polya's theorem.* We approach it from the standpoint of *colorings* of some geometric or combinatorial object. The proofs in this section follow Stanley [**1**, Chapter 7].

Definition 3.1. A *coloring* of a finite set X is a function $f : X \to C$ where C denotes the set of colors (which may be infinite). Given G, a subgroup of S_X, the symmetric group on X, define f and g to be G-equivalent if there exists an element $\pi \in G$ such that $g(\pi(x)) = f(x)$ for all $x \in X$.

More compactly, we write $g\pi = f$ when g and f are equivalent colorings on X. It can be checked that the above definition for equivalent colorings is an equivalence relation. We notice several facts:

(1): Whenever G is the trivial group then two colorings are equivalent if and only if they are the same.

(2): The size of a class of a coloring f is the index in G of the subgroup fixing some fixed coloring in that class, as follows from the orbit-stabilizer formula and knowing that a coloring is determined once we know its value on elements in X.

(3): For any set X if the group G is the symmetric group S_X then two colorings are equivalent if and only each color appears the same number of times.

Our objective in general is to count the number of equivalence classes of colorings which use each color *a specified number of times* and apply the results to understanding how to count the number of inequivalent necklaces under rotational symmetry and the number of elements of a given rank in the quotient algebra B_n/G. For convenience, we put the relevant information into a *generating function*, a polynomial whose coefficients are the numbers we seek:

Example 3.2. Consider the case where X is a 2×2 chessboard and $C=\{r,b,y\}$. Let G be the dihedral group of all rotations and reflections of X. Think of r, b, y as *variables* and form the polynomial

$$F_G(r,b,y)=\sum_{i+j+k=4}\tau(i,j,k)r^ib^jy^k \tag{3.3}$$

where $\tau(i,j,k)$ is the number of inequivalent colorings when we use r i times, b j times, and y k times, in a total of 4 times. The reader should check that $F_G(r,b,y)=(r^4+b^4+y^4)+(r^3b+rb^3+r^3y+ry^3+b^3y+by^3)+2(r^2y^2+r^2b^2+b^2y^2)+2(r^2by+rb^2y+rby^2)$. For example, the coefficient of r^4 is 1 because there is only one coloring (under G action) using the color r four times. Notice the fact that $F_G(r,b,y)$ is a symmetric function because it does not matter how we name the colors r,b,y as long as the summation of their exponents adds to 4. Hence, setting $r=b=y=1$, one sees that the total number of inequivalent colorings with four colors is $F_G(1,1,1)=3+6+2\cdot 3+2\cdot 3=21$.

The basic tool for counting the number of inequivalent objects subject to a group of permutations is a result from group theory, called *Burnside's lemma*, whose proof can be found in any elementary group theory book.

Theorem 3.4. *(Burnside's lemma). Let Y be a finite set and G a subgroup of S_Y. For each $\pi\in G$, let Fix(π)=$\{y\in Y:\pi(y)=y\}$, so $|Fix(\pi)|$ is the number of cycles of length one in the permutation π. Let Y/G be the set of orbits of G. Then $|Y/G|=\frac{1}{|G|}\sum_{\pi\in G}|Fix(\pi)|$.*

In other words, the average number of elements of Y fixed by an element of G is equal to the number of orbits.

Example 3.5. How many inequivalent colorings of the vertices of a regular hexagon H are there using n colors, under cyclic symmetry, i.e. rotation? Let C_n denote the set of all n-colorings of H. Let G be the group of cyclic symmetries (so $G\cong\mathbb{Z}_6$). We want to know the number of orbits of G. Therefore, let π be a generator of G and we want to figure out its fixed points.
If τ is identity: all n^6 colorings are fixed by τ.
If τ is π,π^{-1}: Only n colorings with all colors equal are fixed.
If τ is π^2,π^4: any coloring of the form *ababab* is fixed so there are n^2 colorings in all.
If τ is π^3: colorings of the form *abcabc* are fixed, so n^3 in total.
By Burnside's lemma, we have the number of orbits, $\frac{1}{6}(n^6+n^3+2n^2+2n)$.

One might note in the final expression above that the coefficient in front of each n^i is equal to the number of elements in G with i cycles. For example, there are two permutations (π and π^{-1}) with only one cycle. We prove a general theorem of this kind below, using only the number of cycles a permutation has.

Theorem 3.6. *Let G be a group of permutations of a finite set X. Then the number $N_G(n)$ of inequivalent n-colorings (under G action) of X is given by*

$$N_G(n) = \frac{1}{|G|} \sum_{\pi \in G} n^{c(\pi)},$$

where $c(\pi)$ denotes the number of cycles of π.

Proof. Let π act on the set of n-colorings of X. If a coloring f belongs to $\mathrm{Fix}(\pi)$, then $f(x) = \pi \cdot f(x) = f(\pi(x))$. Hence, $f \in \mathrm{Fix}(\pi)$ if and only if $f(x) = f(\pi^k(x))$ for all $k \geq 1$. Notice that $\pi^k(x)$ are elements in a cycle of π. Therefore, we need f to color each element in the cycle the same color. In total, there are $n^{c(\pi)}$ choices for these f so π fixes $n^{c(\pi)}$ colorings. Applying Burnside's formula generates the result.

□

Now, we want to *specify the number of occurrences of each color.* Before doing so, we define the *type* of π as

$$\mathit{type}(\pi) = (c_1, ..., c_n)$$

where π has c_i i-cycles and we define the *cycle indicator* of π to be the monomial $Z_\pi = z_1^{c_1} z_2^{c_2} ... z_n^{c_n}$. Note that $\sum_i ic_i = n$. For example, if π is (1,4,8)(2,7,11,5)(3)(6,10,9), then π has the type (1,0,2,1) and indicator $z_1 z_3^2 z_4$.

Given a subgroup G of S_X, the cycle indicator of G is defined by

$$Z_G = Z_G(z_1, ..., z_n) = \frac{1}{|G|} \sum_{\pi \in G} Z_\pi,$$

so Z_G is a polynomial in the variables $z_1, ..., z_n$.

