Investigations into the Symmetric group

Isaac Kochman

May 2, 2018

Contents

1 Introduction

The symmetric group is a central object in finite group theory. Its study is illuminative in this and other branches of mathematics. This paper will explore and dissect the symmetric group. We will start with a brief introduction to group theory to provide essential background knowledge and define the object in question. Then we will spend some considerable time developing ideas in representation theory, which is the main focus of this study. We will develop sufficient background in the field to understand the concepts we are working with. This will include demonstrating some interesting results along the way to get comfortable working with these ideas. Then we will construct all the irreducible representations of the symmetric group. In a sense these are the smallest objects from which all other representations are built, and understanding how they work, especially for the symmetric group, is a first step towards many other results in both group theory and representation theory, which are beyond the scope of this paper. This study of the symmetric group will lay a solid foundation for further study in both theories and present some interesting results in their own right.

The purpose of this paper is to summarize an extended study in the symmetric group. This study covered both its composition and applications. It is not meant to be exhaustive in any way, but it pulls together essential information from multiple areas of study presenting enough information to grasp some high level concepts and interesting results. It will approach this task from different levels including a high level view of the fields, outlines of how development and particular topics are approached, and more detailed examples of actual results achieved during this study. We start with the basics of group theory.

2 Group Theory

In this section we will lay down some background, in the related subject of group theory, necessary to understand concepts discussed later in this paper.

2.1 Definitions

To start off with, group theory is the study of a class of mathematical objects known as groups. In order to study these objects we need to have a clear conception of what they are.

Definition 2.1.1 *A **group** is an ordered pair $(G, \bullet)$ where G is a set and $\bullet$ is a binary operation on G satisfying the following axioms*

1. *$(a \bullet b) \bullet c = a \bullet (b \bullet c)$ for all $a, b, c \in G$. I.e., $\bullet$ is **associative**.*
2. *There exists an element $\epsilon \in G$, called the **identity**, such that for all $a \in G$ we have $a \bullet \epsilon = \epsilon \bullet a = a$.*
3. *For each $a \in G$ there is an element $a^{-1} \in G$ such that $a \bullet a^{-1} = a^{-1} \bullet a = \epsilon$. In general we will refer to the group $(G, \bullet)$ as G and assume the operation is understood. A generic group operation will often be referred to as multiplication when none is specified, but it need not refer to the multiplication we are used to.*

A **binary operation** is a function $\bullet : G \times G \to G$. In other words it maps an ordered pair of elements from a set to another element from the set. The particular object of interest in this paper is as follows.

Definition 2.1.2 *A **permutation** of $1, \cdots, n$ is a bijective function $\sigma : \{1, \cdots, n\} \to \{1, \cdots, n\}$, we can also write the function values from 1 to n as a sequence, and that is considered a permutation. **The symmetric group** on n elements, or S_n is the set of all permutations of the numbers $1, \cdots, n$ with the operation of composition. The precise group notation label is $(S_n, \circ)$.*

Example 2.1.3 *The identity in S_4 is the identity function; $\sigma(1) = 1$, $\sigma(2) = 2$, $\sigma(3) = 3$, $\sigma(4) = 4$. If we map 2 to 3 and 3 to 2 we get the sequence $1, 3, 2, 4$, which is an element of S_4. If*

we have two elements, σ and π, defined by

$$\sigma : \begin{array}{l} 1 \mapsto 3 \\ 2 \mapsto 4 \\ 3 \mapsto 1 \\ 4 \mapsto 2 \end{array} \quad \textit{and} \quad \pi : \begin{array}{l} 1 \mapsto 2 \\ 2 \mapsto 4 \\ 3 \mapsto 3 \\ 4 \mapsto 1 \end{array}$$

they can be written as sequences $3, 4, 1, 2$ and $2, 4, 3, 1$. Then composing the permutations is just like composing the functions and we get the following:

$$\sigma \circ \pi = 3, 4, 1, 2 \circ 2, 4, 3, 1 = 4, 2, 1, 3.$$

This notation will quickly become cumbersome so we simplify it with cycle notation. A cycle such as, $(1, 2, 3)$. This cycle is equivalent to

$$\sigma : \begin{array}{l} 1 \mapsto 2 \\ 2 \mapsto 3 \\ 3 \mapsto 1 \\ 4 \mapsto 4 \end{array}.$$

In S_4 this refers to the same permutation as $2, 3, 1, 4$. Four was not permuted so it is left out of the notation. A permutation can be constructed from multiple non trivial cycles such as, $(1, 2)(3, 4) = 2, 1, 4, 3$. This notation is easier to work with and also leads to some useful results as we will see later on.

Now that we have the object of interest defined it is natural to be interested in the associated sub objects.

Definition 2.1.4 *A subset H of a group G is called a **subgroup** denoted $H \leq G$ if it satisfies the group axioms with the operation on G restricted to H.*

Now there are some interesting properties of subgroups which will be of interest. And a particular type of subgroup which will be of use later on. But first we need to define some more operations that can be done within a group.

Definition 2.1.5 *Let G be a group with elements $g, h \in G$. The **conjugation** of h by g is ghg^{-1}. The **conjugacy class** of an element $g \in G$ is $\{h \in G \mid h = xgx^{-1} \text{ for some } x \in G\}$. If g and h are in the same conjugacy class we say they are **conjugates**.*

Now the type of subgroup mentioned above is as follows.

Definition 2.1.6 *A subgroup $H \leq G$ is **normal** if for all $g \in G$ we have $gHg^{-1} \stackrel{def}{=} \{ghg^{-1} \mid h \in H\} = H$. If H is normal we also say H is closed under conjugation by elements of G and write $H \trianglelefteq G$.*

We are continuing to build more complex objects from those we have already laid out to set up the necessary understanding in group theory. The following is what happens if we multiply a subgroup by an element of the parent group only on one side.

Definition 2.1.7 *Given a subgroup $H \leq G$ and an element $g \in G$ we define the **left coset** $gH \stackrel{def}{=} \{gh \mid h \in H\}$ and we denote by G/H the set of all left cosets of H in G.*

Note that a coset can be represented by any of its elements, i.e. if $h \in gH$ then $hH = gH$, and either is an appropriate name for that particular coset. This will be an issue we have to deal with when ensuring that certain functions are well defined on the set of cosets. We can also get at these same objects in another way.

Definition 2.1.8 *Let $H \leq G$ then the set $\{t_1, \cdots, t_l\}$ is a **transversal** for H in G if each left coset is represented exactly once in the list $t_1H, \cdots, t_lH$.*

We have two more important definitions, and then we can start looking at some of their consequences.

Definition 2.1.9 *Given two groups $(G, \bullet)$ and $(H, \star)$ a function $f : G \to H$ is a* ***homomorphism*** *if for all $g, h \in G$ we have $f(g) \star f(h) = f(g \bullet h)$. Again much of the more formalized notation will be omitted later on. We often write $f(g)f(h) = f(gh)$.*

Definition 2.1.10 *An* ***isomorphism*** *is a homomorphism that is also a bijection (i.e. one-to-one mapping). If there exists an isomorphism between two groups G and H we say G and H are* ***isomorphic*** *and write $G \cong H$.*

From these definitions we can start building some basic results. The following are covered in any intro to group theory course, and most of the proofs are omitted here.

