Classifying Representations of Symmetric Groups

Satyaki Sen

July 12, 2026

1 Background

> Groups, like men, will be judged by their actions, not their words. That is why we should study their representations, not their presentations.

Representation theory is at the most basic sense, the study of group actions on vector spaces. Despite the simple definition, we can derive many unexpected and rich results by examining the ways that a certain group acts on a vector space. It is a topic that rears its head in many different fields, including parts of math like algebraic number theory, but also in more unexpected places, like in chemistry and even to describe particles in quantum physics.

Representation theory as we know it was invented mostly by Frobenius in the late 19th and early 20th centuries. In fact, throughout much of the 18th century *all* groups were to be considered as subgroups of the group of matrices over a vector space, or else the permutations of a set. Frobenius's original motivation for introducing many of the concepts he invented(such as the concept of *characters*). The problem Frobenius was working on was originally about the decomposition of a certain determinant associated with group into irreducible polynomials. The theory he ended up needing was the decomposition of the *regular representation* into irreducibles, which we encounter in 3 later. Due to the work of Weyl and Chevalley, the focus shifted away from looking at finite groups to more complicated lie groups and lie algebras; the original motivating problem proved much less important than the theory developed.

In this paper we will restrict our attention entirely to representations of finite groups over fields of characteristic 0, which are very well understood. However, much attention in the field falls on the representations of certain infinite groups known as *Lie groups*: groups that are also manifolds. The representations of finite groups are in some sense mostly "solved", but there is some very interesting structure hiding within them. In particular, the symmetric group S_n has some very combinatorial properties, and we will see later many relations between certain combinatorial objects known as *Young tableaux* and these representations. The second major connection we will mention here is the relation between representations of S_n and *symmetric polynomials*. Another connection that we will not cover in any detail is the so-called *Schur-Weyl duality* : the representations of $GL(n)$ correspond cleanly with representations of S_n.

2 Basics

Throughout this paper, we assume that all our fields have characteristic 0(like the real and complex numbers). We will usually work with vector spaces over $\mathbb{C}$, but we will use $\mathbb{R}$ as well on ocassion. Representation theory over a field of characteristic p is known as *modular* representations, and modular representations behave differently from those over $\mathbb{C}$. With that, we can define a group representation

Definition. A *representation* of a group G is a group action of G onto a vector space V. Formally, we can say that the representation is a group homomorphism ρ from G to $GL(V)$; the group of invertible linear transformations from V to V.

In a piece of somewhat confusing notation which is used very frequently, we will write V *is a representation of* G and omit the homomorphism ρ itself when saying this.

Example. At the very beginning, we alluded to symmetries of a plane figure being group representations. As an example, the square has symmetry group D_4. Corresponding to this, we have a representation of D_4 in $\mathbb{R}^2$, taking the rotation r to the rotation matrix

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

and the reflection s to the matrix

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

One useful property of matrices is that they can be added together. Although $GL(V)$ is not closed under addition, the set of *all* linear transformations is. This set is actually an *algebra*; we can add, multiply, and multiply by scalars. This $\mathbb{C}$-algebra is denoted $\operatorname{End} V$. We'd love to be able to add elements of the groups to correspond with this structure. To do this, define the *group algebra*, which allows addition in the most artificial way.

Definition. The *group algebra* $\mathbb{C}[G]$ is defined as a vector space over $\mathbb{C}$, where we take the group elements as a basis. Accordingly, we let ourselves multiply elements of the vector space using normal group multiplication.

As an example, for some g, h in some G, $2g + 3h$ and $(4 + i)g - h$ are elements of $\mathbb{C}[G]$. The sum is $(6 + i)g + 2h$, and the product is $(2g + 3h)((4 + i)g - h) = (8 + 2i)g^2 - 3h^2 + (12 + 3i)hg - 2gh$. Now, the representations of G into $GL(V)$ clearly correspond to a representation of $\mathbb{C}[G]$ into $\operatorname{End} V$, by simply taking a linear combination of the matrices $\rho(g)$.

Now, we would like to combine representations in some way, analogous to how we combine vector spaces. The two main operations of the direct sum and tensor product are defined in the obvious way. Let V and W be two representations of G. Then $V \oplus W$ is a representation by letting

$$g(v, w) = (gv, gw)$$

and $V \otimes W$ is another representation by taking

$$g(v \otimes w) = (gv) \otimes (gw),$$

and extending this linearly to the rest of $V \otimes W$. We sometimes also use the tensor product in a more general sense: if V and W are representations of two groups G and H, $V \boxtimes W$ is the representation of $G \times H$ on the tensor product space by $(g, h)(v \otimes w) = (gv \otimes hw)$. This is known as the *exterior* tensor product. In terms of matrices, the direct sum looks like

$$\begin{pmatrix} \rho_V(g) & 0 \\ 0 & \rho_W(g) \end{pmatrix}$$

A slightly more involved construct is the representation into the dual space V^* of V. We define ρ^* so that

$$\rho^*(g) = \rho(g^{-1})^\top$$

(recall that the transpose $T^\top$ of a map $T : V \to W$ is really a map from W^* to V^* defined so that $T^\top$ takes each $f \in W^*$ to $f \circ T \in V^*$). One important use-case of this construction that will come up is the space of homomorphisms $V \to W$. We can construct a representation into this space by using the identity $\operatorname{Hom}(V, W) = V^* \otimes W$.

Definition. A *homomorphism of* G*-modules* is a homomorphism f between two representations V and W that respects the action of G, that is, $f(\rho_V(g)v) = \rho_W(g)f(v)$ for each $v \in V$. Such a map is sometimes called an *intertwining*, *equivariant*, or a G*-linear* map.

We can now start asking about how to classify all representations of a group G. However, we immediately run into a problem; because there are infinitely many reps obtained from just taking direct sums. Therefore, we should try to look at the representations that cannot be expressed as a direct sum $V \oplus W$. Clearly any direct sum will have a subspace that's fixed by the action of G. The converse is the following theorem.

Theorem (Maschke's Theorem)**.** *Any representation V with a nontrivial subspace W fixed by the action of G can be written as a direct sum $W \oplus W'$.*

Proof. The key point is that if we can define an inner product on V $\langle v_1, v_2 \rangle$ that is invariant under the action of g(i.e $\langle g\cdot, g\cdot \rangle = \langle \cdot, \cdot \rangle$), then there will be a perpendicular subspace $W^\perp$ of vectors orthogonal to all $w \in W$. But this subspace must be preserved by G, because if $w' \in W^\perp$, then for all $u \in W$, $\langle gw', u \rangle = \langle gw', g(g^{-1}u) \rangle = \langle w', g^{-1}u \rangle = 0$, which means $gw' \in W^\perp$. Now we will just have $V = W \oplus W^\perp$, so it remains to construct the inner product.

The key point is to take *any* hermitian form $H_0(v, w)$ and average it over all g(this idea will come back later). Define:

$$\langle v, w \rangle = \sum_{g \in G} H_0(gv, gw),$$

and multiplying v and w by any h does not change this quantity, because $hG = G$ as a set. □

One of the reasons that much of our theory will not work over fields of nonzero characteristic is that this theorem is only true when the characteristic of the field does not divide the order of the group. The property that ay fixed subspace implies a decomposition is known as begin *semisimple.*

Definition. A representation which *cannot* be written as a direct sum(equivalently, one that does not have a fixed subspace) is known as *irreducible.*

Although the following result about maps between irreducible representations is quite simple, it will actually be key to many proofs later on. Using an averaging trick like the above, we can construct a map that must be a G-module homomorphism, and use the result.

Lemma (Schur's Lemma)**.** *If ϕ is a nontrivial(i.e $\phi \neq 0$) G-module homomorphism between two irreducible representations V and W, then V and W are isomorphic. If $V = W$, then ϕ must be multiplication by some constant λ.*

Proof. Note that $K = \ker \phi$ is a subspace of V. But for any k in K, $\phi(gk) = g\phi(k) = 0_W$, which means that K is a fixed subspace by G. By Maschke's Theorem, $\ker \phi$ must be either 0 or all of V. If $\ker \phi = V$, then ϕ is trivial, and if $\ker \phi = 0$, then ϕ is an isomorphism. To see the second part of the theorem; note that ϕ must have at least one eigenvalue λ. But now consider the λ-eigenspace E, where $\phi w = \lambda w$ for $w \in E$. We see $\phi(gw) = g\phi(w) = g\lambda w = \lambda(gw)$, so W is a fixed subspace, hence $E = V$. □

If we let $\mathrm{Hom}_G(V, W)$ be the space of G-module homomorphisms from V to W, Schur's Lemma states that this space is trivial for two distinct irreps, and dimension 1 when they are the same. As a corollary we see that every representation of a finite group can be uniquely written down as a direct sum of irreducibles. For if we had two decompositions $a_1V_1 \oplus a_2V_2 \cdots \oplus a_kV_k$ and $b_1W_1 \dots b_jW_j$, then the identity map is a G-module homomorphism, and hence takes each irreducible factor to an isomorphic irreducible factor in the second decomposition.