We now prove the main result:

Theorem 3.7. *(Polya's Theorem): Let G be a group of permutations of the n-element set X. Let $C = \{r_1, r_2, ...\}$ be a set of colors. let $\tau(i_1, i_2, ...)$ be the number of inequivalent (under G action) colorings $f : X \to C$ such that color r_j is used i_j times. Define*

$$F_G(r_1, r_2, ...) = \sum_{i_1, i_2, ...} \tau(i_1, i_2, ...) r_1^{i_1} r_2^{i_2} ....$$

Then

$$F_G(r_1, r_2, ...) = Z_G(r_1 + r_2 + r_3 + ..., r_1^2 + r_2^2 + r_3^2 + ..., ...).$$

In other words, substitute $\sum_i r_i^j$ for z_j in Z_G.

Proof. We want to apply Burnside's formula to count the number of inequivalent colorings with color r_j used i_j times. Therefore, consider the sequence $\mathbf{i}$=$(i_1, i_2, ...)$ where $i_1 + i_2 + ... = |X| = n$. Apply π to the set of $C_\mathbf{i}$ colorings, which are colorings using the color r_j i_j times and find $|Fix(\pi_\mathbf{i})|$.

Consider the product

$$H_\pi = \prod_j (r_1^j + r_2^j + ...)^{c_j(\pi)},$$

where we substitute $\sum_i r_i^j$ for z_j in the cycle indicator of π. After we expand out this product, we see that

$$\prod_j (r_1^j + r_2^j + ...)^{c_j(\pi)} = H_\pi = \sum_{\mathbf{i}} |Fix(\pi_{\mathbf{i}})| r_1^{i_1} r_2^{i_2} ....,$$

whose precise interpretation is as follows: we have colored each cycle of π the same and monomials of the form $r_1^{i_1} r_2^{i_2}$.... are colorings that use each color the designated times and are fixed by π. Hence, not only have we found $|Fix(\pi_{\mathbf{i}})|$ for *the specific sequence* $\mathbf{i}$, we in fact have found *all possible colorings for all sequences* $\mathbf{i}$ that π fixes. Now, summing both expressions over all $\pi \in G$ and dividing by $|G|$, we see that the left-hand side becomes:

$$\frac{1}{|G|} \sum_{\pi \in G} \prod_j (r_1^j + r_2^j + ...)^{c_j(\pi)} = Z_G(r_1 + r_2 + r_3 + ..., r_1^2 + r_2^2 + r_3^3 + ..., ...)$$

and the right-hand side becomes

$$\sum_{\mathbf{i}} [\frac{1}{|G|} \sum_{\pi \in G} |Fix(\pi_{\mathbf{i}})|] r_1^{i_1} r_2^{i_2} ...$$

By Burnside's lemma, the expression in brackets is the number of inequivalent colorings using color r_j a total of i_j times and if we sum over all sequences $\mathbf{i}$, we get the desired result as stated in the theorem. $\square$

Notice before we move on that suppose if n colors are available and if we substitute $r_i = 1$ for all i, then Polya's theorem counts the total number of inequivalent n-colorings, proving **Theorem 3.6** as an immediate corollary. This is true because if we replace each r_i with 1, we no longer care about the specific occurrence of any single color but only about the number of times a combination of colorings occurs.

We present two applications of Polya's theorem, one to counting the number of inequivalent *n-colored* necklaces of length l, another to counting the number of elements of a given rank in a quotient poset B_X/G.

Example 3.8. We want to generalize **Example 3.5** in which we discussed the n-coloring of a necklace of length 6. In general, a necklace has length l and $\pi = (1, 2, ..., l)$ is the generator of the cyclic group $G = \{1, \pi, ...\pi^{l-1}\}$. If d is the greatest common divisor of m and l, then π^m has d cycles of length l/d. An integer m satisfies $1 \le m \le l$ and gcd(m, l)=d if and only if $1 \le m/d \le l/d$ and gcd$(m/d, l/d)$=1. The number of such m can be calculated by using the Euler phi-function $\phi(l/d)$ where

$$\phi(k) = k \prod_{p|k} (1 - \frac{1}{p}),$$

where p in the above sum is a *prime divisor* of k (this formula can be proved using inclusion-exclusion principle). Therefore, the cycle indicator $Z_G(z_1, ..., z_l)$ is:

$$Z_G(z_1, ..., z_l) = \frac{1}{l} \sum_{d|l} \phi(l/d) z_{l/d}^d,$$

or (substituting l/d for d),

$$Z_G(z_1, ..., z_l) = \frac{1}{l} \sum_{d|l} \phi(d) z_d^{l/d}.$$

We thus have the two theorems below:

Theorem 3.9. *(a) The number $N_l(n)$ of n-colored necklaces of length l is given by*

$$N_l(n) = \frac{1}{l} \sum_{d|l} \phi(l/d) n^d$$

(b) When we specify the use of each color,

$$F_G(r_1, r_2, ...) = \frac{1}{l} \sum_{d|l} \phi(d)(r_1^d + r_2^d + ...)^{l/d}.$$

Regarding counting the number of elements of a given rank in a quotient poset B_X/G, we observe the following connection between colorings and subsets:

When B_X is the boolean algebra of all subsets of a finite set X and G is a group of permutations of X (with induced action on B_X), a two-coloring $f : X \to \{0, 1\}$ corresponds to a subset S_f of X by the rule:

$$s \in S_f \Leftrightarrow f(s) = 1.$$

Note that two 2-colorings f and g are G-equivalent if and only if S_f and S_g are in the same orbit of G (we know how f or g color X if we know S_f or S_g, respectively). Therefore, the cardinality $|B_X/G|_i$ (or the number of orbits of rank i) equals to the number of inequivalent 2-colorings f of X with i spots of X equal to 1 because in this case a coloring with i values equal to 1 is the same as a subset of X with i elements. We immediately have the corollary below:

Corollary 3.10. $\sum_i \#(B_X/G)_i q^i = Z_G(1 + q, 1 + q^2, 1 + q^3, ..)$

Proof. Let $\tau(i, j)$ denotes the number of inequivalent 2-colorings of X with color x and y such that x is used i times and y j times (so $i + j = |X|$). By Polya's theorem, $\sum_{i+j=|X|} \tau(i, j)x^i y^j = Z_G(x + y, x^2 + y^2, x^3 + y^3, ...)$ as polynomials in x and y. Therefore, setting $x = q$ and $y = 1$ maintains the equality and yields the result. □

Notice, using the rank-unimodality and rank-symmetry of B_X/G, we have the following result: *For any finite group G of permutations of a finite set X, the polynomial $Z_G(1+q, 1+q^2, 1+q^3, ..)$ has symmetric, unimodal, integer coefficients.* This is an interesting result that is worth mentioning.