2.2 Basic Results

First of all we want to define another object related to subgroups, but its existence is not obvious.

Proposition 2.2.1 *Given a group G with subgroup H such that $H \trianglelefteq G$, the set of left cosets G/H is a group where multiplication is defined by $gHhH = ghH$. We call this the* ***quotient group*** *of G by H.*

Next we have a result for identifying elements of the quotient group, which also applies to left cosets in general.

Proposition 2.2.2 *If $H \leq G$ and $g, h \in G$ then, $gH = hH$ if and only if $h^{-1}g \in H$.*

All of our discussion from here on will be based on finite groups which have the following nice property:

Proposition 2.2.3 *If G is a finite group, then for each $g \in G$ there exists some $n \in \mathbb{N}$ such that $g^n = \epsilon$, the identity of the group.*

All of these general properties of groups will be useful, but we will also make extensive use of the following more specific results. Furthermore, we note at this point that the standard field of scalars for matrices and vector spaces, unless otherwise stated, is assumed to be the complex numbers $\mathbb{C}$.

Proposition 2.2.4 *The set of all invertible $d \times d$ matrices is a group with the operation of matrix multiplication. This is referred to as the* ***general linear group*** *of dimension d and is written $GL_d(\mathbb{C})$.*

Proof We check the three properties required to be a group: **1.** Matrix multiplication is associative. **2.** The identity matrix I which has 1's on the diagonal and 0's everywhere else satisfies the identity property for groups. **3.** Every invertible matrix has an inverse which is also an invertible matrix so it is in the group. □

As a point of clarification here, the **diagonal** of a matrix refers to the entries main diagonal which runs from top left to bottom right. Now to finish off this background section we have two results specifically about the symmetric group. But the first requires a little more build up.

Definition 2.2.5 *A* ***partition*** *λ of a number n of length l is an ordered l-tuple, $\lambda = (\lambda_1, \cdots, \lambda_l)$ such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l$ and $\sum_{i=1}^{l} \lambda_i = n$. If λ is a partition of n, we write $\lambda \vdash n$.*

Given an element π of S_n there is a standard way of writing it in cycle notation such that all cycles are disjoint (i.e. they have no numbers in common) and they appear in decreasing order of cycle size. In this way we can associate π with a partition of n which we refer to as its **cycle type**. It is a convenient result from abstract algebra that conjugation preserves cycle type. This observation leads nicely to the next proposition.

Proposition 2.2.6 *The number of conjugacy classes of S_n is equal to the number of partitions of n.*

Finally we mention the biggest reason this focused study is of such great interest, and has such far reaching consequences. This also explains some of the significance of the symmetric group in the overall field of group theory.

Proposition 2.2.7 *Every finite group is isomorphic to a subgroup of S_n for some $n \in \mathbb{N}$.*

With that background set up we can move on to the main focus of this paper. It involves another subject which will allow us to further understand the symmetric group. And continued study on this topic would see further applications of the following back in group theory.

3 Representation Theory

Representation theory is an area of mathematics focused on studying abstract algebraic objects such as general groups, by mapping them to simpler, better understood objects like matrices and vector spaces. In this section we will develop a solid background in the theory, look at some specific problems that can be solved within it, and then look at how we can break the symmetric group up into its irreducible representations. In this section we follow the treatment given in [1].

3.1 Building the Theory

We start by defining matrix and then module representations.

Definition 3.1.1 *A **matrix representation** of a group G is a group homomorphism,*

$$X : G \to GL_d,$$

which assigns a matrix $X(g) \in GL_d$ to each element $g \in G$ such that

1. *$X(\epsilon) = I$ the identity matrix, and*
2. *$X(gh) = X(g)X(h)$ for all $g, h \in G$.*

The parameter d is called the degree, or dimension, of the representation and is denoted by $deg(X)$.

For some specific examples of representations and some results about them see Section 3.2.

Since matrices correspond to linear transformations we get another way to think of representations. If V is any vector space, let $GL(V)$ stand for the set of all invertible linear transformations from V to itself, called the *general linear group of V*. If $\dim V = d$, then $GL(V)$ and GL_d are isomorphic as groups and we have the following:

Definition 3.1.2 *Let V be a vector space and G be a group. Then V is a **G-module** if there as a group homomorphism*

$$\rho : G \to GL(V)$$

Where the element $\rho(g)$ acts on V as a linear transformation which can be thought of as a multiplication, $\rho(g)\boldsymbol{v} = g\boldsymbol{v}$, of elements in V by elements in G such that

1. *$g\boldsymbol{v} \in V$,*
2. *$g(c\boldsymbol{v} + d\boldsymbol{w}) = c(g\boldsymbol{v}) + d(g\boldsymbol{v})$,*
3. *$(gh)\boldsymbol{v} = g(h\boldsymbol{v})$, and*
4. *$\epsilon\boldsymbol{v} = \boldsymbol{v}$*

for all $g, h \in G$, $\boldsymbol{v}, \boldsymbol{w} \in V$, and scalars $c, d \in \mathbb{C}$.

We note that because of the relationship between matrices and linear transformations it is quite easy to transition between these two types of representations, and as we will see they will both be useful in different situations.

Example 3.1.3 *Let G be a group with subgroup $H \leq G$. An example of a G-module is the (left) coset representation of G with respect to H. Let $t_1, \cdots, t_k$ be a transversal for H and let $\mathcal{H} = \{t_1 H, \cdots, t_k H\}$ be the set of left cosets. Then G acts on $\mathcal{H}$ by letting*

$$g(t_i H) = (g t_i) H$$

for all $g \in G$. The corresponding module

$$\mathbb{C}\mathcal{H} = \{c_1 t_1 \boldsymbol{H} + \cdots + c_k t_k \boldsymbol{H} \mid c_i \in \mathbb{C} \text{ for all } i\}$$

inherits the action. And it is easy to check that this space fulfills the definition of a G-module.

Now that we've established what a representation is we will look at reducibility. Some representations, as with many other mathematical objects, can be broken down into smaller parts. Doing this can help us study particular objects in the simplest terms possible. We start with the following definition.

Definition 3.1.4 *Let V be a G-module. A **submodule** of V is a subspace W that is closed under the action of G. Alternatively we say that W is a G-invariant subspace. In other words, W is a subset of V that is also a G-module. And we write $W \leq V$.*

A function f on a G-module is said to be G**-invariant** if for an element $g \in G$, $f(gx) = gf(x)$ for all x.

Definition 3.1.5 *We say that a G-module V is **reducible** if it contains a nontrivial submodule W.*

With this in mind we consider what it looks like for a representation to be built from smaller ones then we will see some results on the nature of reducibility.