One natural question to ask is how representations of a subgroup H relate to representations of G. Obviously, each rep of G can be restricted to a representation of H, by only considering the elements of H. We denote this restricted representation as $\mathrm{Res}_H^G(V)$. The more complicated direction is the reverse. How can a representation V of H be lifted into a representation of G? We'll construct this *induced* representation as the "reverse" of the restriction. In a restriction, all the cosets gH which had been isomorphic parts of V are collapsed down to one coset H. In order to expand them, we actually define the induced representation as

$$\mathrm{Ind}_H^G V = \bigoplus_{\text{cosets } \sigma H} V^\sigma,$$

where the V^σ are copies of V. H fixes each copy, but each element of g simply takes one V^σ to some other $V^{g\sigma}$, depending on which coset it is a part of.

Theorem (Frobenius Reciprocity)**.** *Let V be any representation of G and let W be any representation of a subgroup H. Then*

$$\mathrm{Hom}_H(W, \mathrm{Res}_H^G V) = \mathrm{Hom}_G(\mathrm{Ind}_H^G W, V).$$

Proof. We just need to show how any homomorphism $\varphi\colon W \to \mathrm{Res}\, V$ extends to a homomorphism $\phi\colon \mathrm{Ind}\, W \to V$. To take some $g \in W^\sigma$ to V, define ϕ as the map $\sigma \circ \varphi \circ \sigma^{-1}$, which takes an element of W^σ, moves it back into W, maps it to V, and then moves it back to a coset. This is a G-module homomorphism because if we look at $\phi(ghw)$ for w in W, then this lies in coset gh, and so

$$gh\varphi((gh)^{-1}ghw) = gh\varphi(w) = g\varphi(hw),$$

as desired $\square$

Finally, let's look at two representations that will show up for any group G. The first is of course the trivial representation triv, with $\rho(g) = I$ for any ρ. The second is the *regular representation*, which we define as the action of G on the $\mathbb{C}[G]$ by left multiplication. This second representation is an example of a permutation representation: it takes a basis for $\mathbb{C}[G]$(the group elements themselves) and simply permutes it.

3 Characters

Definition. Let (V, ρ) be a representation of G. The *character* of V, denoted χ_V is defined as the function from G to $\mathbb{C}$ with $\chi_V(g) = \mathrm{tr}(\rho(g))$.

The characters are actually *class functions* on G, which means that if $g = xhx^{-1}$, $\chi(g) = \chi(h)$. This simply follows from the properties of the trace: $\mathrm{tr}(\rho(xhx^{-1})) = \mathrm{tr}(\rho(x)\rho(h)\rho(x)^{-1}) = \mathrm{tr}(\rho(h)\rho(x)^{-1}\rho(x)) = \mathrm{tr}(\rho(x))$. We will denote the vector space of class functions from G to $\mathbb{C}$ as $\mathrm{Class}(G)$. Not all elements of $\mathrm{Class}(G)$ are actually characters, the one's that aren't are known as *virtual characters*. Clearly the character of a direct sum is the sum of the characters of the summands, since we simply put the matrices together. The tensor product is slightly more complicated, but we can see that the if $v \otimes w$ is an eigenvalue of $\rho_{V\otimes W}(g)$, then v is a eigenvector of $\rho_V(g)$ and w is an eigenvector of $\rho_W(g)$. Since we can get all $\dim V \cdot \dim W$ eigenvectors in this way, each eigenvalue is a product of an eigenvalue from V and one from W. Thus, if $\rho_V(g)$ has eigenvalues λ_i and $\rho_W(g)$ has eigenvalues μ_j, the sum of eigenvalues is

$$(\lambda_1 + \cdots + \lambda_n)(\mu_1 + \cdots + \mu_m) = \chi_V(g)\chi_W(g)$$

Finally, let's try to find what the character of V^* is. The transpose in the definition can safely be ignored, since the $\mathrm{tr}(A^\top) = \mathrm{tr}(A)$. The eigenvalues of g^{-1} are of course λ_i^{-1}. Now since for any $g \in G$, $g^{|G|} = 1$, the same is true for the eigenvalues(that is, they are on the unit circle). Thus, we can replace the sum $\lambda_1^{-1} + \lambda_2^{-1} \dots \lambda_n^{-1}$ by $\overline{\lambda_1} + \cdots + \overline{\lambda_n} = \overline{\lambda_1 + \cdots + \lambda_n} = \overline{\chi_V(g)}$.

Here is where we will make use of the averaging trick we saw before. Note that for any representation (V, ρ), the linear map $f\colon V \to V$ given by

$$\frac{1}{|G|}\sum_{g\in G}\rho(g)$$

has image V^G, the set of v such that $gv = v$ for all $g \in G$. For, $h\sum_{g\in G} gv = \sum_g (hg)v = \sum gv$. For $v \in G^V$, we see that $\frac{1}{|G|}\sum gv = \frac{1}{|G|}\sum v = v$. Thus, $f \circ f = f$, and so f is the projection map onto the subspace V^G. As such, it has trace of $\dim V^G$, hence

$$\frac{1}{|G|}\sum_g \mathrm{tr}(\rho(g)) = \frac{1}{|G|}\sum_g \chi(g) = \dim V^G.$$

On it's own, this is not so interesting, as it only relates the characters of a single representation, and doesn't tell us anything about how representations are related. But consider applying this to $\mathrm{Hom}(V, W)$. The subspace $\mathrm{Hom}_G(V, W)$ contains all of the G-module homomorphisms, so

$$\dim \mathrm{Hom}_G(V, W) = \frac{1}{|G|}\sum_g \chi_{\mathrm{Hom}(V,W)}(g) = \frac{1}{|G|}\sum_g \chi_{V^*\otimes W}(g) = \frac{1}{|G|}\sum_g \overline{\chi_V(g)}\chi_W(g).$$

Definition. As in the previous equation, *define* an hermitian inner product on Class(G) by

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

The upshot of this particular definition is that the characters of the irreducible representations form an orthonormal set under this inner product. This is just Schur's Lemma: If V and W are distinct irreducibles, $\mathrm{Hom}_G(V, W)$ is empty, so $\langle \chi_V, \chi_W \rangle = 0$. If $V = W$, then $\mathrm{Hom}_G(V, W)$ is the one dimensional vector space of λI, so $\langle \chi_V, \chi_V \rangle = 0$.

In fact, we will be able to show that these characters form an orthonormal *basis*, which implies the number of irreducibles is the same as the number of conjugacy classes. For now, we at least know that there are at most this many irreps, and hence a finite number of them.

Theorem. *The characters of irreducible representations form an orthonormal basis for* Class(G)*.*

Proof. Suppose that there exists a nontrivial $\alpha \in \mathrm{Class}(G)$ with $\langle \alpha, \chi_V \rangle = 0$ for all irreps V. Since all representations can be decomposed into irreps, α is orthogonal to the character of any representation, With the same idea as before, consider the map from V to V

$$\varphi_\alpha = \sum_g \alpha(g) g$$

φ_α is a G-module homomorphism, as $\varphi_\alpha(hv) = \sum_g \alpha(g) g(hv)$ can be reindexed by substituting g with it's expression as $hg'h^{-1}$ to

$$\sum_{g'} \alpha(hg'h^{-1}) hgh^{-1}hv = \sum_{g'} \alpha(g') hgv = h\varphi_\alpha(v).$$

Again, Schur's Lemma tells us this map is just multiplication, and so it has trace $\dim V \cdot \lambda$. But $\mathrm{tr}(\varphi_A) = \sum_g \alpha(g) \chi_V(g) = |G| \langle \chi_{V^*}, \alpha \rangle = \dim V \cdot \lambda$, where we use V^* to get rid of the conjugate in the definition. As α and χ_{V^*} are orthogonal. $\lambda = 0$ and $\varphi_A = 0$.

To obtain the punchline $\alpha = 0$, consider defining this map when the representation is the regular representation R. The sum $\sum_g \alpha(g) g$ can be thought of as an element of $\mathbb{C}[G]$, so there is a linear combination of the g that is 0. But since they're defined to be linearly independent, $\alpha(g) = 0$ for each $g \in G$. $\square$

Since we know this, we can do all the things we'd normally do with an orthonormal basis. In particular if $V = a_1 V_1 \oplus a_2 V_2 \dots a_k V_k$, $\langle \chi_V, \chi_V \rangle = a_1^2 + a_2^2 + \cdots + a_n^2$, and this means V is irreudcible if and only if $\langle \chi_V, \chi_V \rangle = 1$. We can look at the regular representation in particular to derive some interesting results. Since it's a permutation representation, and the trace of a permutation matrix is the number of fixed points, $\chi_R(g) = 0$ if $g \neq 1$, and $|G|$ otherwise. Taking the inner product with any irreducible character χ_V, $\langle \chi_R, \chi_V \rangle = \frac{1}{|G|} |G| \chi_V(1) = \chi_V(1) = \dim V$. This means the multiplicity of V is $\dim V$. Considering the dimension of R, this gives us

$$\sum_i \dim V_i^2 = |G|, \tag{1}$$

where the sum goes over all irreps V_i.