4. Counting Walks in Young's Lattice

Before transiting into representation theory of the symmetric groups, we take a detour and count the number of Hasse walks of a given type in *Young's lattice*, a Hasse digram whose vertices are partitions of positive integers. Not only will this discussion further reinforce the readers' understanding of certain combinatorial strategies, the results in this section will also motivate the discussion of representation theory whose connections with combinatorics are often intimate. The proofs in this section follow Stanley [**1**. Chapter 8].

As usual, we give some definitions:

Definition 4.1. A *partition* of an integer $n \geq 0$ is a sequence $\lambda = (\lambda_1, \lambda_2, ...)$ of integers $\lambda_i \geq 0$ satisfying $\lambda_1 \geq \lambda_2 \geq ...$ and $\sum_{i \geq 1} \lambda_i = n$. Thus all but finitely many λ_i are 0. The *Young diagram* of a partition λ is a left-justified array of squares, with λ_i squares in the ith row.

It is those Standard Young Tableaux (or SYT) that interest us the most and relate mostly closely with representation theory and walk counting.

Definition 4.2. A *SYT* τ consists of the Young diagram D of some partition λ of an integer n, together with the numbers $1, 2, ..., n$ inserted into the squares of D, so that each number appears exactly once, and the numbers in every row and column are increasing. In this case λ is called the *shape* of the SYT τ.

For instance, there are five SYT of shape (2,2,1). From now on, we call f^λ the number of SYT of shape λ. A general formula for the number of SYT of a certain partition is given by the *hook-length formula*, whose statement and proof we omit here.

Now, Young's lattice, Y, is the poset of all partitions of all nonnegative integers, ordered by containment of their Young diagrams (one Young diagram contains another if every row of the contained digram is shorter than the corresponding row containing it). We are interested here in counting certain walks in the Hasse diagram (considered as a graph) of Y.

Note that Y is also a graded poset (of infinite rank) with Y_i consisting of all partitions of i, so Y is the disjoint union of all Y_i's. We also call Y_i the ith level of Y.

We call a walk w in the Hasse diagram of a poset (in this case Y) a Hasse walk, which is composed of letters U and D, denoting "up" and "down," respectively. In fact, as we did in Section 2, we define and use U_i and D_i as linear transformations from Y_i to Y_{i+1} or Y_{i-1}:

$$U_i(\lambda) = \sum_{\mu \vdash i+1 \ and \ \lambda < \mu} \mu$$

$$D_i(\lambda) = \sum_{v \vdash i-1 \ and \ \lambda > v} v.$$

If the walk w has steps of type $A_1, A_2, ..., A_n$, where each A_i is either U or D, with subscripts i omitted, then we say w is of type $A_n A_{n-1} ... A_1$. *Notice the type of w is written in opposite order to that of the walk*

We now discuss the connection between SYT and counting walks in Young's lattice. If w is some word $A_n A_{n-1} ... A_1$ in U and D and λ partitions n, then we write $\alpha(w, \lambda)$ for the number of Hasse walk in Y of type w which starts at the empty partition $\emptyset$ and end at λ. For instance $\alpha(UDUU, 11) = 2$, where $\lambda = (11) \vdash 2$ and $\vdash$ is the symbol for partitioning. In particular, we notice two facts:

$$\alpha(U^n, \lambda) = f^\lambda;$$

$$\alpha(D^n U^n, \emptyset) = \sum_{\lambda \vdash n} (f^\lambda)^2.$$

The first equality follows because each SYT of shape λ, $\lambda \vdash n$, is a distinct U^n walk from $\emptyset$ to n. The second equality follows because given a λ which partitions n, there are f^λ ways to walk from $\emptyset$ to n and f^λ ways to return to the starting point. Summing over all λ generates the result.

We want to find a general formula for $\alpha(w, \lambda)$ of the form $f^\lambda c_w$ where c_w does not depend on λ. We first present a necessary and sufficient condition for the existence of any Hasse walk of type w from $\emptyset$ to λ

Lemma 4.3. *Suppose $w = D^{a_n}U^{b_n}...D^{a_1}U^{b_1}$, where $a_i, b_i \geq 0$. Let $\lambda \vdash n$. Then there exists a Hasse walk of type w from $\emptyset$ to λ if and only if:*

$$\sum_{i=1}^{n}(b_i - a_i) = n,$$

$$\sum_{i=1}^{j}(b_i - a_i) \geq 0 \ for\ 1 \leq j \leq n.$$

Proof. Since each U moves up one level and each D moves down one level, we see that $\sum_{i=1}^{n}(b_i - a_i)$ is the level at which a walk of type w beginning at $\emptyset$ ends. Hence $\sum_{i=1}^{n}(b_i - a_i) = |\lambda| = n$.

After $\sum_{i=1}^{j}(b_i + a_i)$ steps we will be at level $\sum_{i=1}^{j}(b_i - a_i)$. Since the lowest level of Y is 0, we must have $\sum_{i=1}^{j}(b_i - a_i) \geq 0$ for $1 \leq j \leq n$. □

The details of the proof that these two conditions are sufficient are omitted. One merely needs to observe that for $\lambda \vdash n$ and a word w satisfying the condition for the lemma, we can fill up the parts of λ successively from right to left. For example, if $\lambda = (3, 2) \vdash 5$ and $w = DU^5DU^2$, then the walk can be $(0, 0) \to (0, 2) \to (0, 1) \to (0, 2) \to (4, 2) \to (3, 2)$.

Based on how we defined U_i and D_i, it is clear that if r is the number of *distinct* (unequal) parts of λ, then $U_i(\lambda)$ is a sum of $r + 1$ terms and $D_i(\lambda)$ is a sum of r terms. The next lemma resembles **Lemma 2.7** for B_n.

Lemma 4.4. *For any $i \geq 0$ we have*

$$D_{i+1}U_i - U_{i-1}D_i = I_i,$$

the identity transformation on $\mathbb{R}Y_i$.

Proof. As usual, we apply the left-hand side, i.e. $D_{i+1}U_i - U_{i-1}D_i$, to a given partition of i and expand in terms of the basis Y_i. Suppose we act on a partition $\mu \in \mathbb{R}Y_i$. If μ differs from another basis partition λ in a way that we can add a square and subtract a necessarily different one to obtain λ from μ, then this process can be reversed and the coefficient in the final expansion is zero; if μ differs from another partition τ such that we cannot obtain τ via the method we just described, then the coefficient is also zero. Finally, when λ is μ, then there are $r + 1$ ways to add one square and then delete it and r ways to delete it and add it back in. Thus, applying the left hand side to each basis partition gives the result.