Definition 3.1.6 *Let V be a vector space with subspaces U and W. Then V is the **direct sum** of U and W, written $V = U \oplus W$, if every $\boldsymbol{v} \in V$ can be written uniquely as a sum*

$$\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}, \text{ for } \boldsymbol{u} \in U, \text{ and } \boldsymbol{w} \in W.$$

*If V is a G-module and U, W are submodules, then we say that U and W are **complements** in V.*

*If X is a matrix, then X is the **direct sum** of matrices A and B, written $X = A \oplus B$, if X has the block diagonal form*

$$X = \left(\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right).$$

An **Inner Product**, $\langle \cdot, \cdot \rangle$ is a type of multiplication that takes two vectors as inputs and outputs a scalar. It has several nice properties for a function on a vector space. With it we can find submodule complements. Specifically, given a G-module V with submodule W we can construct the **orthogonal complement**:

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

It follows from this construction that whenever the inner product is G-invariant, $W^{\perp}$ is also a submodule and we have $V = W \oplus W^{\perp}$.

This result leads us nicely into **Maschke's Theorem**:

Theorem 3.1.7 *Let G be a finite group and let V be a nonzero G-module. Then*

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

where each $W^{(i)}$ is an irreducible G-submodule of V.

Proof We outline the proof here. It is done by induction on the dimension $d = \dim V$ of the module. If $d = 1$ the module is irreducible and we are done. If $d > 1$ and V happens to be irreducible then again we are done. Otherwise V is reducible so it has some nontrivial submodule W. If we can construct a G-invariant inner product, then we can find $V = W \oplus W^{\perp}$ and by the induction hypothesis W and $W^{\perp}$ can be broken down into irreducibles and we have our desired result.
To construct a G-invariant inner product, take a bases $\mathcal{B} = \{\mathbf{v}_1, \cdots, \mathbf{v}_d\}$ for V and consider the unique inner product that satisfies

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{i,j}$$

for elements of $\mathcal{B}$. This my not be G-invariant, but we can define

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

It is easy to verify that is still an inner product, and it is now G-invariant which completes the proof. □

The $\delta_{i,j}$ above is known as the **Kronecker delta**. This is a function defined for any two variables x and y such that:

$$\delta_{x,y} = \begin{cases} 1 : \text{if } x = y \\ 0 : \text{if } x \neq y \end{cases}$$

A similar result to the previous one applies to matrix representations, and there is a particular name for any representations which can be broken down as in the above theorem.

Definition 3.1.8 *A representation is* ***completely reducible*** *if it can be written as a direct sum of irreducibles.*

In light of this, Maschke's Theorem can be rewritten as: *Every representation of a finite group having positive dimension is completely reducible.*

Irreducible representations are useful when trying to get as much information as possible about an abstract group from a much simpler object. And, it turns out that much of the information contained in a representation can be simplified even more with characters.

Definition 3.1.9 *Let $X(g)$, $g \in G$, be a matrix representation. Then the* ***character*** *of X is*

$$\chi(g) = trX(g),$$

where tr denotes the trace of a matrix. If $x_{i,j}$ denotes the element in row i and column j of the $d \times d$ matrix $X(g)$. Then the ***trace*** *is $trX(g) = x_{1,1} + \cdots + x_{d,d}$.*

Note: We will use the expression $X \cong Y$, read: X and Y are **isomorphic** as representations. This means there is some other matrix T such that for each $g \in G$ we have $X(g) = TY(g)T^{-1}$.

If V is a G-module we can find a corresponding matrix representation X by choosing a basis and the trace of that will be the character of V. Any two matrix representations generated in this way from the same G-module will be isomorphic. And, since we know from linear algebra that trace is preserved under conjugation, the character is well defined for module representations.

The terminology we have developed to discuss representations will from now on be used interchangeably to refer to characters. Here are some basic properties of characters that we will not prove in this paper. However, most of them follow easily from the definitions.

Proposition 3.1.10 *Let X be a matrix representation of a group G of degree d with character χ.*

1. *$\chi(\epsilon) = d$.*
2. *If K is a conjugacy class of G, then*

$$g, h \in K \implies \chi(g) = \chi(h).$$

3. *If Y is a representation of G with character ψ, then*

$$X \cong Y \iff \chi(g) = \psi(g)$$

for all $g \in G$.

The reverse implication in part 3 is not obvious, but we will discuss it more later when we have the tools necessary to prove it.

Definition 3.1.11 *A **class function** is a function $f : G \to \mathbb{C}$ which is constant on conjugacy classes K of G.*

From the proceeding proposition we can see that the group character is a class function. This is important for the following proposition.

Proposition 3.1.12 *The number of irreducible characters of a group is equal to the number of conjugacy classes of that group.*

A proof of this fact requires considerable development beyond what we have room to present in this paper. Thus for our purposes we will take it as a given, but we recognize the results we derive from that fact represent a much greater development than can be seen in these pages. That being said, we can now construct a convenient way to display all the information we can get about a group from representations.

Definition 3.1.13 *Let G be a group. The **character table** of G is an array with rows indexed by the inequivalent irreducible characters of G and columns indexed by the conjugacy classes. The entry in row χ and column K is the value of that character on that conjugacy class denoted χ_K.*

Next we will define an inner product on characters which is useful in the proof that the number of irreducible characters is the same as the number of conjugacy classes, but will also be used in several other results which we will touch on here. There are a few possible equivalent definitions, but for our purposes the following will suffice.

Definition 3.1.14 *Let χ and ψ be any two complex-valued characters of a group G. Then the **inner product** of χ and ψ is:*

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

Now we will see our first theorem about characters, the proof is omitted. **Character Relations of the First Kind**:

Theorem 3.1.15 *Let χ and ψ be irreducible characters of a group G. Then*

$$\langle \chi, \psi \rangle = \delta_{\chi,\psi}.$$

This has several consequences which will come in handy later on.

Corollary 3.1.16 *Let X be a matrix representation of G with character χ. Suppose*

$$X \cong m_1 X^{(1)} \oplus \cdots \oplus m_k X^{(k)},$$

where the $X^{(i)}$ are pairwise inequivalent irreducibles with characters $\chi^{(i)}$.

1. $\chi = m_1\chi^{(1)} + \cdots + m_k\chi^{(k)}$.
2. $\langle \chi, \chi^{(j)} \rangle = m_j$ *for all* j.
3. $\langle \chi, \chi \rangle = m_1^2 + \cdots + m_k^2$.
4. *X is irreducible if and only if* $\langle \chi, \chi \rangle = 1$.

5. *Let Y be another matrix representation of G with character ψ. Then, as noted earlier,*

$$X \cong Y \text{ if and only if } \chi(g) = \psi(g)$$

for all $g \in G$.