Let's now try to find the characters of an induced representation. Note that $g \in G$ will take the coset W^σ to the coset $W^{g\sigma}$. If we consider this in the matrix $\rho(g)$, this effectively means that only the cosets with $\sigma = g\sigma$ will count toward the trace, as otherwise the corresponding block will not be on the main diagonal. If $\sigma = sH$ for some representative s(our choice doesn't matter), then $gsH = sH \implies s^{-1}gs \in H$. The action of G on W^σ is identical to the action of $g_0 = s^{-1}gs$ on W, as $g(sh) = sg_0 s^{-1} sh = sg_0 h$ for any $h \in H$. Hence, we have

$$\chi_{\mathrm{Ind}\, W}(g) = \sum_{g\sigma = \sigma} \chi_W(s^{-1}gs).$$

We can in fact extend this to any virtual character, as in

Definition. The *induced character* of $\chi \in \text{Class}(H)$ is $\chi\!\uparrow^G$ or $\text{Ind}_H^G \chi$, defined by

$$\chi\!\uparrow^G(g) = \sum_{g\sigma=\sigma} \chi(s^{-1}gs)$$

for an arbitrary representative s of σ. Similarly, define the *restricted character* $\text{Res}_H^G \chi = \chi\!\downarrow_H$ by just restricting χ to H.

Since these virtual characters aren't actually representations, we can't use the $\text{Hom}_G(V, W)$ idea that we used before to prove Frobenius Reciprocity. However, the result

Theorem (Frobenius Reciprocity for Characters)**.**

$$\langle \chi, \psi\!\downarrow_H \rangle_H = \langle \chi\!\uparrow^G, \psi \rangle_G$$

is still true. Since this proof is purely algebraic, we'll be able to show it in a more general setting, where χ, ψ are class functions from G, H to some arbitrary $\mathbb{C}$-algebra A.

Proof. To simplify our argument, we extend χ to G by just making $\chi(g) = 0$ when g is in G but not H. We can write the definition of the induced character as

$$\frac{1}{|H|} \sum_{s \in G} \chi(s^{-1}gs),$$

since the terms when $s^{-1}gs$ isn't in H are all 0, and each coset σ is counted $|H|$ times for each of its elements.

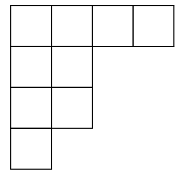
$$\begin{aligned} \langle \chi\!\uparrow^G, \psi \rangle_G &= \frac{1}{|G|} \sum_{g \in G} \chi\!\uparrow^G(g) \overline{\psi(g)} \\ &= \frac{1}{|G|} \sum_{g \in G} \frac{1}{|H|} \sum_{s \in G} \chi(s^{-1}gs) \overline{\psi(g)} \\ &= \frac{1}{|G|} \frac{1}{|H|} \sum_{s \in G} \sum_{g \in G} \chi(s^{-1}gs) \overline{\psi(g)} \end{aligned}$$

Here, reindex the inner sum by writing in in terms of $t := s^{-1}gs$. We see

$$\begin{aligned} &\frac{1}{|G|} \frac{1}{|H|} \sum_{s \in G} \sum_{g \in G} \chi(s^{-1}gs) \overline{\psi(g)} \\ &= \frac{1}{|G|} \frac{1}{|H|} \sum_{s \in G} \sum_{t \in G} \chi(t) \overline{\psi(sts^{-1})} \\ &= \frac{1}{|G|} \frac{1}{|H|} \sum_{s \in G} \sum_{t \in G} \chi(t) \overline{\psi(t)} \\ &= \frac{1}{|G|} \frac{1}{|H|} |G| \sum_{t \in G} \chi(t) \overline{\psi(t)} \\ &= \frac{1}{|H|} \sum_{t \in G} \chi(t) \overline{\psi(t)} \\ &= \frac{1}{|H|} \sum_{t \in H} \chi(t) \overline{\psi(t)} = \langle \chi, \psi\!\downarrow_H \rangle_H \end{aligned}$$

where we have use the fact that $\psi(sts^{-1}) = \psi(t)$ because ψ is a class function, and in the last step, used the fact that $\chi(t)$ is only zero inside H. $\square$

4 Tableaux



The above figure is an example of a *Young Diagram*, which is a set of boxes corresponding to a partition of some integer n. Each row has length of a term of the partition, in nonincreasing order, so the above diagram is of the partition $(4, 2, 2, 1)$ of 9. When we fill this Young Diagram with objects in the boxes, we get a Young tableau.

Definition. A *Young tableau* is a filling of a λ Young diagram with the numbers from 1 to n. We say that a tableau is standard when the numbers in the boxes strictly increase from left to right, and from top to bottom.

Below, we see a Young tableau of $\lambda = (4, 2, 2, 1)$.

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

The partitions are ordered by lexicographical order: two partitions $\lambda = (\lambda_1, \dots)$, $\mu = (\mu_1, \dots)$ of n have $\lambda > \mu$ if the first index at which λ_i, μ_i differ has $\lambda_i > \mu_i$. We will also define a partial order known as dominance, where λ dominates μ, written $\lambda \trianglerighteq \mu$, iff for all i,

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

Observe that this partial order can also be defined in the same way on any sequences, not just partitions.

Definition. A *semistandard* Young tableau t is a filling of the young diagram of λ that is weakly increasing from left to right, but strongly increasing from top to bottom. We are allowed to put any integer in the boxes, not just from 1 to n. The *content* of t is the composition $\mu = (\#\text{ of ones}, \#\text{ of twos}, \#\text{ of threes}, \dots)$

The number of semistandard tableau of shape λ and content μ is known as the *Kostka Number* $K_{\lambda\mu}$. We have one more variant of the Young tableaux that we will need.

Definition. Say two tableaux are *row-equivalent* if they have the same numbers in each row, but possibly in different columns. A *tabloid* is an equivalence class of Young tableaux under row-equivalence. We can think of this as a normal Young tableau $t$except we remove the dividing lines within rows. We denote the tabloid of t as $\{t\}$.

For instance, the tabloid

$$\begin{array}{cc}\hline 1 & 2 \\ \hline 3 & \\ \cline{1-1} \end{array}$$

is the equivalence class containing

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array} \qquad \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 \\ \cline{1-1} \end{array}.$$

We have not yet seen anything about the symmetric group, or how it should connect to these definitions. The first hint is to recall the conjugacy classes of S_n; they are actually the cycle-types of the elements: that is, if π and σ have all the same cycle lengths, they are conjugate to each other. But each cycle type is just a partition, so the conjugacy classes(and hence the number of irreducible representations) correspond to partitions of n. We can also see how S_n acts on a Young tableaux T; by simply permuting the numbers 1 to n within.

Definition. Let S_n act on the vector space whose basis is the λ-tabloids by permuting the numbers 1 to n inside the tabloid. This representation is known as the *permutation module*, and we denote it U_λ.

As an example of this action, we look at an element of S_5 acting on an tabloid

$$(145)(23)\begin{array}{cc} \hline 1 & 2 \\ \hline 3 & 5 \\ \hline 4 & \\ \cline{1-1} \end{array} = \begin{array}{cc} \hline 4 & 3 \\ \hline 2 & 1 \\ \hline 5 & \\ \cline{1-1} \end{array}$$

These representations of S_n are not irreps, but they are closely related to the true irreducible representations. First, we shall define a more algebraic definition of U_λ. Given a tableau t, let the *row symmetrizer* R_t be the subgroup of S_n containing elements that preserve the rows, and similarly define the column symmetrizer C_t. Consider the canonical young tableaux T_λ with shape λ where we just number boxes consecutively from 1 to n, for example,

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

[1] Let P_λ be defined as R_{T_λ}, and $Q_\lambda = C_{T_\lambda}$. If $g \in P_\lambda$, the tableau gT is row-equivalent to T_λ. Additionally, define two elements a_λ and b_λ in $\mathbb{C}[S_n]$, where $a_\lambda = \sum_{g \in P_\lambda} g$, and $b_\lambda = \sum_{g \in Q_\lambda} \operatorname{sign}(g) g$. These definitions clearly show that for $p \in P_\lambda$, $pa_\lambda = a_\lambda$, and for $q \in Q_\lambda$, $qb_\lambda = \operatorname{sign}(q) b_\lambda$.

Proposition. $U_\lambda = \operatorname{Ind}_{P_\lambda}^{S_n} \operatorname{triv}$, *and this is equivalent to the action of* S_n *on* $\mathbb{C}[S_n]a_\lambda$ *by left multiplication.*

Proof. Firstly, it's obvious since P_λ is defined to preserve the rows that each $g \in P_\lambda$ acts trivially on the λ-tabloids in the span of the canonical tabloid that we've chosen to define it with. Moreover, every tabloid has itself a row symmetrizing set in S_n which is a coset of P_λ, hence each coset acts trivially on a certain one-dimensional subspace of the space with the tabloids as bases, and the result follows. For the second part, note that $\mathbb{C}[S_n]a_\lambda$ is a vector space that can be generated by one element per coset of P_λ in S_n(since for any $p \in P_\lambda$, $pa_\lambda = a_\lambda$). Taking the vector space generated by group algebra element ha_λ to the subspace in the induced representation corresponding to hP_λ is an isomorphism. $\square$

Now, we come to the main definition of this paper, involving a different action of S_n on the same vector space with tabloids as basis.