□

We now prove the central theorem of this section:

Theorem 4.5. *Let λ be a partition and $w = A_n...A_1$ be a valid λ word (whose existence is assumed by **Lemma 4.3**). Let $S_w = \{i : A_i = D\}$. For each $i \in S_w$, let a_i be the number of D's in w to the right of A_i and let b_i be the number of U's in w to the right of A_i. Then*

$$\alpha(w, \lambda) = f^{\lambda} \prod_{i \in S_w}(b_i - a_i).$$

Proof. Consider the special case where w consists exclusively of parts of the form DU^i, $i \in \mathbb{Z}$, and $w = DU^aDU^b$. The general case follows because we can write D^k as $DD...DD$ for k times and interchange D and U^a each time we see a single D. We interchange D and U^a as follows: based on **Lemma 4.4**, we can prove by induction on i that $DU^i = U^iD + iU^{i-1}$. Therefore, $w(\emptyset) = DU^aDU^b(\emptyset) = DU^a(U^bD + bU^{b-1})(\emptyset) = (DU^{a+b}D+bDU^{a+b-1})(\emptyset) = (b(a+b-1)U^{a+b-2})(\emptyset) = (b(a+b-1))f^\lambda$, where $(DU^{a+b}D + bDU^{a+b-1})(\emptyset) = (b(a+b-1)U^{a+b-2})(\emptyset)$ by $DU^{a+b}D(\emptyset) = 0$ and $\alpha(U^n, \lambda) = f^\lambda$. □

We are finally able to compute the value of $\alpha(D^nU^n, \emptyset) = \sum_{\lambda \vdash n}(f^\lambda)^2$

Corollary 4.6. $\alpha(D^nU^n, \emptyset) = \sum_{\lambda \vdash n}(f^\lambda)^2 = n!$

Proof. When $w = D^nU^n$, we have $S_w = \{n+1, n+2, ..., 2n\}, a_i = i - n - 1$, and $b_i = n$, from which we see the result because $f^\emptyset = 1$. □

As we transit into representation theory, we want to ultimately establish that those f^λ are *the degrees of the irreducible representations of the symmetric group* S_n and that this corollary is a special case of the result that the sum of the squares of the degrees of the irreducible representations of a finite group G is equal to the order of G. Before doing so, we present another bijective proof, called the RSK algorithm, of the above result for the sake of its ingenuity.

Define a *near Young tableau* (NYT) to be the same as a SYT, except that the entries can be any distinct integers, not necessarily $1, ..., n$. Let P_{ij} denote the entry in i-th row and j-th column of P. The RSK operation *row inserts* a positive integer k into P as follows: let r be the least integer such that $P_{1r} > k$. If no such r exists, then place k at the end of the first row of P. If such an r does exist, then replace P_{1r} by k and insert the element r into the second row according to the rule just described. Continue until an element is inserted at the end of a row. The resulting array is $P \leftarrow k$, where $\leftarrow$ is the row insertion process. Now, let $\pi = a_1a_2...a_n \in S_n$. We show that π produces a pair of SYT's by the end of the algorithm. We inductively construct a sequence $(P_0, Q_0), ..., (P_n, Q_n)$ of pairs of NYT of the same shape as follows: First, define $(P_0, Q_0) = (\emptyset, \emptyset)$. If (P_{i-1}, Q_{i-1}) has been defined, then set $P_i = P_{i-1} \leftarrow a_i$. Now, define Q_i to be the NYT obtained from Q_{i-1} by inserting i so that Q_i and P_i have the same shape (not row-inserting i into Q_{i-1} but simply placing i in a new place so Q_i and P_i have the same shape). We write $\pi \rightarrow (RSK)(P, Q)$.

Theorem 4.7. *The RSK algorithm defines a bijection between S_n and the set of all pairs (P, Q) of SYT of the same shape, where the shape λ is a partition of n.*

Proof. (Sketch) We want to define the inverse of RSK. Note that the position occupied by n in Q is the last position to be occupied in the insertion process. Suppose k in P occupies this position. It was bumped to this position by an element l in the row above k such that *l is currently the largest element of its row less than k.* Inductively proceed, we finally bump an element in the first row out, which is $a_n \in \pi$. Perform the same procedure for $n-1, n-2, ...$ to get the desired result.

Therefore, we have another bijective proof of the fact that $\sum_{\lambda \vdash n}(f^\lambda)^2 = n!$ because we can recover π from any pair of (P, Q). □

5. Representations of the Symmetric Group

We continue here our brief foray into the study of representations of symmetric groups, eventually arriving at the conclusion that $\sum_{\lambda \vdash n}(f^\lambda)^2 = |S_n|$ by understanding the irreducible representations (modules) of S_n.

Recall that for a finite group G, the number of its irreducible representations is equal to the number of its conjugacy classes and, when G=S_n, the symmetric group on n-letters, this is equal to the number of partitions of n. This happens because two permutations of S_n are in the same conjugacy class if and only if they have the same cycle type.

We begin by defining the permutation modules M^λ:

Definition 5.1. Suppose $\lambda = (\lambda_1, ..., \lambda_l) \vdash n$. The *Ferrers diagram* of λ is an array of n dots having l left-justified rows with row i containing λ_i dots for $1 \leq i \leq l$. A *Young tableau of shape* λ is an array formed by replacing the dots of the Ferrers diagram of λ with the numbers $1, 2, ..., n$ bijectively. Two λ-tableaux (tableaux is the plural form of tableau and a λ-tableau is a tableau associated with $\lambda \vdash n$) 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 $\{t\} = \{t_1 | t_1 \sim t\}$, i.e. a row equivalence class.

Regarding notations, we will write a λ-tableau where $\lambda = (3, 2)$ and $\lambda \vdash 5$ in the form

$$t = \begin{array}{ccc} 4 & 1 & 2 \\ 3 & 5 & \end{array}$$

while its associated tabloid will have the form

$$\begin{array}{ccc} \hline 4 & 1 & 2 \\ \hline 3 & 5 & \\ \cline{1-2} \end{array},$$

with long vertical bars above the numbers.