Proof

1. Using the fact that the trace of a direct sum is the sum of the traces, we see that

$$\chi = trX = tr\bigoplus_{i=1}^{k} m_i X^{(i)} = \sum_{i=1}^{k} m_i \chi^{(i)}.$$

2. We have, by the previous theorem,

$$\langle \chi, \chi^{(j)} \rangle = \langle \sum_i m_i \chi^{(i)}, \chi^{(j)} \rangle = \sum_i m_i \langle \chi^{(i)}, \chi^{(j)} \rangle = m_j.$$

3. By another application of the same theorem:

$$\langle \chi, \chi \rangle = \langle \sum_i m_i \chi^{(i)}, \sum_j m_j \chi^{(j)} \rangle = \sum_{i,j} m_i m_j \langle \chi^{(i)}, \chi^{(j)} \rangle = \sum_i m_i^2$$

4. The forward implication is exactly 3.1.15. For the other direction suppose

$$\langle \chi, \chi \rangle = \sum_i m_i^2 = 1.$$

Then there must be exactly one index j such that $m_j = 1$ and the rest of the m_i must be zero, because the m_i are all non-negative integers. But then $X = X^{(j)}$, which is irreducible by assumption.

5. The forward implication which, we mentioned earlier, is obvious because trace is preserved under conjugation. For the reverse implication let $Y = \oplus_{i=1}^{k} n_i X^{(i)}$. There is no harm in assuming that the X and Y expansions contain the same irreducibles: Any found in one but not the other can be added in with multiplicity 0. Now $\chi = \psi$, so $\langle \chi, \chi^{(i)} \rangle = \langle \psi, \chi^{(i)} \rangle$ for all i. But then, by part 2, $m_i = n_i$ for all i. Thus the two direct sums are equivalent-i.e., $X \cong Y$.

□

Our final set of results in this section will be concerned with how we relate representations of groups and their subgroups. Given a group G with subgroup H we can get a representation of H from a representation of G and vice versa.

Definition 3.1.17 *Consider $H \leq G$ and a matrix representation X of G. The **restriction** of X to H, $X \downarrow_H^G$, is given by*

$$X \downarrow_H^G (h) = X(h)$$

for all $h \in H$. If X has character χ, then denote the character of $X \downarrow_H^G$ by $\chi \downarrow_H^G$.

It is trivial to verify that this satisfies the properties of a representation for H. The other direction is a little more involved.

Definition 3.1.18 *Consider $H \leq G$ and fix a transversal $t_1, \cdots, t_l$ for the left cosets of H. So $G = t_1 H \uplus \cdots \uplus t_l H$, where $\uplus$ denotes disjoint union. If Y is a representation of H, then the corresponding **induced representation** $Y \uparrow_H^G$ assigns to each $g \in G$ the block matrix*

$$Y \uparrow_H^G (g) = (Y(t_i^{-1} g t_j)) = \begin{pmatrix} Y(t_1^{-1} g t_1) & Y(t_1^{-1} g t_2) & \cdots & Y(t_1^{-1} g t_l) \\ Y(t_2^{-1} g t_1) & Y(t_2^{-1} g t_2) & \cdots & Y(t_2^{-1} g t_l) \\ \vdots & \vdots & \ddots & \vdots \\ Y(t_l^{-1} g t_1) & Y(t_l^{-1} g t_2) & \cdots & Y(t_l^{-1} g t_l) \end{pmatrix},$$

Where $Y(g)$ is the zero matrix if $g \notin H$. As with the previous definition we will use similar notation for the induced character.

It remains to show that induction does actually define a representation and that this representation is unique. We also note here that a restricted or induced representation may not be irreducible even if the original was.

Theorem 3.1.19 *Suppose $H \leq G$ has transversal $\{t_1, \cdots, t_l\}$ and let Y be a matrix representation of H. Then $X = Y \uparrow_H^G$ is a representation of G.*

Proof First we note that each row and column of $X(g)$ has exactly one non-zero block because for each t_i there is exactly one t_j such that $gt_i \in t_jH$ and hence $t_j^{-1}gt_i \in H$.
To prove that this is a representation we need to confirm that $X(\epsilon)$ is the identity matrix, but this is easy to see based on the definition of X and the previous observation.
Next we need to show that $X(g)X(h) = X(gh)$ for all $g, h \in G$. Considering the (i, j) block on both sides, following the definition of matrix multiplication it suffices to prove

$$\sum_k Y(t_i^{-1}gt_k)Y(t_k^{-1}ht_j) = Y(t_i^{-1}ght_j).$$

For ease of notation, let $a_k = t_i^{-1}gt_k, b_k = t_k^{-1}ht_j$, and $c = t_i^{-1}ght_j$. Note $a_kb_k = c$ for all k and that the sum can be rewritten as

$$\sum_k Y(a_k)Y(b_k) \stackrel{?}{=} Y(c).$$

Now we have two cases.
If $Y(c) = 0$, then $c \notin H$, and so either $a_k \notin H$ or $b_k \notin H$ for all k because H is a subgroup. Thus $Y(a_k)$ or $Y(b_k)$ is zero for each k, which forces the sum to be zero as well. If $Y(c) \neq 0$, then $c \in H$. Let m be the unique index such that $a_m \in H$. Thus $b_m = a_m^{-1}c \in H$, and so

$$\sum_k Y(a_k)Y(b_k) = Y(a_m)Y(b_m) = Y(a_mb_m) = Y(c),$$

completing the proof. □

Now we will see the first half of uniqueness, the rest will be shown as part of a result in the next section.

Proposition 3.1.20 *Consider $H \leq G$ and a matrix representation Y of H. Let $\{t_1, \cdots, t_l\}$ and $\{s_1, \cdots, s_l\}$ be two transversals for H giving rise to representations X and Z, respectively, for $Y \uparrow_H^G$. Then X and Z are equivalent.*

Proof Let χ, ψ and ϕ be the characters of X, Y, and Z, respectively. Then it suffices to show that $\chi = \phi$ by corollary 3.1.16. Now

$$\chi(g) = \sum_i trY(t_i^{-1}gt_i) = \sum_i \psi(t_i^{-1}gt_i),$$

where $\psi(g) = 0$ if $g \notin H$. Similarly,

$$\phi(g) = \sum_i \psi(s_i^{-1}gs_i).$$

Since the t_i and s_i are both transversals, we can permute subscripts if necessary to obtain $t_iH = s_iH$ for all i. Now $t_i = s_ih_i$, where $h_i \in H$ for all i, and so

$$t_i^{-1}gt_i = h_i^{-1}s_i^{-1}gs_ih_i.$$

Thus $t_i^{-1}gt_i \in H$ if and only if $s_i^{-1}gs_i \in H$, and when both lie in H, they are in the same conjugacy class. It follows that $\psi(t_i^{-1}gt_i) = \psi(s_i^{-1}gs_i)$, since ψ is constant on conjugacy classes of H and zero outside. Hence the sums for χ and ϕ are the same. □

And finally to wrap up this section we will prove a result about the relationship between the restricted and induced representations. But first we note that the character of the induced representation can be written as follows

$$\psi \uparrow_H^G (g) = \sum_i \psi(t_i^{-1} g t_i) = \frac{1}{|H|} \sum_i \sum_{h \in H} \psi(h^{-1} t_i^{-1} g t_i h) = \frac{1}{|H|} \sum_{x \in G} \psi(x^{-1} g x).$$

This is because the trace of a block form matrix is the sum of the characters of the matrices on the diagonals. And we know conjugation doesn't affect the character, so the middle equation effectively adds each of the character values to itself $|H|$ times then divides by $|H|$. Finally the $t_i h$ represent every element of a transversal multiplied by every element of H which gives us every element of G exactly once.