Definition. Given a tableau t with shape λ, define the *polytabloid* as the *linear combination* of λ-tabloids given by

$$e_t = \sum_{\pi \in C_\lambda} \operatorname{sign}(\pi) \pi \{t\},$$

where the action of π on $\{t\}$ is as before by permuting the numbers 1 to n, and C_λ is the subgroup of elements in S_n that preserve the columns of t.

For instance with t as the example

$$t = \begin{array}{|c|c|c|} \hline 1 & 2 & 4 \\ \hline 3 & 5 \\ \cline{1-2} \end{array},$$

we see $C_\lambda = \{1, (13), (25), (13)(25)\}$, and

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

We can observe that for the tableau T_λ, the e_{T_λ} is simply b_λ times the corresponding tabloid $\{T_\lambda\}$

[1] This choice is entirely irrelevant; if we made some other choice we could just relabel the elements to see the representations are isomorphic.

Definition. The *Specht Module* $V_\lambda$is the submodule of the permutation module spanned by the e_t for all tableaux t.

Although it initially seems like V_λ and U_λ have the same dimension, the polytabloids are not, in general, independent. We will show later that the e_t over all *standard* tableau t do in fact form a basis. Even before this, we can show that the Specht Modules are irreducible representations. We'll prove one preliminary lemma about polytabloids:

Lemma 1. *If t is a tableau and π is a permutation, $e_{\pi t} = \pi e_t$.*

Proof. Observe that $C_{\pi t} = \pi C_t \pi^{-1}$, which can be viewed as a relabelling by π. Then we have

$$e_{\pi t} = \sum_{\sigma \in C_{\pi t}} \operatorname{sign}(\sigma)\sigma\{\pi t\} = \sum_{\sigma \in \pi C_t \pi^{-1}} \operatorname{sign}(\sigma)\sigma\pi t$$

$$= \sum_{\sigma' \in C_t} \operatorname{sign}(\pi\sigma'\pi^{-1})\pi\sigma'\pi^{1}\pi\{t\} = \pi \sum_{\sigma' \in C_t} \operatorname{sign}(\sigma')\sigma'\{t\} = \pi e_t$$

□

5 Specht Modules

Our goals in this section will be to prove that the Specht Modules v_λ are irreducible, and have a basis given by e_t over all standard Young tableau t.

Our proof that V_λ are irreducible is adapted from [1], but there is a less algebraic proof presented in [4]. We actually prove that the module V'_λ defined as the action of S_n by left multiplication on the ideal $\mathbb{C}[S_n]a_\lambda b_\lambda \coloneqq \mathbb{C}[S_n]c_\lambda$ is irreducible, and then prove that this is actually the Specht Module V_λ we were after.

Lemma 2. *For any $x \in \mathbb{C}[S_n]$, $a_\lambda x b_\lambda$ is a multiple of c_λ.*

Proof. Get rid of the subscript λ within the proof for convenience. If $g = (p \in P)(q \in Q)$ we just get $agb = apqb = a\operatorname{sign}(q)b = \operatorname{sign}(q)c$. Moreover, we show that if g is not of this form, $agb = 0$, and our result follows since the group elements are a basis.

To do this, we find a transposition t such that $t \in P_\lambda$, and $g^{-1}tg \in Q_\lambda$, since now $agb = atgb = ag\operatorname{sign}(g^{-1}tg)g^{-1}tgb = -\operatorname{sign}(g^{-1}g)agg^{-1}tgb = -atgb = -agb$. Suppose t permutes two elements x and y, which must be in the same row of the "canonical" tableau $T = T_\lambda$. $g^{-1}tg$ is the transposition that swaps $g^{-1}x$ and $g^{-1}y$. Thus these two elements are in the same column, or alternatively, x and y are in the same column of $gT = T'$. We'll show that such x, y exists unless there's some $p \in P$, $q' \in gQg^{-1} = C_{T'}$ with $pT = q'T'$, which is clearly equivalent to $g \in PQ$.

Now, we construct these elements using an algorithm-like approach. Assume that no pair x, y exists with x, y in the same row of T and x, y in the same column of T'. Then the numbers in the first row of T all end up in different columns of T'. This implies there exists some q'_1 in $C_{T'}$ that brings all these elements to the first row, and this still preserves the columns. Now we can choose some p_1 in R_T that permutes the first row to bring it in the same order as q'_1T'. Thus, p_1T and q'_1T' have the same first row. Now consider the second row, and with the same process construct p_2, q'_2, so on and so forth. But after we finish, $p_k p_{k-1} \dots p_1 T$ and $q'_k q'_{k-1} \dots q'_1 T'$ have the same first row, second row, etc : hence they are the same and we win. □

Lemma 3. *If $\lambda > \mu$, then $a_\lambda x b_\mu = 0$ for all $x \in \mathbb{C}[S_n]$. Recall that $\lambda > \mu$ means λ is greater than μ in lexicographical order.*

Proof. We adopt the same approach as in the previous lemma, by showing that there exists $t \in P_\lambda$ with $g^{-1}tg \in Q_\mu$. As before let $T = T_\lambda$, and $T' = gT_\mu$, and want to find two integers x, y in the same row of T, and in the same column of T'.

If $\lambda_1 > \mu_1$, we are already done, since by the pigeonhole principle there are more elements in the first row than distinct columns they may be part of. Otherwise, we use the same trick as before and construct

$p_1 \in R_T$ and $q_1' \in C_{T'}$ so that $p_1 T$ and $q_1' T'$ have the same first row. Then, try again with the second row; if $\lambda_2 > \mu_2$ we win, and otherwise we can construct p_2 and q_2'. Eventually, we will end up at some row with $\lambda_i > \mu_i$ by definition of lexicographical order, so we will able to find the desired x and y. □

As a consequence of 2, by plugging in $x = b_\lambda a_\lambda$ we see that $a_\lambda b_\lambda a_\lambda b_\lambda = c_\lambda^2$ is a multiple of c_λ, which implies that c_λ is proportional to an idempotent. Although this does not seem very useful, the following abstract lemma will allow us to use it.

Lemma 4. *Let A be an algebra and e an idempotent within it. If M is any A-module, then $\mathrm{Hom}_A(Ae, M) \cong eM$. To be precise, an element $x \in eM$ corresponds to the homomorphism f_x given by $f_x(a \in Ae) = ax$.*

Proof. First, we can note that $1-e$ is idempotent as well: $(1-e)^2 = 1 - 2e + e^2 = 1 - 2e + e = 1 - e$. Clearly $\mathrm{Hom}_A(A, M)$ is just M, because for any homomorphism f, we can take f to $f(1 \in A)$. This is injective because if $f(1) = g(1)$, then $(f-g)(1) = 0$, but since this we can pass elements of A through, $(f-g)(a) = a \cdot (f-g)(1) = 0$, and $f - g$ is identically zero, and it is clearly surjective. Now write $A = Ae \oplus A(1-e)$. We can write any $x \in M$ as $ex + (1-e)x$, and $f_x(a)$ for any $x \in M$ has a component of $aex + a(1-e)x$, and the component eM just corresponds to having only an Ae part. □

Lemma 5. *V_λ' is an irreducible representation.*

Proof. If we show that $\mathrm{Hom}_{S_n}(V_\lambda', V_\mu')$ is trivial for λ and μ distinct, and $\mathrm{Hom}_{S_n}(V_\lambda', V_\lambda')$ has dimension 1, we will get that this is a set of irreducibles, since this is equivalent to showing that $\langle \chi_{V_\lambda'}, \chi_{V_\mu'} \rangle = 0$ and $\langle \chi_{V_\lambda'}, \chi_{V_\mu'} \rangle$. Note that by 3 and 4,

$$\mathrm{Hom}_{S_n}(V_\lambda', V_\mu') = \mathrm{Hom}_{S_n}(\mathbb{C}[S_n]c_\lambda, \mathbb{C}[S_n]c_\mu) = c_\lambda \mathbb{C}[S_n]c_\mu = a_\lambda b_\lambda \mathbb{C}[S_n]a_\mu b_\mu = 0,$$

and by 2,

$$\mathrm{Hom}_{S_n}(V_\lambda', V_\lambda') = \mathrm{Hom}_{S_n}(\mathbb{C}[S_n]c_\lambda, \mathbb{C}[S_n]c_\lambda) = c_\lambda \mathbb{C}[S_n]c_\lambda = a_\lambda b_\lambda \mathbb{C}[S_n]a_\lambda b_\lambda$$

which is a one dimensional subspace of the multiples of c_λ. □

Theorem 1. *The Specht Module V_λ equals V_λ', and hence the Specht modules are irreducible.*

Proof. Establish an map from $\mathbb{C}[S_n]$ to the vector space spanned by the tabloids by letting $g \mapsto g\{T_\lambda\}$ where T_λ is our canonical choice of λ-tableau. What is the image of $gb_\lambda a_\lambda$? We see

$$gb_\lambda a_\lambda \mapsto gb_\lambda a_\lambda \{T_\lambda\} = gb_\lambda |P_\lambda| \{T_\lambda\} = |P_\lambda| ge_T = |P_\lambda| e_{gT},$$

where since every term of a_λ does nothing to $\{T_\lambda\}$, $a_\lambda\{T_\lambda\}$ is simply $|P_\lambda|\{T_\lambda\}$. Thus, $\mathbb{C}[S_n]b_\lambda a_\lambda$ is equivalent to V_λ. However, we care about $\mathbb{C}[S_n]c_\lambda = \mathbb{C}[S_n]a_\lambda b_\lambda$, which is in a different order. We will demonstrate that these are isomorphic as well.