Now $\pi \in S_n$ acts on a tableau $t = (t_{i,j})$ of shape $\lambda \vdash n$ as follows:

$$\pi t = (\pi(t_{i,j})).$$

This induces an action on tabloids by letting $\pi\{t\} = \{\pi t\}$. This is well-defined because π of any two equivalent tableaux will still be equivalent as elements in any row for those two tableaux will still be the same. Therefore, we have a S_n-module in a natural way:

Definition 5.2. Suppose $\lambda \vdash n$. Let

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

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

Notice that M^λ is cyclic, meaning that any λ-tabloid $\in M^\lambda$ can be sent to another arbitrary λ-tabloid $\in M^\lambda$ by the action of a suitable $\pi \in S_n$.

Now, we define a partial order on the partitions and present a dominance lemma for those partitions:

Definition 5.3. Suppose $\lambda = (\lambda_1, ..., \lambda_l) \vdash n$ and $\mu = (\mu_1, ..., \mu_m) \vdash n$ are two partitions of n. Then λ *dominates* μ, written as $\lambda \trianglerighteq \mu$, if, for all $i \geq 1$

$$\lambda_1 + ... + \lambda_i \geq \mu_1 + ... + \mu_i.$$

If $i > l$ (or m), then we take λ_i(or μ_i) to be 0.

Lemma 5.4. ***Dominance Lemma for Partitions*** *Let t^λ and s^μ be two tableaux of shape λ and μ, respectively. If, for each index i, the elements of row i of s^μ are all in different columns in t^λ, then $\lambda \trianglerighteq \mu$.*

Proof. By hypothesis, t^λ contains elements of s^μ (although in a different order) and we can sort the entries in each column of t^λ so that elements of the first i rows of s^μ all occur in the first i rows of t^λ. Thus $\lambda_1 + ... + \lambda_i$ is equal to number of elements in the first i rows of t^λ which is greater or equal to the number of elements in the first i rows of s^μ which is equal to $\mu_1 + ... + \mu_i$ for all $i \geq 1$. □

We now construct all the irreducible modules of S_n, which are called Specht modules, S^λ.

Definition 5.5. Suppose that a tableau t has rows $R_1, ..., R_l$ and columns $C_1, ..., C_k$. Then

$$R_t = S_{R_1} \times S_{R_2} \times ... \times S_{R_l}$$

and

$$C_t = S_{C_1} \times S_{C_2} \times ... \times S_{C_k}$$

are the *row-stabilizers* and *column-stabilizers* of t, respectively.

Notice that $\{t\} = R_t t$ for any tableau t since two row-equivalent tableaux are in the same equivalence class.

We give two other definitions, the first of which takes group algebra sums over any subset of S_n:

Definition 5.6. Given Q, a subset of S_n, define $Q^- := \sum_{\pi \in Q} \operatorname{sgn}(\pi)\pi$.

Definition 5.7. If t is a tableau, define $\tau_t := C_t^-$ and form the associated *polytabloid*

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

In other words, a polytabloid is a linear combination of tabloids (signs included) formed by permuting the column entries of the original tabloid$\{t\}$ among themselves.

Example 5.8. For $t = \begin{matrix} 4 & 1 & 2 \\ 3 & 5 & \end{matrix}$, $\tau_t = (\epsilon - (3,4))(\epsilon - (1,5))$ and

$$e_t = \begin{array}{ccc} \hline 4 & 1 & 2 \\ \hline 3 & 5 & \\ \cline{1-2} \end{array} - \begin{array}{ccc} \hline 3 & 1 & 2 \\ \hline 4 & 5 & \\ \cline{1-2} \end{array} - \begin{array}{ccc} \hline 4 & 5 & 2 \\ \hline 3 & 1 & \\ \cline{1-2} \end{array} + \begin{array}{ccc} \hline 3 & 5 & 2 \\ \hline 4 & 1 & \\ \cline{1-2} \end{array}$$

Note: a tableau is represented in the typical form

$$t = \begin{matrix} 4 & 1 & 2 \\ 3 & 5 & \end{matrix}$$

whereas its equivalence class, distinguished by bolding it, is represented in the typical form

$$\begin{array}{ccc} \hline 4 & 1 & 2 \\ \hline 3 & 5 & \\ \cline{1-2} \end{array}.$$

We also note the following facts: $R_{\pi t} = \pi R_t \pi^{-1}$, $C_{\pi t} = \pi C_t \pi^{-1}$, $\tau_{\pi t} = \pi \tau_t \pi^{-1}$, and $e_{\pi t} = \pi e_t$. The first three relations follow because we notice that for instance $\alpha \in R_{\pi t} \leftrightarrow \alpha\{\pi t\} = \{\pi t\} \leftrightarrow \pi^{-1}\alpha\pi\{t\} = \{t\} \leftrightarrow \pi^{-1}\alpha\pi \in R_t \leftrightarrow \alpha \in \pi R_t \pi^{-1}$ while $e_{\pi t} = \tau_{\pi t}\{\pi t\} = \pi\tau_t\pi^{-1}\{\pi t\} = \pi\tau_t\{t\} = \pi e_t$. In other words, we obtain a row (or column stabilizer) of πt by sending πt back to t, applying a stabilizer of t to t, and applying π again to arrive at a stabilizer for πt.

Definition 5.9. For any partition λ, the corresponding *Specht module*, S^λ, is the submodule of M^λ spanned by the polytabloids e_t, where t is of shape λ.

Notice also that S^λ is a cyclic module generated by any given *polytabloid* while M^λ is generated by any *tabloid*.

Our main focus lies in proving that S^λ's form a full set of irreducible modules for S_n and understanding the basis elements for S^λ so that we can apply the well-known formula $\sum_i (dimV^{(i)})^2 = |G|$. Here, $V^{(i)}$ will be S^λ for a particular $\lambda \vdash n$ and $G = S_n$.

We begin by introducing a sequence of important lemmas. We will also need the unique inner product on M^λ for which

$$\langle \mathbf{t}, \mathbf{s} \rangle = \delta_{\{t\},\{s\}}.$$

Theorem 5.10. *(Sign Lemma): Let $H \leq S_n$ be a subgroup.*

1. If $\pi \in H$, then $\pi H^- = H^- \pi = (\operatorname{sgn}\pi) H^-$

2. For any $\boldsymbol{u}, \boldsymbol{v} \in M^\lambda$, $\langle H^- \boldsymbol{u}, \boldsymbol{v}\rangle = \langle \boldsymbol{u}, H^- \boldsymbol{v}\rangle$ where $\langle \{\mathbf{t}\}, \{\mathbf{s}\}\rangle = \delta_{\{\mathbf{t}\},\{s\}}$ is the unique inner product on M^λ.