The following result due to Frobenius is known as **Frobenius Reciprocity**.

Theorem 3.1.21 *Let $H \leq G$ and suppose that ψ and χ are characters of H and G, respectively. Then*

$$\langle \psi \uparrow_H^G, \chi \rangle = \langle \psi, \chi \downarrow_H^G \rangle,$$

Proof We have the following string of equalities:

$$\begin{aligned}
\langle \psi \uparrow_H^G, \chi \rangle &= \frac{1}{|G|} \sum_{g \in G} \psi \uparrow_H^G (g) \chi(g^{-1}) \\
&= \frac{1}{|G||H|} \sum_{x \in G} \sum_{g \in G} \psi(x^{-1} g x) \chi(g^{-1}) && \text{(note preceding this theorem)} \\
&= \frac{1}{|G||H|} \sum_{x \in G} \sum_{y \in G} \psi(y) \chi(x y^{-1} x^{-1}) && \text{(let } y = x^{-1} g x) \\
&= \frac{1}{|G||H|} \sum_{x \in G} \sum_{y \in G} \psi(y) \chi(y^{-1}) && (\chi \text{ is a class function}) \\
&= \frac{1}{|H|} \sum_{y \in H} \psi(y) \chi(y^{-1}) && (\psi \text{ is zero outside } H) \\
&= \langle \psi, \chi \downarrow_H^G \rangle.
\end{aligned}$$

□

That concludes the necessary background. Now we will prove some results to get more comfortable with the theory before constructing the irreducibles of the symmetric group.

3.2 Interesting Problems

The following problem deals with the properties of kernels and faithful representations. This is useful because faithful representations can better help us understand abstract groups because the map each element to a more concrete object, and as stated above this is the general goal of representation theory.

Problem 1 If X is a matrix representation of a group G, then its **kernel** is the set $N = \{g \in G \mid X(g) = I\}$ where I is the identity matrix. A representation is **faithful** if it is one-to-one.

1. Show that N is a normal subgroup of G and find a condition on N equivalent to the representation being faithful.

   Let $g \in G$ and $n \in N$ to show N is normal it suffices to show $g^{-1} n g \in N$. But representations are multiplicative so with the definition of N we have

   $$X(g^{-1} n g) = X(g^{-1}) X(n) X(g) = X(g^{-1}) I X(g) = X(g^{-1}) X(g) = X(g^{-1} g) = X(\epsilon) = I$$

   so $g^{-1} n g \in N$ as desired. We know from group theory that a function is injective if and only if it's kernel is trivial so if $N = \{\epsilon\}$ then the representation is faithful.

2. Suppose X has character χ and degree d. Prove that $g \in N$ if and only if $\chi(g) = d$.

   If $g \in N$ then $X(g) = I_d$ which is exactly the matrix with d ones on the diagonal so $\chi(g) = d$. If $\chi(g) = d$ we need a little development. Notice for each $g \in G$ we have $X(g)$ is invertible so we know from linear algebra that any complex matrix is unitarily equivalent to an upper triangular matrix. So there exists a $d \times d$ matrix Y that is upper triangular such that $X(g)$ is similar to Y. Furthermore for each $g \in G$ there exists some $n \in \mathbb{N}$ such that $g^n = \epsilon$ because Y is similar to $X(g)$ we have some matrix T such that

   $$I = X(g)^n = (TYT^{-1})^n = TYT^{-1}TYT^{-1} \cdots TYT^{-1} = TY^nT^{-1}$$

   Conjugating both sides by T^{-1} the following. $T^{-1}IT = T^{-1}TY^nT^{-1}T$ hence $Y^n = I$ because I commutes with everything. Now notice $Y(g)^n = I$ and the diagonal entries $(y_{i,i}^n)$ for a product of upper triangular matrices are just the product of the diagonal entries $(y_{i,i})^n$ which in this case equals 1 so $\chi(g)$ is a sum of roots of unity. Now for any root of unity x we know $|x| \leq 1$ so if the sum of d roots of unity is d the they must all be one.

   Now we will induct on d to show $X(g) = Y = I$ hence $g \in N$. The case where $d = 1$ obvious because the matrix has only one entry. Now assume this is true for $d = k$, and let Y be an $k + 1 \times k + 1$ matrix with all other properties the same as above. We note that when multiplying upper triangular matrices the elements in the last column, $k + 1$ in this case, only impact elements in the last column of the product matrix due to the nature of matrix multiplication. So ignoring the last row and column we can invoke the induction hypothesis and see that Y must have ones on the diagonal and zeros everywhere else except potentially in the last column. But then for each element of the last column $0 = (y_{i,k+1}^n) = n(y_{i,k+1})$ for $1 \leq i \leq k$. So Y must be the identity matrix as desired.

3. Show that the kernel of the coset representation is $N = \bigcap_i g_i H g_i^{-1}$, where the g_i form a transversal.

   The coset representation is defined based on the action of G on the set of cosets, so saying $g \in N$ is equivalent to saying g acts trivially on all cosets g_iH. So let $g \in \bigcap_i g_i H g_i^{-1}$ we want to show that $g(g_iH) = g_iH$ for all i. But for any g_iH we know $g \in g_iHg_i^{-1}$ so $g = g_ihg_i^{-1}$ for some $h \in H$ hence $g(g_iH) = g_ihg_i^{-1}(g_iH) = g_ihH = g_iH$. So $g \in N$ and $\bigcap_i g_i H g_i^{-1} \subseteq N$.

   Now assume $g \in N$. We want to show $g \in g_iHg_i^{-1}$ for all i. But for any i we know $(gg_i)H = g(g_iH) = g_iH$ so from group theory we have $h \overset{def}{=} g_i^{-1}(gg_i) \in H$ so $g = g_ihg_i^{-1} \in g_iHg_i^{-1}$ and $N \subseteq \bigcap_i g_i H g_i^{-1}$ which completes the result.