Define f as a map from $\mathbb{C}[S_n]a_\lambda b_\lambda$ to $\mathbb{C}[S_n]b_\lambda a_\lambda$ by $f(x) = xa_{lambda}$, which maps to the desired ideal because $xa_\lambda = ya_\lambda b_\lambda a_\lambda$. This is clearly a homomorphism of S_n-modules, and it is clearly nonzero by evaluating it at c_λ(for example), so Schur's Lemma tells us that f is an isomorphism. □

Now, we embark on the proof that the standard tableaux form a basis for V_λ. Looking at the dimension, this means that the number of standard tableau of shape λ, denoted f^λ is equal to the dimension. Using a result from section 2, we get the very surprising

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

To prove this result, we introduce order on the tabloids.

Definition. Let $\{T\}$ and $\{S\}$ be two tabloids of shape λ Then $\{S\} \geq \{T\}$ if there exists some i such that

- for all $j > i$, j appears in the same row of both $\{S\}$ and $\{T\}$.
- i appears in a higher row of $\{S\}$ than in $\{T\}$.

1	2	3	6
4	7	5	
8	9		
10	11		

Figure 1: The row descent is 7 and 5, $A = \{7, 9, 11\}$, and $B = \{3, 5\}$. Note the ordering $11 > 9 > 7 > 5 > 3$

Essentially, this means we take i to be the largest number for which $\{S\}$ and $\{T\}$ differ, and compare the row for which i is in. We can clearly also extend this order to tableaux by letting $S > T$ iff $\{S\} > \{T\}$.

Lemma 6. *Let $\{T\}$ be the tabloid associated with a standard tableau T. Let π be an element of C_T. Then $\{T\} \geq \{\pi T\}$.*

Proof. Let i be the largest element that π will move to a different column. Since nothing lower than i is permuted(since i was the largest, and T increases as we go down along the columns), i must have been moved to a higher column, as desired. □

Theorem 2. *The polytabloids associated with the standard Young tableaux are linearly independent.*

Proof. Order the tableau using the above ordering, so that $T_1 < T_2 < \cdots < T_k$, and suppose

$$c_1 e_{T_1} + c_2 e_{T_2} \dots c_k e_{T_k} = 0.$$

By definition of e_{T_k}, $\{T_k\}$ appears when we write it as a sum. We must have that this $\{T_k\}$ is cancelled out by one of the previous terms. However, each previous polytabloid term will look like $\pm\{\pi T_j\}$ for $\pi \in C_{T_j}$, and by 6, this means that the term is less than $\{T_j\}$, which is less than $\{T_k\}$. But this means that they cannot be equal, and hence cannot cancel out! We have thus shown $c_k = 0$, and now the exact same argument can be repeated to show the other coefficients are also 0. □

The second thing we need to show is that the standard polytabloids actually span the space of all polytabloids. In fact, we provide an explicit algorithm to write any e_t as a a linear combination of standard polytabloids for any t. We can assume that the columns increase from top to bottom, as applying a column-preserving permutation only changes the polytabloid up to sign. The key point is that if we can write e_t as a combination of polytabloids that are "closer" to being standard, we can win by induction. The so-called straightening algorithm is as follows:

In a tableau t, pick a particular row descent that we will eliminate, say the topmost, leftmost one. Let A be the set of numbers below the left out-of-order element, and let B be the set of numbers above the right out-of-order element. Then the Garnir Element $g_{A,B} \in \mathbb{C}[S_n]$ is the sum of $\operatorname{sign}(\pi)\pi$ over all π that only permute $A \cup B$ and maintain the increasing order of the columns. The key point is that for each of these π except the identity, the row descent is eliminated. This is true because if there is still a row descent in the same spot, then $A \cup B$ is still decreasing from the bottom of A to the top of B, and so there is only one such ordering.

Lemma 7. *Let $g_{A,B}$ be a Garnir element of a tableau T. Then $g_{A,B} e_T = 0$.*

Proof. Look at the larger sum $\sum_{\sigma \in S_{A\cup B}} \operatorname{sign}(\sigma)\sigma\{T\}$. Considering the transposition of two elements a, b that lie in the same row of $\sigma\{T\}$, we see that $(ab)\sigma\{T\} = \sigma\{T\}$, but $(ab)\sigma$ and σ have opposite signs and hence these two terms cancel out, making the sum equal 0. Each permutation in $S_{A\cup B}$ can be identified as the product of one of the permutations $\pi \in S_{A\cup B}$ that preserve the increasing order of columns (those that we include in the Garnir element), and a column preserving permutation in C_T. But now,

$$0 = \sum_{\sigma \in S_{A\cup B}} \operatorname{sign}(\sigma)\sigma\{T\} = \sum_{\pi} \sum_{\tau \in C_T} \operatorname{sign}(\pi\tau)\pi\tau\{T\} = \left(\sum_{\pi} \operatorname{sign}(\pi)\pi\right)\left(\sum_{\tau \in C_T} \operatorname{sign}(\tau)\tau\{T\}\right) = g_{A,B} e_T$$

□

Thus, $g_{A,B}e_T$ let us write e_T as the sum of other polytabloids in this expansion. We must now quantify what it means for a tableau to be closer to standard.

Definition. Consider the *column tabloid* $[t]$ corresponding to t, where the ordering of elements within *columns* is ignored. For a column tabloid $[t]$ of shape λ, let λ_t^i be the composition given by (# of elements $\leq i$ in the first column, # of elements $\leq i$ in the second column, ...). Say $[s]$ dominates $[t]$. written $[s] \trianglerighteq [t]$, if $\lambda_t^i \trianglerighteq \lambda_s^i$ for all i.

Theorem 3. *The polytabloids associated with standard young tableau span V_λ.*

Proof. To see the base case of the induction, we must prove that that the only maximal elements in this poset[2] are the $[t]$ corresponding to standard t. Consider the standard tableaux T_0, formed by numbering 1 to n consecutively, going down the columns, for example $\begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 4 \\ \cline{1-2}\end{array}$. Luckily, T_0 is clearly a maximum element, so it dominates all other $[t]$, and hence nothing else can be maximal.

by our above results,

$$e_T = \left(\sum_{\pi \neq 1} \text{sign}(\pi)\pi\right) e_T = \sum_{\pi} \text{sign}(\pi) e_{\pi T}$$

over all π used to construct the Garnir element. For the inductive hypothesis, assume that if $[s] \triangleright [t]$, then e_s can be written in the span of standard polytabloids. It suffices to prove that for each non-identity π included in the construction of $g_{A,B}$, $[\pi t] \triangleright [e_T]$. If $[t]$ is the class of a standard tableau, we have nothing to show, hence assume then that we have a row descent.

In rows (say) j and $j+1$ that we had the row descent, the picture is like 5, and in particular the elements in column j are the largest elements of $A \cup B$. This means that after the action of π, the numbers of column j were smaller than they were before. As a result, for each choice of i, there are at least as many elements $\leq i$ in column j than there were before. Since there have been no changes to any columns except j and $j+1$, this means that $[\pi t] \triangleright [t]$. $\square$

Theorem 4 (Young's Rule)**.** *The permutation modules decompose as*

$$U_\mu = \bigoplus_{\lambda \geq \mu} K_{\lambda\mu} V_\lambda$$

This can be proven in a way that is very similar to the above, but we defer a proof to section 8 where we will use a different idea. The key idea of the first is to find a basis set for $\text{Hom}(U_\mu, V_\lambda)$ with elements that corresponds to semistandard tableau with shape λ and content μ. The same trick of considering the maximal element in some ordering to show independence, and writing each homomorphism as a combination of ones that are closer to the desired will still work. We can already see here that only the terms with $\lambda \geq \mu$ can be nonzero, for

$$\text{Hom}(U_\mu, V_\lambda) = \text{Hom}(\mathbb{C}[S_n]a_\mu, \mathbb{C}[S_n]a_\lambda b_\lambda) = a_\mu \mathbb{C}[S_n]a_\lambda b_\lambda,$$

and this is zero for $\mu > \lambda$ by 3.

6 Symmetric Functions

The representations of S_n are closely related to the ring of *symmetric functions*. We will describe the relationship in the next section, but here, we provide some background about this ring.