3. If the transposition (b,c) belongs to H, then we can factor $H^- = k(\epsilon - (b,c))$ where $k \in \mathbb{C}[S_n]$.

4. If t is a tableau with b, c in the same row of t and $(b,c) \in H$, then $H^-\{\boldsymbol{t}\} = \boldsymbol{0}$.

Proof. 1. Given $\pi \in H$,

$$\pi H^- = \pi \sum_{\tau \in H} \operatorname{sgn}(\tau)\tau = \sum_{\tau \in H} \operatorname{sgn}(\tau)\pi\tau = \sum_{\sigma \in H} \operatorname{sgn}(\pi^{-1}\sigma)\sigma = \operatorname{sgn}(\pi^{-1}) \sum_{\sigma \in H} \operatorname{sgn}(\sigma)\sigma = \operatorname{sgn}(\pi) H^-$$

where we use the facts that $\sigma = \pi\tau \in H$ and that $\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi^{-1})$.
2. Since the inner product is S_n-invariant ($\pi(\{\mathbf{t}\}) = \pi(\{\mathbf{s}\})$ if and only if $\{\mathbf{t}\} = \{\mathbf{s}\}$), we see that

$$\langle H^- \mathbf{u}, \mathbf{v}\rangle = \sum_{\pi \in H} \langle (\operatorname{sgn}\pi)\pi \mathbf{u}, \mathbf{v}\rangle = \sum_{\pi \in H} \langle (\mathbf{u}, \operatorname{sgn}\pi)\pi^{-1}\mathbf{v}\rangle = \langle \mathbf{u}, H^- \mathbf{v}\rangle \tag{5.11}$$

since $\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi^-)$.
3. Consider the subgroup $K = \{\epsilon, (b,c)\}$ of H. Then we can find a transversal, which is the set consisting of representatives from each coset of K, and write $H = \uplus_i k_i K$. Then $H^- = (\sum_i k_i^-)(\epsilon - (b,c))$.
4. By assumption, $(b,c)(\{\mathbf{t}\}) = \{\mathbf{t}\}$. Thus

$$H^-(\{\mathbf{t}\}) = k(\epsilon - (b,c))(\{\mathbf{t}\}) = k(\{\mathbf{t}\} - \{\mathbf{t}\}) = \mathbf{0}.$$

We have two immediate corollaries: ∎

Corollary 5.12. *Let $t = t^\lambda$ and $s = s^\mu$ be λ and μ tableaux, where λ and μ partition n. If $\tau_t\{\boldsymbol{s}\} \neq \boldsymbol{0}$, then $\lambda \trianglerighteq \mu$. If $\lambda = \mu$, then $\tau_t\{\boldsymbol{s}\} = \pm e_t$.*

Proof. First, since τ_t is in essence a linear combination of elements in S_n, including the sgn of each element, we can apply τ_t to $\{\mathbf{s}\}$, even if s^μ and t^λ correspond to different partitions. Suppose b and c are two elements in the same row of s^μ. Then they cannot be in the same column of t^λ since otherwise $\tau_t = k(\epsilon - (b,c))$ and $\tau_t(\{\mathbf{s}\}) = \mathbf{0}$ by the preceding lemma. So by the dominance lemma for partitions, $\lambda \trianglerighteq \mu$.

If $\lambda = \mu$, then we must have $\{s\} = \pi\{t\}$ for some $\pi \in C_t$. Hence, by the preceding lemma,

$$\tau_t\{\mathbf{s}\} = \tau_t\pi\{\mathbf{t}\} = (\operatorname{sgn}\pi)\tau_t\{\mathbf{t}\} = \pm e_t$$

□

Corollary 5.13. *If $\boldsymbol{u} \in M^\mu$ and shape of t is μ, then $\tau_t\boldsymbol{u}$ is a multiple of e_t.*

Proof. We have $\mathbf{u} = \sum_i c_i\{s_i\}$, where the s_i are μ-tableaux. By the previous corollary, $\tau_t\mathbf{u} = \sum_i c_i\tau_t(\{s_i\}) = \sum_i \pm c_i e_t$. □

Theorem 5.14. *(Submodule Theorem) Let U be a submodule of M^μ, then*

$$U \subset S^\mu \text{ or } U \subset S^{\mu\perp}$$

Proof. Our proof mainly relies on the cyclicity of S^μ and the second part of the sign lemma.

Consider $\mathbf{u} \in U$ and a μ-tableau t. We know that $\tau_t\mathbf{u} = ce_t$ for some field element c by the preceding corollary. We now consider two cases where c can arise.

First, suppose there exists $\mathbf{u}$ and t such that $c \neq 0$. Then, since U is a submodule containing $\mathbf{u}$, we have $e_t = \frac{1}{c}\tau_t\mathbf{u} \in U$. Since S^μ is cyclic, we have $S^\mu \subset U$. Second, if we always have $\tau_t\mathbf{u} = \mathbf{0}$, then consider any $\mathbf{u} \in U$. Given an arbitrary μ-tableau t, we see that $\langle \mathbf{u}, e_t\rangle = \langle \mathbf{u}, \tau_t\{\mathbf{t}\}\rangle = \langle \tau_t\mathbf{u}, \{\mathbf{t}\}\rangle = \langle \mathbf{0}, \{\mathbf{t}\}\rangle = \mathbf{0}$.

Now, when we work over the complex field, we see that $S^\mu \cap S^{\mu\perp} = \{0\}$ so that given any non-trivial submodule $U \in S^\mu$, $U = S^\mu$ by double inclusion. Hence, S^μ is irreducible. □

Theorem 5.15. *If $\theta \in$ Hom(S^λ, M^μ) is nonzero then λ dominates μ. If $\lambda = \mu$, then θ is multiplication by a scalar*

Proof. Since $\theta \neq 0$, there exists some basis vector e_t such that $\theta(e_t) \neq \mathbf{0}$. Because we have $M^\lambda = S^\lambda \bigoplus S^{\lambda\perp}$ over the complex field, we can extent θ to an element of $\mathrm{Hom}(M^\lambda, M^\mu)$ by letting $\theta(S^{\lambda\perp})$=$\mathbf{0}$.

$$\mathbf{0} \neq \theta(e_t) = \theta(k_t\{\mathbf{t}\}) = k_t\theta(\{\mathbf{t}\}) = k_t(\sum_i c_i\{\mathbf{s}\}_i).$$

Therefore, where $\{\mathbf{s}\}_i$ are basis elements in M^μ. By **Corollary 5.12**, we have that λ dominates μ.