4. For each of the following representations, under what conditions are they faithful:

   - Trivial: The trivial representation of any group is the one which maps every group element to the 1×1 identity matrix (1). The trivial representation is only faithful on a group of one element.
   - Regular: The regular representation is a module. Given a finite group $G = \{g_1, \cdots, g_n\}$, then we have the corresponding G-module

     $$\mathbb{C}[G] = \{c_1\mathbf{g_1} + \cdots + c_n\mathbf{g_n} \mid c_i \in \mathbb{C}\}$$

     This is called the **group algebra** of G. And G acts on it in a natural way as follows.

     $$g(c_1\mathbf{g_1} + \cdots + c_n\mathbf{g_n}) = c_1(\mathbf{gg_1}) + \cdots + c_n(\mathbf{gg_n})$$

     for all $g \in G$. The regular representation is always faithful.
   - Coset: The coset representation we have already seen. This is faithful when $\bigcap_i g_i H g_i^{-1} = \{\epsilon\}$.
   - Sign for S_n: Every element of S_n can be written as a product of 2-cycles, which may not be disjoint. The **Sign Function**, denoted $\text{sgn}(\pi)$, on elements of S_n, is equal to 1 if π can be written as an even number of 2-cycles and -1 if the number is odd. This leads naturally to the sign representation of S_n which maps each element $\pi \in S_n$ to $\text{sgn}(\pi)$. The sign representation for S_n is faithful for all $n \leq 2$.

- Defining for S_n: The defining representation for S_n maps each element to a so called "permutation matrix" which has exactly one 1 in each row and column corresponding to the position of each number $1, \cdots, n$ in the permutation. The defining representation is always faithful.
- Degree 1 for C_n: If C_n is written as $\{\epsilon, g, g^2, \cdots, g^{n-1}\}$ then all degree one representations can be obtained by mapping g to an nth root of unity. These are faithful when $g \mapsto e^{2k\pi i/n}$ where $gcd(k, n) = 1$.

5. Define a function Y on the group G/N by $Y(gN) = X(g)$ for $gN \in G/N$.

   - Prove that Y is a well-defined faithful representation of G/N.

     First we show Y is well defined. Let g_1 and g_2 be two representatives for the same coset (i.e. $g_1N = g_2N$), we need to prove $Y(g_1N) = Y(g_2N)$, but that's equivalent to saying $X(g_1) = X(g_2)$. But they are in the same coset so there is some $n \in N$ such that $g_1 = g_2n$ so by the definitions of X and N we have

     $$X(g_1) = X(g_2n) = X(g_2)X(n) = X(g_2)I = X(g_2)$$

     so Y is well defined.

     Next to prove Y is a representation of G/N we check two conditions. First $Y(N) = X(\epsilon) = I$. Second if gN and hN are cosets then $Y(gN)Y(hN) = X(g)X(h) = X(gh) = Y(ghN) = Y(gNhN)$. So Y is a representation of G/N.

     Lastly to show Y is faithful we need to show that its kernel N_Y is trivial. Let $gN \in N_Y$ then $I = Y(gN) = X(g)$ so $g \in N$ by definition, hence $gN = N$ and Y is faithful.
   - Show that Y is irreducible if and only if X is.

     Notice that for each $g \in G$, $Y(gN) = X(g)$ so if χ is the character associated with X and ψ is the character for Y then $\psi(gN) = trY(gN) = trX(g) = \chi(g)$. Each $g \in G$ can be mapped to a coset with the same character value so

     $$Y \text{ is irreducible} \iff \langle \psi, \psi \rangle = 1 \iff \frac{1}{|G|/|N|} \sum_{gN \in G/N} \psi(gN)\psi(gN^{-1}) = 1$$

     $$\iff \frac{1}{|G|} \sum_{gN \in G/N} |N|\chi(g)\chi(g^{-1}) = 1 \iff \frac{1}{|G|} \sum_{g \in G} \chi(g)\chi(g^{-1}) = 1$$

     $$\iff \langle \chi, \chi \rangle = 1 \iff X \text{ is irreducible.}$$
   - If X is the coset representation for a normal subgroup H of G, what is the corresponding representation Y?

     From 5 above we know $N = \bigcap_i g_i H g_i^{-1}$ but H is normal so $\bigcap_i g_i H g_i^{-1} = H$, thus $G/N = G/H$. So the cosets gN are the same as the cosets gH on which G acts so Y is the regular representation of G/H (the sepcial case of a coset representation in which the whole group acts on itself).

Next we will expand our results about induced characters. This will further expand our ability to relate representations of groups and subgroups, and increase our functional ability to work within representation theory.

Problem 2 Show that induction is transitive as follows. Suppose we have groups $G \geq H \geq K$ and a matrix representation X of K. Then

$$X \uparrow_K^G \cong (X \uparrow_K^H) \uparrow_H^G .$$

We will do this in two steps. First we will finish the problem from earlier proving that any two induced representations from a subgroup K to a group G are isomorphic. We have already shown that regardless of the transversal chosen the cosets are equivalent, so we can permute indexes prior to building the block matrices and the resulting representations are the same, but what if we don't have that level of control. Let $\{t_1, \cdots, t_l\}$ be a transversal for K in G. And let π be a

permutation of the numbers 1 through l. Then if X is a matrix representation for K and Y and Z are induce representations stemming from $\{t_1, \cdots, t_l\}$ and $\{t_{\pi(1)}, \cdots, t_{\pi(l)}\}$ respectively then we want to show that $Y \cong Z$. But this just requires looking at their respective characters. Notice for $g \in G$, $trY(g) = \sum_{i=1}^{l} trX(t_i^{-1} g t_i)$ and $trZ(g) = \sum_{i=1}^{l} trX(t_{\pi(i)}^{-1} g t_{\pi(i)})$. But addition is commutative in $\mathbb{C}$ so the two expressions are equivalent. Thus Y and Z are isomorphic. This completes the proof that any two induced representations from a subgroup K to a parent group G are isomorphic. In other words $X \uparrow_K^G$ is well defined.
Now to solve the problem at hand we just construct the required representation. Let X be a representation of K with transversal $\{t_1, \cdots, t_n\}$ for K in H and $\{s_1, \cdots, s_m\}$ for H in G. Then we have

$$(X \uparrow_K^H) \uparrow_H^G (g) = ((X \uparrow_K^H)(s_i^{-1} g s_j)) = (X(t_k^{-1} s_i^{-1} g s_j t_l)) =$$

$$\begin{pmatrix} X(t_1^{-1}s_1^{-1}gs_1t_1) & \cdots & X(t_1^{-1}s_1^{-1}gs_1t_n) & & X(t_1^{-1}s_1^{-1}gs_mt_1) & \cdots & X(t_1^{-1}s_1^{-1}gs_mt_n) \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ X(t_n^{-1}s_1^{-1}gs_1t_1) & \cdots & X(t_n^{-1}s_1^{-1}gs_1t_n) & & X(t_n^{-1}s_1^{-1}gs_mt_1) & \cdots & X(t_n^{-1}s_1^{-1}gs_mt_n) \\ & \vdots & & \ddots & & \vdots & \\ X(t_1^{-1}s_m^{-1}gs_1t_1) & \cdots & X(t_1^{-1}s_m^{-1}gs_1t_n) & & X(t_1^{-1}s_m^{-1}gs_mt_1) & \cdots & X(t_1^{-1}s_m^{-1}gs_mt_n) \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ X(t_n^{-1}s_m^{-1}gs_1t_1) & \cdots & X(t_n^{-1}s_m^{-1}gs_1t_n) & & X(t_n^{-1}s_m^{-1}gs_mt_1) & \cdots & X(t_n^{-1}s_m^{-1}gs_mt_n) \end{pmatrix},$$

and this is the definition of the induced representation from K to G with transversal $\{s_1t_1, \cdots, s_1t_n, \cdots, s_mt_1, \cdots, s_mt_n\}$ so together with the last result we have

$$X \uparrow_K^G \cong (X \uparrow_K^H) \uparrow_H^G .$$

as desired.