Definition. The *graded ring* Λ is the ring of formal power series in $\mathbb{C}$, with an *infinite* set of variables $\mathbf{x} = \{x_1, x_2, \dots\}$. Additionally, we require that under any permutation of variables, the polynomials in Λ are invariant, and the degrees of the monomials in any symmetric function are bounded. The ring is graded into Λ^n, the elements of degree n.

[2]In a poset, maximal elements have no element greater than them, but a maximum is greater than every element.

Remark. Λ is formally defined with polynomials on an infinite set of variables, but we typically won't have to think about constructions like this. When we work over Λ^n to look at S_n, all our polynomials can be thought of as being over n variables $x_1, x_2, \dots x_n$, by simply forgetting about all terms with any other variables.

Realistically, we only care about the first condition in this definition, although the second prohibits certain infinite but symmetric polynomials. Since the ring is in fact a $\mathbb{C}$-algebra, we can ask about a basis for Λ. Corresponding to any partition $\lambda = (\lambda_1, \dots, \lambda_\ell)$, define the *monomial function* m_λ as the sum of *all* monomials that have exponents $\lambda_1, \lambda_2, \dots$ as their exponents, in any order. For example,

$$m_{(2,1)} = x_1^2x_2 + x_2^2x_1 + x_1^2x_3 + x_3^2x_1 \dots$$

Essentially, this means that whenever we want to include a monomial with this pattern of exponents, we are required to include m_λ in our symmetric function. Denote $x_1^{\lambda_1}x_2^{\lambda_2}\dots$ as $\mathbf{x}^\lambda$, so that the coefficient of this is the same as the coefficient of m_λ.Thus, the m_λ form a basis; over Λ^n, we can take m_λ over partitions λ of n as a basis. Let's rapid-fire some other bases for Λ^n:

- The power sum symmetric functions p. We define $p_{(k)} = m_{(k)} = x_1^k + x_2^k \dots$, and let $p_\lambda \coloneqq \prod p_{(\lambda_i)}$.
- The elementary symmetric functions e(which are familiar from Vieta's Formulas). Define $e_{(k)} = m_{(1,1,\dots 1)}$ as the sum of all monomial terms of degree k with no exponents higher than 1, and as above $e_\lambda \coloneqq \prod e_{(\lambda_i)}$.
- The complete homogenous symmetric functions h. Similar to the above, let $h_{(k)}$ be the sum of *all* monomial terms of degree k, and $h_\lambda \coloneqq \prod h_{\lambda_i}$.

For all these symmetric polynomials, we will allow a slight abuse of notation in writing(for example) p_π for $\pi \in S_n$, when we really mean $p_{\text{the conjugacy class of }\pi}$.

However, we have not yet encountered the most important basis for Λ^n. For any Young tableaux T, define $\mathbf{x}^T$ as $x_1^{\#\text{ of ones}}x_2^{\#\text{ of twos}}\dots$. With this,

Definition. The Schur Polynomial s_λ is defined as the sum over all semi-standard tableaux T of shape λ of $\mathbf{x}^T$, i.e

$$\sum_{T\text{ is a SSYT with shape }\lambda} \mathbf{x}^T$$

Using the Kostka Numbers from earlier, we know that the coefficient of $x_1^{\mu_1}x_2^{\mu_2}\dots$ will be exactly $K_{\lambda\mu}$. However, it's not obvious that s_λ is even symmetric at all. In order to show that $K_{\lambda\mu} = K_{\lambda\mu'}$ whenever μ' is a permutation of μ. Yet by Young's Rule, this means we want to show that the multiplicity of V_λ in U_μ and $U_{\mu'}$ is the same, but this is obvious; U_μ and $U_{\mu'}$ are isomorphic representations. Because of the definition of the monomial functions, we clearly have

$$s_\lambda = \sum_{\mu\text{ is a partition of n}} K_{\lambda\mu}m_\mu. \tag{2}$$

This also shows the s_λ form a basis, using the following trick. Because we know that $K_{\lambda\mu}$ vanishes for $\mu > \lambda$, when we write the s_λ as a matrix in terms of the basis m_μ, this matrix is triangular, and so must have linearly independent columns.

In fact, the same triangular trick works to show that the p and the e are also bases. For the p_λ, recall that the coefficient for $x_1^{\mu_1}x_2^{\mu_2}\dots$ will be the coefficient of m_μ. Therefore, it suffices to show that $\prod_{\lambda_i}(x_1^{\lambda_1} + x_2^{\lambda_1}\dots)[x_1^{\mu_1}\dots] = 0$ whenever $\mu < \lambda$(so here our matrix is upper instead of lower triangular). To see this all we need to do is note that the largest term of μ must include at least λ_1(since λ_1 must be selected in the first $p_{(\lambda_1)}$ term), and so on. A similar argument suffices to show that the e form a basis, a result known as the *Fundamental Theorem of Symmetric Functions.* To see the h form a basis, define the generating functions $H(t) = \sum_i h_it^i$, and $E(t) = \sum_i e_it^i$.It's clear that since all variables of e have exponents 0 and 1,

$$E(t) = \prod_k (1 + x_kt), \tag{3}$$

and similarly we get

$$H(t) = \prod_k (1 + x_k t + x_k^2 t^2 + \dots) = \prod_k \frac{1}{1 - x_k t}. \tag{4}$$

As a result $H(t)E(-t) = 1$, so matching coefficients of t^n gives $\sum_{r=0}^{n} (-1)^r h_{n-r} e_r = 0$ for $n \geq 1$. We can use this to write $e_n = h_1 e_{n-1} - h_2 e_{n-2} + h_3 e_{n-3} \dots$, and by induction this means that the e can be written in terms of the h_λ, as desired.

Definition. The *Hall inner product* is defined on Λ^n by requiring $\langle h_\lambda, m_\mu \rangle = \delta_{\lambda\mu}$, i.e h_λ and m_μ are orthogonal when $\lambda \neq \mu$.

Proposition. *The Schur Polynomials are an orthonormal basis under this inner product.*

Proof. The main idea here is to use *Cauchy's Identity*,

$$\prod_{i,j} \frac{1}{1 - x_i y_j} = s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}).$$

Note that here we consider *two* infinite sets of variables $\mathbf{x}$ and $\mathbf{y}$. We will not prove it here, but there is a combinatorial proof based on the RSK algorithm (that we use later to prove 8) presented in [5]. Since $\sum_i h_i(x) y_j^i = \prod_i \frac{1}{1 - x_i y_j}$ by (4),

$$s_\lambda(\mathbf{x}) s_\lambda(\mathbf{y}) = \prod_{i,j} \frac{1}{1 - x_i y_j} = \prod_j \sum_i h_i(\mathbf{x}) y_j^i = \sum_\lambda h_\lambda(\mathbf{x}) m_\lambda(\mathbf{y})$$

Now if we write $s_\lambda = \sum_\gamma a_{\lambda\gamma} h_\gamma = \sum b_{\gamma\lambda} m_\gamma$, expanding the above identity and equating coefficients tells us $\sum_\gamma b_{\rho\lambda} a_{\lambda\gamma} = \delta\rho, \gamma$. But expanding the inner product $\langle \sum a_{\lambda\gamma} h_\gamma, \sum b_{\gamma\lambda} m_\gamma \rangle$, we obtain an equivalent expression. $\square$

7 Representations and Symmetric Functions

Lemma. *Let μ be a fixed partition of n.*

$$p_\mu = \sum_{\lambda \vdash n} \chi_{U_\lambda}(\mu) m_\lambda.$$

Proof. Expand p into the product $\prod_i (x_1^{\mu_1} + x_2^{\mu_2} \dots)$. Because U_λ is a permutation module, the character on μ is the number of tabloids fixed by an element with type μ. Such a tabloid must have each cycle acting independently on a row. Thus, in order to construct such a tabloid, we need to select which row each cycle μ_i will go into. But this is exactly enumerated by the coefficients of the LHS; choosing a term $x_k^{\mu_j}$ corresponds to placing μ_j entirely in the kth row. With this interpretation, the coefficient of $\mathbf{x}^\lambda$ corresponds with such tabloids that have λ_1 boxes in the first row, λ_2 in the second, etc. That is, $p_\mu[\mathbf{x}^\lambda] = \chi_{U^\lambda}(\mu)$, which is what we wanted to show. $\square$

Theorem (Frobenius Character Formula)**.** *The character table χ_{V_λ} is the change of basis from the Schur polynomials to the power sum polynomials. That is,*

$$s_\lambda = \frac{1}{n!} \sum_\pi p_\pi \chi_{V_\lambda}(\pi)$$

Proof. We will leverage the previous identity to expand p_π. By manipulating the sum, we get

$$\begin{aligned}
&\frac{1}{n!}\sum_{\pi\in S_n} p_\pi \chi_{V_\lambda}(\pi) \\
&= \frac{1}{n!}\sum_{\pi\in S_n}\left(\sum_{\mu\in S_n}\chi_{U_\mu}(\pi)m_\mu\right)\chi_{V_\lambda}(\pi) \\
&= \frac{1}{n!}\sum_{\pi\in S_n}\sum_{\mu\in S_n}\chi_{U_\mu}(\pi)\chi_{V_\lambda}(\pi)m_\mu \\
&= \sum_{\mu\in S_n} m_\mu \frac{1}{n!}\sum_{\pi\in S_n}\chi_{U_\mu}(\pi)\chi_{V_\lambda}(\pi) \\
&= \sum_{\mu\in S_n} m_\mu \left\langle \chi_{U_\mu}, \chi_{V_\lambda}\right\rangle \\
&= \sum_{\mu\in S_n} m_\mu \dim \operatorname{Hom}(U_\mu, V_\lambda) \\
&= \sum_{\mu\in S_n} m_\mu K_{\lambda\mu} \\
&= s_\lambda.
\end{aligned}$$