If $\lambda = \mu$, by **Corollary 5.13** $\theta(e_t) = ce_t$ for some constant c. Given any permutation π,

$$\theta(e_{\pi t}) = \theta(\pi e_t) = \pi\theta(e_t) = \pi(ce_t) = ce_{\pi t}.$$

□

Theorem 5.16. *S^λ for λ partitioning n form a complete list of irreducible S_n-modules over the complex field.*

Proof. Because the number of irreducible representations of S_n is equal to the number of partitions of n, we have the right number of modules for a full set. Thus, we want to know that Specht modules associated with different partitions are inequivalent. Instead, we prove the fact that $S^\lambda \cong S^\mu$ forces $\lambda = \mu$. If this happens, then there exists $0 \neq \theta \in Hom(S^\lambda, M^\mu)$ and $\lambda \trianglerighteq \mu$ by the previous theorem. Similarly, there exists $0 \neq \beta \in Hom(S^\mu, M^\lambda)$ and $\mu \trianglerighteq \lambda$. Hence, $\lambda = \mu$. $\square$

Finally, we want to find a basis for S^λ so that we can understand their degrees. More specifically, we want to prove that $\{e_t : t \text{ is a standard } \lambda\text{-tableau}\}$ forms a basis for S^λ.

We first prove that the set of e_t's are independent and then show that they span S^λ.

We begin with a definition that allows us to state the dominance lemma for tabloids:

Definition 5.17. A *composition* of n is an ordered sequence of nonnegative integers $\lambda = (\lambda_1, ..., \lambda_l)$ such that $\sum_i \lambda_i = n$ and λ_i are parts of the composition. Given a tabloid $\{t\}$ with shape $\lambda \vdash n$, for each index $1 \leq i \leq n$, let $\{t^i\}$, in the sense introduced earlier, be the tabloid formed by all elements $\leq i$ in $\{t\}$ and λ^i be the composition which is the shape of $\{t^i\}$.

Example 5.18. For

$$\{t\} = \begin{array}{cc} \hline 2 & 4 \\ \hline 1 & 3 \\ \hline \end{array},$$

$$\{t^1\} = \begin{array}{c} \hline \emptyset \\ \hline 1 \\ \hline \end{array}, \{t^2\} = \begin{array}{c} \hline 2 \\ \hline 1 \\ \hline \end{array}, \{t^3\} = \begin{array}{cc} \hline 2 & \\ \hline 1 & 3 \\ \hline \end{array}, \{t^4\} = \begin{array}{cc} \hline 2 & 4 \\ \hline 1 & 3 \\ \hline \end{array}$$

$$\lambda^1 = (0,1), \lambda^2 = (1,1), \lambda^3 = (1,2), \lambda^4 = (2,2)$$

Definition 5.19. When $\{s\}$ and $\{t\}$ are tabloids with composition series s^i and t^i, $\{s\}$ dominates $\{t\}$ if $s^i \trianglerighteq t^i$ for all i.

Lemma 5.20. *(Dominance Lemma for Tabloids): Given a tabloid $\{t\}$, if $k < l$ and k appears in a lower row than l in $\{t\}$, then $(k,l)\{t\}$ dominates $\{t\}$.*

Proof. When $i < k$ and $i > l$, the composition series for $(k,l)\{t\}$ and $\{t\}$ are the same. When $k \leq i < l$, suppose k appears in r-th row and l in k-th row. Then the l-th part of the composition for $(k,l)\{t\}$ increases by 1 and the r-th part of it decrease by 1. Since we have assumed that l is a higher row than r, the result follows. $\square$

In fact, the point of this dominance ordering is that it will give us an ordering that we can induct upon, as the corollary below explains.

Corollary 5.21. *If t is standard and $\{s\}$ appears in e_t, then $\{t\}$ dominates $\{s\}$.*

Proof. let $s = \pi t$, where $\pi \in C_t$. We induct on the number n of column inversions of s, where n is the number of pairs of k, l in the same column with $k < l$ and k in the lower row than l. When n=1, the result follows directly from the previous lemma. Assume the result for $n-1$ is true and consider when the number of inversions equals to n. Given any such pair a, b, $(a,b)\{s\}$ dominates $\{s\}$ by the previous lemma. Since $(a,b)\{s\}$ has fewer inversions than $\{s\}$ does, by induction $\{t\}$ dominates $\{s\}$. $\square$

Before we prove the independence of the $\{e_t\}$ for t being a SYT, we introduce a last lemma, in which $\{t_i\} \in v_i$ means that when we write v_i in terms of the tabloids, the coefficient of $\{t_i\}$ is nonzero:

Lemma 5.22. *Let $v_1, ..., v_m$ be elements of M^μ. Suppose, for each v_i, we can choose a maximum tabloid $\{t_i\} \in v_i$ with respect to the dominance ordering and that all $\{t_i\}$'s are distinct, then $v_1, ..., v_m$ are independent.*

Proof. Suppose $\sum_i c_i v_i = 0$, we want to show $c_i = 0$ for all i. We again proceed by induction on m, the number of v_i. If $m = 1$, then $c_m = c_1 = 0$. If $m = n-1$ is true, consider when $m = n$. Choose a *maximal* element among the *maximum* $\{t_i\}$'s, call it $\{t_1\}$ (by permuting the indices). We claim that $\{t_1\} \in v_1$ and exists in no other element $v_2, ..., v_n$. Suppose $\{t_1\} \in v_i, i \neq 1$, then $\{t_i\}$ dominates $\{t_1\}$, contradicting the maximality of $\{t_1\}$. Thus, we know that $c_1 = 0$ because we cannot cancel $\{t_1\}$ in any other way. Hence, by induction, we have all other coefficients zero so the vectors are independent. $\square$

Proposition 5.23. *The set of elements $\{e_t : t$ is a standard young λ-tableau $\}$ is linearly independent.*