Now that we've built a solid background in representation theory and gotten comfortable with its use we will look at a way to construct the irreducible representations of the symmetric group.

3.3 Breaking Down S_n

We already know that the number of irreducible representations of a group is equal to the number of conjugacy classes. And we know that the number of conjugacy classes of S_n is the number of partitions of n. So in this section we will find a way to construct an irreducible representation of the symmetric group from each partition of n and then show that they are each unique.

We know that S_n is the set of all permutations of the set $\{1, \cdots, n\}$. In general if A is a set we can define S_A to be the set of all permutations of the elements of A. With this in mind we start by finding a subgroup of S_n based on a given partition of n.

Definition 3.3.1 *Let $\lambda = (\lambda_1, \cdots, \lambda_l) \vdash n$. Then the corresponding* ***Young subgroup*** *of S_n is*

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

In general it is easy to see that $S_\lambda \cong S_{\lambda_1} \times \cdots \times S_{\lambda_l}$. This subgroup lends itself to a particular representation which can be used to define an induced representation of S_n, but we will construct and equivalent one by other means which will provide a greater understanding of the tools we are working with. First we need a couple more objects.

Definition 3.3.2 *Suppose $\lambda \vdash n$. A* ***Young tableau*** *of shape λ, is an array t obtained by replacing the dots of the Ferrers diagram for λ with the numbers $1, \cdots, n$ bijectively.*

A **Ferrers diagram** is a way of visually representing a partition. For example the Ferrers diagram of $\lambda = (2, 1)$ is

$$\begin{array}{ll} \bullet & \bullet \\ \bullet & \end{array}$$

For notational purposes we let $t_{i,j}$ stand for the entry of t in position (i, j). A Young tableau of shape λ is also called a λ-**tableau** and denoted by t^λ. We also write shape $t = \lambda$. Clearly there are $n!$ Young tableaux of shape $\lambda \vdash n$. For example

$$\lambda = (2, 1)$$

then a list of all possible tableaux of shape λ is

$$t: \begin{matrix} 1 & 2 \\ 3 & \end{matrix}, \begin{matrix} 2 & 1 \\ 3 & \end{matrix}, \begin{matrix} 1 & 3 \\ 2 & \end{matrix}, \begin{matrix} 3 & 1 \\ 2 & \end{matrix}, \begin{matrix} 2 & 3 \\ 1 & \end{matrix}, \begin{matrix} 3 & 2 \\ 1 & \end{matrix}.$$

To consolidate some of these possibilities we have the following:

Definition 3.3.3 *Two λ-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, or λ-**tabloid**, is then*

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

where shape $t = \lambda$.

If $\lambda = (\lambda_1, \cdots, \lambda_l) \vdash n$, then the number of tableaux in any given equivalence class is $\lambda_1!\lambda_2!\cdots\lambda_l! \overset{def}{=} \lambda!$. Thus the number of λ-tabloids is just $n!/\lambda!$.

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

$$\pi\{t\} = \{\pi t\}.$$

It is easy to confirm that this action is well defined, and this gives rise to an S_n-module.

Definition 3.3.4 *Suppose $\lambda \vdash n$. Let*

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

*where $\{t_1\}, \cdots, \{t_k\}$ is a complete list of λ-tabloids regarded as a basis for a complex vector space. Then M^λ is called the **permutation module** corresponding to λ.*

The M^λ enjoy the following general property of modules.

Definition 3.3.5 *Any G-module M is **cyclic** if there is a $v \in M$ such that*

$$M = \mathbb{C}Gv,$$

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

Since any λ-tabloid can be taken to any other tabloid of the same shape by some permutation, M^λ is cyclic. We summarize in the following proposition.

Proposition 3.3.6 *If $\lambda \vdash n$, then M^λ is cyclic, generated by any given λ-tabloid. In addition, $dimM^\lambda = n!/\lambda!$, the number of λ-tabloids.*

Notice that these tabloids resemble the subgroups described earlier.

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

contains the same information as the tabloid

$$\{t^\lambda\} = \begin{matrix} 1 & 2 & \cdots & \lambda_1 \\ \lambda_1+1 & \lambda_1+2 & \cdots & \lambda_1+\lambda_2 \\ & \vdots & & \\ n-\lambda_l+1 & \cdots & n & \end{matrix}.$$

Each permutation contained in S_λ can be written in a tableaux contained in the equivalence class that is the above tabloid.

Now that we have constructed an S_n-module, M^λ, for each partition of n, it would be nice if they were all inequivalent and irreducible, but we're not quite there yet. We will, however, be able to find irreducible modules from them. But first we will introduce two useful orderings on partitions, one partial and one total, and some properties.

Definition 3.3.7 *Suppose $\lambda = (\lambda_1, \cdots, \lambda_l)$ and $\mu = (\mu_1, \cdots, \mu_m)$ are partitions of n. Then λ* ***dominates*** *μ, written $\lambda \trianglerighteq \mu$, if*

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

for all $i \geq 1$. If $i > l$ (respectively, m), then we take λ_i (respectively, μ_i) to be zero.

The dominance order just described can be related to tableaux in the following way with the **Dominance Lemma for Partitions**:

Lemma 3.3.8 *Let t^λ and s^μ be 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 3.3.8.1 *By hypothesis, we can sort the entries in each column of t^λ so that the elements of rows $1, \cdots, i$ of s^μ all occur in the first i rows of t^λ. Thus*

$$\begin{aligned}\lambda_1 + \cdots + \lambda_i &= \textit{number of elements in the first } i \textit{ rows of } t^\lambda \\ &\geq \textit{number of elements of } s^\mu \textit{ in the first } i \textit{ rows of } t^\lambda \\ &= \mu_1 + \cdots + \mu_i.\end{aligned}$$

The dominance ordering on partitions is only a partial order (there are elements that are not related), the next ordering is a total order similar to that of a dictionary.

Definition 3.3.9 *Let $\lambda = (\lambda_1, \cdots, \lambda_l)$ and $\mu = (\mu_1, \cdots, \mu_m)$ be partitions of n. Then $\lambda < \mu$ in* ***lexicographic order*** *if, for some index i,*

$$\lambda_j = \mu_j \textit{ for all } j < i \textit{ and } \lambda_i < \mu_i.$$

The lexicographic order is a refinement of the dominance order in the following way which is easy to prove and will not be done here.

Proposition 3.3.10 *If $\lambda, \mu \vdash n$ with $\lambda \trianglerighteq \mu$, then $\lambda \geq \mu$.*

Now that we have this we will identify a few more objects and then start constructing the submodules.