Above, the equality $\dim \operatorname{Hom}(U_\mu, V_\lambda) = K_{\lambda\mu}$ follows from Young's Rule, $U_\mu = \bigoplus_{\lambda>\mu} V_\lambda^{\oplus K_{\lambda\mu}}$. We have also used the fact that all the representations of S_n are entirely real, and have real character values, so we don't have to take the conjugate to compute the inner product of characters. □

Remark. The above theorem is often stated(see [2], [1]) in the more explicit(but less intuitive) formulation

$$\chi_{V_\lambda} = \Delta(\mathbf{x})p_\lambda[x_1^{\lambda_1+k-1}x_2^{\lambda_2+k-2}\dots],$$

where λ is a partition into k parts, and Δ is the discriminant, given by $\prod_{i<j}(x_i - x_j)$.

We can already see many connections between the symmetric functions and the representations of S_n. Based on the form of the above two results, it is reasonable to assume that the Specht Modules somehow correspond to the Schur Polynomials. We can actually formalize this idea using the following map.

Definition. The *Characteristic Map* is the the following map ch from the space of class functions on S_n, to $\Lambda^n(\mathbf{x})$ defined as

$$\operatorname{ch}(\chi) = \frac{1}{n!}\sum_{\pi\in S_n}\chi(\pi)p_\pi,$$

for χ in Class(G).

Of course, we can also regard ch as a map from the representations of S_n by taking the character. With this definition in mind, the Frobenius Character Formula just states $\operatorname{ch}(\chi_{V_\lambda}) = s_\lambda$. So, what's good about this particular definition for the characteristic map?

- It preserves addition: $\operatorname{ch}(f+g) = \operatorname{ch}(f) + \operatorname{ch}(g)$, which is clear from the definition.
- It preserves the inner products of the two spaces: $\langle f, g\rangle = \langle \operatorname{ch}(f), \operatorname{ch}(g)\rangle$, since it takes an orthonormal basis of irreducible characters to another orthonormal basis of Schur Functions. Thus, it is an *isometry* of vector spaces.

We can get even more if we extend ch to act on class functions of *any* symmetric group. Just as Λ is the graded ring containing all Λ^n, where multiplying something in Λ^n with something in Λ^m gets you a function in Λ^{m+n}, we'd like to define some kind of multiplication of two representations(equivalently characters) of S_n and S_m respectively. Simply taking the product of f and g gives us a class function $f\times g$ from $S_n \times S_m$ to $\mathbb{C}$. But we are not lost yet, for we can take the character that this representation induces in S_{n+m}(since $S_n \times S_m$) is a subgroup. We will denote this product with $f \cdot g$. With all this structure, we have

Theorem. *The characteristic map is an isomorphism of rings, in addition to being an isometry of vector spaces.*

Proof. We just have to show that with the above definition of fg, $\mathrm{ch}(f \cdot g) = \mathrm{ch}(f)\,\mathrm{ch}(g)$. The key ingredient will be Frobenius Reciprocity, which states $\langle \chi, \psi\downarrow_H \rangle_H = \langle \chi\uparrow^G, \psi \rangle_G$. In this case, we will need it in a more general setting, where our functions are class functions from S_{mn} to Λ. Denote by ρ the function $\rho(\pi) = p_\pi$, i.e ρ takes an element of any symmetric group and outputs the corresponding power sum polynomial. Then $\mathrm{ch}(\chi) = \langle \chi, \rho \rangle$. We have

$$\begin{aligned}\mathrm{ch}(f \cdot g) &= \langle (f \times g)\uparrow^{S_{mn}}, \rho \rangle_{S_{nm}} \\ &= \langle f \times g, \rho\downarrow_{S_m \times S_n} \rangle_{S_n \times S_m} \\ &= \frac{1}{n!m!} \sum_{(\sigma,\pi) \in S_m \times S_n} (f \times g)(\sigma, \pi)\rho\downarrow((\sigma, \pi))\end{aligned}$$

Now recall that since ρ is restricted to $S_m \times S_n$, it's value on (σ, π) is equal to it's value on the corresponding term of S_{mn}, i.e, $\rho((\sigma, \pi)) = p_\sigma p_\pi$. Substituting this in:

$$\begin{aligned}&\frac{1}{n!m!} \sum_{(\sigma,\pi) \in S_n \times S_m} (f \times g)(\sigma, \pi)\rho\downarrow((\sigma\pi)) \\ &= \frac{1}{n!m!} \sum_{(\sigma,\pi)} f(\sigma)g(\pi)p_\sigma p_\pi \\ &= \left(\frac{1}{n!} \sum_{\sigma} f(\sigma)p_\sigma\right)\left(\frac{1}{m!} \sum_{\pi} g(\pi)p_\pi\right) = \mathrm{ch}(f)\,\mathrm{ch}(g).\end{aligned}$$

□

Let us try and compute that symmetric polynomial corresponding to the permutation modules(or its character) of S_n. Let $\mathrm{ch}(\chi_{U_\lambda})$ be the polynomial f_λ. By applying ch to Young's Rule, we see

$$f_\mu = \sum_{\lambda \geq \mu} K_{\lambda\mu} s_\lambda.$$

We can show that the polynomial f must actually be the homogenous polynomial h by verifying that h satisfies the above. To do this, consider the inner product $\langle h_\mu, s_\lambda \rangle$, which we want to show is $K_{\lambda\mu}$. But expanding s_λ, we see

$$\langle h_\mu, s_\lambda \rangle = \left\langle h_\mu, \sum_\nu K_{\lambda\nu} m_\nu \right\rangle = \sum_\nu K_{\lambda\nu} \langle h_\mu, m_\nu \rangle = K_{\lambda\mu}$$

as we wanted. Thus U_λ corresponds to h_λ.

8 More

In this section we collect some beautiful combinatorial results related to the material we have developed, but we omit most of the proofs because they are too long for this paper. A full treatment of these results is proved in [4] chapters 2, 3, 4.

The most famous combinatorial result that is related to representation theory is the *Hook Length Formula*. For a box i, the hook length is the number of boxes directly below or directly to the right of i, including i itself.

Theorem (Hook Length)**.** *The number of ways to fill a standard Young tableaux of λ is given by*

$$\frac{n!}{\prod_{boxes\ i} h_i!}$$

7	4	3	1
5	2	1	
2			
1			

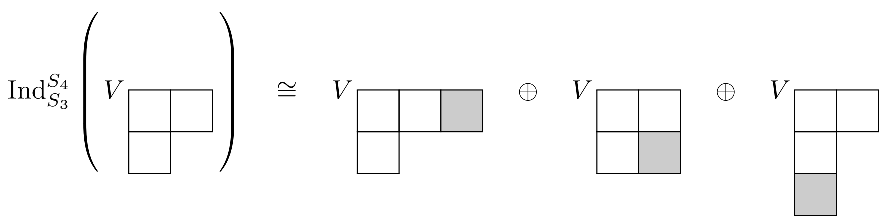
Above we see a young diagram where each box has been labeled with its hook length, and this means there are

$$\frac{9!}{7!4!3!1!5!2!1!2!1!} = 216$$

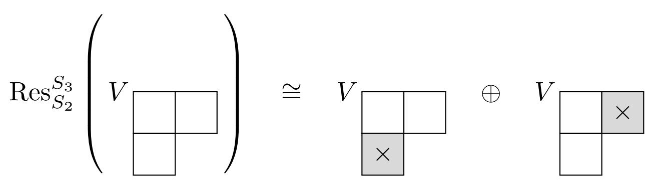
ways to make a standard Young tableau on $\lambda = (4, 3, 1, 1)$.

The hook length formula can be proved starting from the Frobenius Character Formula, because the dimension of the representation V_λ(which by an earlier theorem is the number of ways to fill the young frame) is equal to $\chi_{V_\lambda}(1)$, and the rest of the proof is mostly algebra to get in the desired form but there is also a more elementary probabilistic proof presented in [3]. Our next result is a combinatorial characterization of the induced representations.

Theorem (Branching Rule)**.** *The induced representation $\mathrm{Ind}_{S_n}^{S_{n+1}} V_\lambda$ is the direct sum of V_μ, over all μ that can be obtained from adding a single box to λ.*



Similarly, the restricted representation $\mathrm{Res}_{S_{n-1}}^{S_n} V_\lambda$ is the direct sum of V_μ over all μ that can be obtained from deleting a single box from λ. As an example,



To prove this, we appeal to symmetric functions once again. The well-known lemma we need is

Lemma 8 (Pieri's Formula)**.** *We have*

$$s_\lambda s_{(r)} = s_\lambda(x_1 + x_2 + \dots) = \sum_\mu S_\mu, \tag{5}$$

over all Young diagrams μ that can be obtained from adding r boxes to λ, with no two boxes in the same column.