Proof. By **Corollary 5.21**, $\{t\}$ is the maximum element in e_t and by definition of being standard, different standard young tableaux represent different equivalence classes. Therefore, both assumptions in the previous proposition are met and we have the result. $\square$

Lastly, we want to prove the spanning property of $\{e_t\}$ by introducing the *Garnir elements.*

The steps are as follows: suppose t is an arbitrary tableau. We can assume that the columns of t are increasing. If not, we can permute its column entries with π in C_t and use the first sign lemma to show that the permuted polytabloid and the original one differ only by a ± 1 in front. We want to show that e_t lives in the span of standard polytabloids by using that given a *column tabloid* $[t]$ (follow the definition of a row tabloid except changing the word row to a column), every polytabloid e_s with tableau s and $[s]$ dominating $[t]$ (using dominance of compositions) stays in the span of the standard polytabloids. When t is not standard (meaning there exists a pair of row descents $t_{i,j}$ and $t_{i,j+1}$ such that $t_{i,j} > t_{i,j+1}$), write the polytabloid e_t as a linear combination of polytabloids e_s by applying the Garnir element to e_t, whose definition is below.

Definition 5.24. Let t be a tableau and let A and B be subsets of the j-th and $(j+1)$-th column of t, respectively. The *Garnir element associated with t* (and A, B) is $g_{A,B} = \sum_\pi \mathrm{sgn}(\pi)\pi$, where the π have been chosen so that the elements of $A \cup B$ are increasing down the columns of πt.

In practice, we always take A (respectively B) to be all elements below $t_{i,j}$ (respectively above $t_{i,j+1}$). As the example below:

Example 5.25.

For $t = \begin{matrix} 1 & 2 & 3 \\ 5 & 4 & \\ 6 & & \end{matrix}$, $A = \{5, 6\}, B = \{2, 4\}$,

$$g_{A,B} = \epsilon - (4,5) + (2,4,5) + (4,6,5) - (2,4,6,5) + (2,5)(4,6).$$

When we choose $\operatorname{sgn}(\pi)\pi = -(4,5)$, we get

$$\pi t = \begin{matrix} 1 & 2 & 3 \\ 4 & 5 & \\ 6 & & \end{matrix}.$$

Notice that when we permute the elements of $A \cup B$ according to the rule described above, we have eliminated the row descent $5 < 4$ at the position $t_{2,1} = 5$ and $t_{2,2} = 4$. Nonetheless, we may create other tableaux such as

$$t' = \begin{matrix} 1 & 4 & 3 \\ 2 & 5 & \\ 6 & & \end{matrix}$$

with row descents somewhere else. However, this will not pose a problem since $[t']$ dominates $[t]$ according to the composition rule.

We argue that $g_{A,B}e_t = 0$ under the condition below:

Proposition 5.26. *Let t, A, and B be as in the definition of a Garnir element. If $|A \cup B|$ is greater than the number of elements in column j of t, where j is the column that contains A, then $g_{A,B}e_t = 0$.*

Proof. Consider any $\mu \in C_t$. By the hypothesis, there exist $a, b \in A \cup B$ such that a, b are in the same row of μt. Then $(a, b) \in S_{A\cup B}$ and $S^-_{A\cup B}\{\mu t\} = 0$ by part 4 of the sign lemma. Since this is true for every element in C_t and hence in τ_t, we have that $S^-_{A\cup B}e_t = 0$.

Since $S_A \times S_B$, where S_A (or S_B) is the symmetric group on A, is a subgroup of $S_{A\cup B}$, $S^-_{A\cup B} = g_{A,B}(S_A \times S_B)^-$ where $g_{A,B}$ constitutes a full set of transversals for $S_A \times S_B$. Hence, $g_{A,B}(S_A \times S_B)^- e_t = 0$. Now, since $S_A \times S_B \subset C_t$, by sign lemma we see that $(S_A \times S_B)^- e_t = |S_A \times S_B| e_t$ and dividing the equation $g_{A,B}(S_A \times S_B)^- e_t = 0$ by cardinality yields the result. $\square$

Theorem 5.27. *The set $\{e_t : t$ is a standard young $\lambda-$tableau$\}$ spans S^λ.*

Proof. By induction, we may assume that every tableau s with $[s]$ dominating $[t]$ is in the span of e_t. We can also assume that any arbitrary tableau t has increasing columns. Consider the base case, the maximum column tabloid $[t_0]$, where we number the entries of each column consecutively from top to bottom, starting from the leftmost column and working right. Since $[t_0]$ is standard, we are done for this equivalence class. We also assume that

If t is not standard, then it has at least one pair of row descents. Applying $g_{A,B}$ to e_t, we see that $g_{A,B}e_t = 0$ by the previous proposition. Rearranging the equality yields the result that $e_t = -\sum_{\pi \neq \epsilon} (\operatorname{sgn} \pi)e_{\pi t}$. Notice that $[\pi t]$ dominates $[t]$ by a column analogue of the dominance lemma for (row) tabloids. Hence by induction, we see that every summand in the right hand side of the equation is in the span of standard polytabloids; so is e_t. The result follows. $\square$

We finally arrive at the result:

Theorem 5.28. *1.* $\{e_t : t$ *is a standard young* $\lambda-$*tableau*$\}$ *is a basis for* S^λ
2. dim $S^\lambda = f^\lambda$*=the number of standard young tableaux*
3. $\sum_{\lambda \vdash n} (f^\lambda)^2 = |S_n|$.

We summarize in the end what we have done in this section:

(1): Given a partition $\lambda \vdash n$, we associated a permutation S_n-module, denoted as M^λ with that partition.

(2): We defined an ordering on partitions and gave a dominance lemma on those partitions

(3): We restricted our view to the Specht modules S^λ, which are irreducible submodules of M^λ and form a complete list of irreducible modules for S_n.

(4): We proved that the polytabloids formed from the Standard Young Tabloids constitute a basis for S^λ, achieving our result.

Acknowledgments. I want to thank my mentor, Jonathan Dewitt, for his dedication and helpfulness throughout the program. This paper would not have been possible without his enthusiastic efforts and strategic formulation. I am also grateful for Prof. Peter May and the Math Department for organizing such an amazing program on an annual basis throughout which I was introduced to the grandeur of mathematics.

References

[1] Richard P. Stanley. Algebraic Combinatorics. Springer UTM, 2013.
[2] Bruce E. Sagan. The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions. Springer GTM, 2001.