Definition 3.3.11 *Suppose that the tableau t has row $R_1, \cdots, R_l$ and columns $C_1, \cdots, C_k$. Then*

$$R_t = S_{R_1} \times \cdots \times S_{R_l}$$

and

$$C_t = S_{C_1} \times \cdots \times S_{C_k}$$

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

Note that our tableau equivalence classes can be written as $\{t\} = R_t t$. In addition, these groups are associated with certain elements of the group algebra $\mathbb{C}[S_n] = \{c_1\pi_1 + \cdots + c_{n!}\pi_{n!} \mid c_i \in \mathbb{C}\}$. In general, given a subset $H \subseteq S_n$, we can form the group algebra sums

$$H^+ = \sum_{\pi \in H} \pi$$

and

$$H^- = \sum_{\pi \in H} sgn(\pi)\pi.$$

For a tableau t, the element R_t^+ is already implicit in the corresponding tabloid. But we will also make use of

$$\kappa_t \stackrel{def}{=} C_t^- = \sum_{\pi \in C_t} sgn(\pi)\pi.$$

Note that if t has columns $C_1, \cdots, C_k$, then κ_t factors as

$$\kappa_t = \kappa_{C_1}\kappa_{C_2}\cdots\kappa_{C_k}.$$

Finally, we can pass from t to an element of the module M^λ by the following definition.

Definition 3.3.12 *If t is a tableau, then the associated* ***polytabloid*** *is*

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

The next lemma describes what happens to the previously defined objects passing form t to πt. The proofs are straight forward and will be omitted.

Lemma 3.3.13 *Let t be a tableau and π be a permutation. Then*

1. $R_{\pi t} = \pi R_t \pi^{-1}$,
2. $C_{\pi t} = \pi C_t \pi^{-1}$,
3. $\kappa_{\pi t} = \pi \kappa_t \pi^{-1}$,
4. $e_{\pi t} = \pi e_t$.

Note: The above R and C are algebras defined by subgroups of S_n, and the other two objects are formed from operations on those algebras. Multiplying a permutation by a tableau permutes the elements resulting in a new tableau. The point of the above lemma is to show how the action on a tableau affects the object we build from it.

Now we have everything we need to define the irreducible modules of S_n.

Definition 3.3.14 *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 a tabloid of shape λ.*

Because of lemma 3.3.13 part 4, and similar to M^λ, we have the following

Proposition 3.3.15 *The S^λ are cyclic modules generated by any given polytabloid of shape λ.*

The construction is complete. We have now found one S_n-module for each conjugacy class of S_n. It remains to show that these are irreducible and inequivalent. In doing this we will complete the goal of this section.

Again we must build up several smaller results to achieve this final goal. We start with the **Sign Lemma**.

Lemma 3.3.16 *Let $H \leq S_n$ be a subgroup.*

1. *If $\pi \in H$, then*
$$\pi H^- = H^- \pi = (sgn\pi)H^-.$$
Otherwise put: $\{\pi\}^- H^- = H^-$.
2. *For any $u, v \in M^\lambda$,*
$$\langle H^- u, v\rangle = \langle u, H^- v\rangle.$$
3. *If the transposition $(b, c) \in H$, 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*
$$H^-\{t\} = 0,$$
as a vector.

The proofs are omitted here, but we note that as in previous results it relies on the unique inner product on M^λ for which

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

This leads to two useful corollaries, the proofs are not entirely trivial, but we will accept them for out purposes.

Corollary 3.3.17 *Let $t = t^\lambda$ be a λ-tableau and $s = s^\mu$ be a μ-tableau, where $\lambda, \mu \vdash n$. If $\kappa_t\{s\} \neq 0$, then $\lambda \trianglerighteq \mu$. And if $\lambda = \mu$, then $\kappa_t\{s\} = \pm e_t$.*

Corollary 3.3.18 *If $u \in M^\mu$ and shape $t = \mu$, then $\kappa_t u$ is a scalar multiple of e_t.*

With all of this we can now prove the **Submodule Theorem**.

Theorem 3.3.19 *Let U be a submodule of M^μ. Then*

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

In particular, when the field is $\mathbb{C}$, the S^μ are irreducible.

Proof 3.3.19.1 *To remain consistent we only prove this in terms of complex numbers. Consider $u \in U$ and a μ-tableau t. By the preceding corollary, we know that $\kappa_t u = ce_t$ for some $c \in \mathbb{C}$. There are two cases, depending on which multiples can arise.*

Suppose that there exist u and t with $c \neq 0$. Then since u is in the submodule U, we have $ce_t = \kappa_t u \in U$. Thus $e_t \in U$ (since c is nonzero) and $S^\mu \subseteq U$ (since S^μ is cyclic). On the other hand, suppose we always have $\kappa_t u = 0$. We claim that this forces $U \subseteq S^{\mu\perp}$. Consider any $u \in U$. Given an arbitrary μ-tableau t, we can apply part 2 of the sign lemma to obtain

$$\langle u, e_t\rangle = \langle u, \kappa_t\{t\}\rangle = \langle \kappa_t u, \{t\}\rangle = \langle 0, \{t\}\rangle = 0.$$

Since the e_t span S^μ, we have $u \in S^{\mu\perp}$, as claimed.

We need one more result with proof omitted and that will lead us to our final theorem.

Proposition 3.3.20 *Suppose $\theta \in Hom(S^\lambda, M^m u)$ is nonzero. Then $\lambda \trianglerighteq \mu$, and if $\lambda = \mu$, then θ is multiplication by a scalar.*

Where given any two groups G and H we write $\mathbf{Hom}(G, H)$ for the set of homomorphisms from G to H.

And finally we have the much anticipated theorem about Specht modules.

Theorem 3.3.21 *The S^λ for $\lambda \vdash n$ form a complete list of irreducible S_n-modules over the complex field.*

Proof 3.3.21.1 *The S^λ are irreducible by the submodule theorem and the fact that $S^\lambda \cap S^{\lambda\perp} = 0$ for the field $\mathbb{C}$.*

Since we have the right number of modules for a full set, it suffices to show that they are pairwise inequivalent. But if $S^\lambda \cong S^\mu$, then there is a nonzero homomorphism $\theta \in Hom(S^\lambda, M^m u)$, since $S^\mu \subseteq M^\mu$. Thus $\lambda \trianglerighteq \mu$ by the previous proposition. Similarly, $\mu \trianglerighteq \lambda$, so $\lambda = \mu$.

We have successfully constructed all of the inequivalent irreducible representations of S_n. However we needed to make some choices in the process, so we really just showed that it is possible by this method. The next step to really work with these objects would be to standardize those choices to make later theorems cleaner and use these results in practice. This concludes our study of the symmetric group. It is a useful tool and object of study in many ways in mathematics, and here we have seen a way to break it down and better understand its contents.

References

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

[2] David S. Dummit, Richard M. Foote, Abstract Algebra, John Wiley and Sons Inc., 2004.

Author's Note

While most of this paper is a summary of a year long study of the symmetric group, Section 3.2 represents original work guided by problems in the books referenced.