Proof. The main idea is the following. This takes any semistandard tableau T and inserts an element i to create a new tableau T' that still must be semistandard. Now looking at $s_{(r)}$, we can see that this is the same as the homogenous polynomial $h_{(r)}$, i.e, $s_{(r)}$ contains all monomial terms of degree r, which correspond to weakly increasing lists of r integers. To prove the desired result, we want to find a bijection that takes in a semistandard tableau T and a list of r integers, and outputs a new semistandard tableau T' that contains these new r integers added to T. But this can be done by simply row-inserting all the elements in our list of r integers to add using the above algorithm.

Algorithm 1 Knuth-Robinson-Schensted Row Insertion

Set the insert row R as the first row.
while Some element of R is more than i **do**
  Replace the smallest element of R greater than i with i.
  Set i to be the replaced element.
  Set R as the next row down.
end while
At this point i is greater than every element of R, so place it at the end of R.

The only part left to show is that by doing this, we can never have added two boxes to the same column. Suppose that b_0, b_1 have been inserted into the same column.The first such occurence must occur at the ends of the rows, but then since $b_1 > b_0$, b_1 is greater than or equal to the whole of b_0's row, so we would have added it to the end of b_0's row, contradiction. □

Proof of Branching Rule. Consider the special case $r = 1$ of Pieri's Formula. We see that applying ch^{-1} gives us

$$\chi_{V_\lambda} \cdot \mathbf{1} = \sum_\mu \chi_{V_\mu}.$$

Translating this to its statement in terms of representations gives

$$\mathrm{Ind}_{S_n \times S_1}^{S_{n+1}} V_\lambda \boxtimes \mathrm{triv} = \mathrm{Ind}_{S_n}^{S_{n+1}} V_\lambda = \bigoplus_\mu V_\mu,$$

which is the desired result. The other form in terms of the reduced representation follows directly by utilizing Frobenius Reciprocity. □

We are finally ready for the proof of Young's Rule.

Proof of 4 Young's Rule. By Frobenius Reciprocity, we have

$$\mathrm{Hom}(U_\mu, V_\lambda) = \mathrm{Hom}(\mathrm{Ind}_{P_\mu}^{S_n} \mathrm{triv}, V_\lambda) = \mathrm{Hom}_{P_\mu}(\mathrm{triv}, \mathrm{Res}_{P_\mu} V_\lambda).$$

That is, the space of vectors in $\mathrm{Res}\, V_\lambda$ that are fixed by P_λ. Our approach will be to take this one step at a time. Since $P_\mu = S_{\mu_1} \times S_{\mu_2} \cdots \times S_{\mu_k}$, we can first find the vectors in V_λ that are invariant under the action of S_{μ_k}, then take this space and find the vectors invariant under $S_{\mu_{k-1}}$, etc. For the first step, this means we want to find a form for $\mathrm{Hom}_{S_{\mu_k}}(\mathrm{triv}, \mathrm{Res}_{S_{n-\mu_k} \times S_{\mu_k}} V_\lambda)$. But if $\mu_1 = r$, converting Pieri's Rule using ch tells us that $\mathrm{Ind}^{S_n} V_\lambda \boxtimes \mathrm{triv}_{S_r}$ decomposes as a sum of S_ν for ν formed by adding r boxes to λ not in the same column. By Frobenius Reciprocity for such ν,

$$1 = \dim \mathrm{Hom}_{S_n}(\mathrm{Ind}\, V_\lambda \boxtimes \mathrm{triv}_{S_r}, V_\nu) = \dim \mathrm{Hom}_{S_{n-r} \times S_r}(V_\lambda \boxtimes \mathrm{triv}, \mathrm{Res}\, V_\nu),$$

which effectively means that when we restrict a Specht Module V_λ to $S_{n-r} \times S_r$, it becomes a direct sum of the Specht modules $S_\nu \boxtimes \mathrm{triv}_{S_r}$ corresponding to Young Diagrams formed by removing r boxes.

Because Hom distributes over a direct sum, we see

$$\mathrm{Hom}_{S_r}(\mathrm{triv}, \mathrm{Res}_{S_{n-r} \times S_r} V_\lambda) = \mathrm{Hom}_{S_r}\left(\mathrm{triv}, \bigoplus_{\nu \searrow \lambda} (V_\nu \boxtimes \mathrm{triv}_{S_r})\right) = \bigoplus_\nu \mathrm{Hom}_{S_r}(\mathrm{triv}, V_\nu \boxtimes \mathrm{triv}),$$

but any homomorphism from $\mathbb{C}$ to the tensor product $V_\nu \otimes \mathbb{C}$ is uniquely determined by the image of 1, this is just isomorphic to V_ν! Repeating this argument, we are removing μ_k, then μ_{k-1}, then $\mu_{k-2}, \ldots$ boxes from λ, no two in the same column, and each such decomposition corresponds to a one-dimensional component of this $\mathrm{Hom}(U_\mu, V_\lambda)$. But each such decomposition corresponds to a SSYT of content μ, by filling the first μ_k boxes removed with the number k, the next μ_{k-1} with the number $k - 1$, and so on. □

To generalize the idea in the proof of the Branching Rule, consider the expansion of a product of Schur Polynomials,

$$s_\mu s_\nu = \sum_\lambda c^\lambda_{\mu\nu} s_\lambda$$

for some arbitrary partitions $\mu \vdash m$, $\nu \vdash n$. We call these $c^\lambda_{\mu\nu}$ the *Littlewood-Richardson coefficients*. Equivalently, by using ch^{-1} on this,

$$\mathrm{Ind}^{S_{mn}}_{S_m \times S_n} V_\mu \otimes V_\nu = \bigoplus_\lambda c^\lambda_{\mu\nu} V_\lambda$$

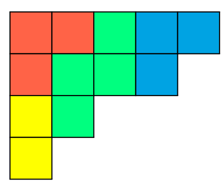
This following theorem is an even wider generalization of the Branching Rule, and provides an interpretation for these numbers.

Theorem (Littlewood-Richardson Rule)**.** *A* skew *tableau of shape λ/μ is the tableau obtained from taking the shape λ and "cutting out" the shape of μ. A semistandard such tableau is defined in the same way as f before, and so is the content. Then $c^\lambda_{\mu\nu}$ is the number of semistandard skew tableau T with shape λ/μ and content μ.*

A *rim hook* is an ordinary hook in a Young tableaux, but projected onto the boundary. For such a rim hook ζ, we let the *leg lenth* $\ell(\zeta)$ be the number of rows minus one. These rim hooks will provide a more combinatorial character formula.

Theorem (Murnaghan-Nakayama)**.** *Let λ be a partition of n. To evaluate the character on g, where g has cycle type μ_1, μ_2, $\ldots \mu_p$, consider all ways s to decompose the Young Tableaux into p rim hooks of lengths μ_1, μ_2, $\ldots \mu_p$. Let $\ell(s)$ be the sum of $\ell(\zeta)$ for each hook ζ in the decomposition. Then*

$$\chi_{V_\lambda}(g) = \sum_s (-1)^{\ell(s)}$$



As an example, the above diagram has $\lambda = (5, 4, 2, 1)$, and $\mu = (3, 4, 2, 3)$. The Young Diagram of λ can be decomposed by first removing the blue hook, then the green, then the yellow, and then finally the red. Here, $\ell(s) = 5$.

References

[1] Pavel Etingof et al. *Introduction to representation theory. With historical interludes by Slava Gerovitch.* English. Vol. 59. Stud. Math. Libr. Providence, RI: American Mathematical Society (AMS), 2011. ISBN: 978-0-8218-5351-1.

[2] William Fulton and Joe Harris. *Representation theory. A first course.* English. Vol. 129. Grad. Texts Math. New York etc.: Springer-Verlag, 1991. ISBN: 0-387-97495-4; 0-387-97527-6.

[3] Curtis Greene, Albert Nijenhuis, and Herbert S. Wilf. "A probabilistic proof of a formula for the number of Young tableaux of a given shape". English. In: *Adv. Math.* 31 (1979), pp. 104–109. ISSN: 0001-8708. DOI: 10.1016/0001-8708(79)90023-9.

[4] Bruce E. Sagan. *The symmetric group. Representations, combinatorial algorithms, and symmetric functions.* English. 2nd ed. Vol. 203. Grad. Texts Math. New York, NY: Springer, 2001. ISBN: 0-387-95067-2.

[5] Richard P. Stanley. *Enumerative combinatorics. Volume 2. With an appendix by Sergey Fomin.* English. 2nd edition. Vol. 208. Camb. Stud. Adv. Math. Cambridge: Cambridge University Press, 2024. ISBN: 978-1-00-926248-4; 978-1-00-926249-1; 978-1-00-926253-8. DOI: 10.1017/9781009262538.