REPRESENTATION THEORY OF S_n

EVAN JENKINS

ABSTRACT. These are notes from three lectures given in MATH 26700, Introduction to Representation Theory of Finite Groups, at the University of Chicago in November 2009. Most of the material comes from chapter 7 of [Ful97], some of it verbatim.

1. Introduction

Most of this course has focused on the general theory for compact groups, but our examples have focused on a particular class of compact groups, namely, the compact Lie groups. The usual constructions of representations of compact Lie groups relies heavily upon special structures that exist in the world of Lie groups, namely maximal tori and Borel subgroups.

Finite groups, while in some ways much simpler than compact Lie groups, lack a nice unified structure theory. As a result, we can't expect that there will be any very powerful general methods for constructing representations of finite groups.

The symmetric groups, however, were one of the first classes of groups studied via the lens of representation theory. The character theory of S_n for arbitrary n was worked out by Frobenius in 1900. Since then, many explicit constructions of the representations of S_n were developed, some of them very closely connected to the study of Lie groups.

2. Representation Theory of Finite Groups

All of our results for compact groups hold in particular for finite groups, which can be thought of as compact groups with the discrete topology. Here, the description of Haar measure is very simple; it is just the normalized counting measure. Explicitly, if G is a group, and f function on G, then

$$\int_G f(g) \, dg = \frac{1}{|G|} \sum_{g \in G} f(g).$$

We'll denote by $\mathbb{C}G$ the vector space of complex-valued functions on G. Note that $\dim \mathbb{C}G = |G|$. This vector space appears in many different guises, so it's important to be able to distinguish how we're thinking about it at any given time. First, we note that we can make $\mathbb{C}G$ into an algebra in two natural ways, namely, by pointwise multiplication of functions or by convolution. In fact, even the convolution product has two conventions, and there are good reasons to use either one, but we'll take our convolution product to be

$$(f_1 * f_2)(g) = |G| \int_G f_1(h) f_2(h^{-1}g) dh = \sum_{h \in G} f_1(h) f_2(h^{-1}g).$$

Under convolution, this ring is called the **group algebra** of G and will play a very important role in the representation theory of finite groups.

There are three natural actions of G on $\mathbb{C}G$, the left and right regular representations L_G and R_G and the conjugation action A_G . These are defined by

$$L_G(g)f(x) = f(g^{-1}x),$$

$$R_G(g)f(x) = f(xg),$$

$$A_G(g)f(x) = f(g^{-1}xg).$$

An element of $\mathbb{C}G$ that is invariant under the conjugation action is called a **class** function. The class functions form a subspace of $\mathbb{C}G$, which will be denoted by $\mathbb{C}G^G$.

Exercise 2.1. Prove that $\mathbb{C}G^G$ is a subalgebra of $\mathbb{C}G$ under both the pointwise product and the convolution product.

Let's now see how our results from the general compact case look in this setting. Recall that if (T, V) is a representation of G, its character $\chi_V \in \mathbb{C}G^G$ is defined as $\chi_V(g) = \operatorname{tr} T(g)$. We can give $\mathbb{C}G$ (and hence $\mathbb{C}G^G$) the L^2 inner product,

$$\langle f_1, f_2 \rangle = \int_G f_1(g) \overline{f_2(g)} \, dg.$$

Under this inner product, we proved the following orthogonality relation for characters.

Theorem 2.2. Let T_1 and T_2 be irreducible representations with characters χ_1 and χ_2 . Then

$$\langle \chi_1, \chi_2 \rangle = \begin{cases} 1 & \text{if } T_1 \sim T_2, \\ 0 & \text{if } T_1 \not\sim T_2. \end{cases}$$

Another big result of the compact theory is the Peter-Weyl theorem.

Theorem 2.3 (Peter-Weyl). The right regular representation $\mathbb{C}G$ decomposes as

$$\mathbb{C}G = \bigoplus_{V \ irred.} d_V V,$$

where d_V is the dimension of V.

Corollary 2.4.
$$\sum_{V irred.} d_V^2 = |G|$$
.

A consequence of Peter-Weyl that we haven't seen yet says that the characters of representations of G form a uniformly dense subspace of continuous class functions on G. In the finite setting, this gives the following result.

Theorem 2.5. $\mathbb{C}G^G$ is equal to the span of the characters of G. In particular, the characters of the irreducible representations form an orthonormal basis.

This result, combined with the following lemma, tells us that the number of (isomorphism classes of) irreducible representations of a finite group G is equal to the number of conjugacy classes of G.

Lemma 2.6. dim $\mathbb{C}G^G$ is equal to the number of conjugacy classes of G.

Proof. A class function is precisely a function on G that is constant on conjugacy classes. We can thus identify $\mathbb{C}G^G$ with the algebra of functions on the set of conjugacy classes of G. The dimension of this algebra is the number of conjugacy classes.

Next, we will develop some theory that is particular to the case of finite groups.

Theorem 2.7. A representation of G is the same as a (finite dimensional) left module over the group algebra $\mathbb{C}G$.

Proof. Recall that the group algebra refers to the convolution product structure on $\mathbb{C}G$. To see this, we first reinterpret the group algebra $\mathbb{C}G$. A function on G can be thought of as a formal \mathbb{C} -linear combination of elements of G via $(f:G\to\mathbb{C})\mapsto \sum_{g\in G}f(g)g$.

Exercise 2.8. If we view an element of $\mathbb{C}G$ as $\sum_{g \in G} a_g g$, where $a_g \in \mathbb{C}$, then the convolution product becomes

$$\left(\sum_{g \in G} a_g g\right) \left(\sum_{h \in G} b_h h\right) = \sum_{k \in G} \sum_{gh=k} a_g b_h k.$$

What does it mean to give a left $\mathbb{C}G$ -module? First, we note that since $\mathbb{C}G$ is a \mathbb{C} -algebra, a left $\mathbb{C}G$ -module is in particular a \mathbb{C} -vector space V, which we are assuming to be finite dimensional. To specify a $\mathbb{C}G$ -module structure on V, it is enough to say how a basis acts. We have a particularly nice basis, namely, the elements $g \in \mathbb{C}G$. Each g must act \mathbb{C} -linearly, as \mathbb{C} lies in the center of $\mathbb{C}G$. The element e acts trivially, as it is the identity for the convolution product. Furthermore, since $g \cdot (h \cdot v) = (gh) \cdot v$, it follows that the action of gh is the composite of the actions of g and h. Finally, the action of each g is invertible, because $gg^{-1} = e$ in $\mathbb{C}G$. We may thus conclude that giving a $\mathbb{C}G$ -module structure on V is the same as giving a homomorphism $G \to GL(V)$, i.e., a giving a representation of G on V.

What does this point of view buy us? Now we can apply the theory of rings and modules to the study of representations of finite groups. Recall that a module is **simple** if it has no proper nontrivial submodules. Thus, a $\mathbb{C}G$ -module is irreducible if and only if its corresponding module is simple.

3. The Symmetric Group

We know from the above discussion how many representations we need to look for, namely, one for every conjugacy class. However, in general, there is no known explicit one-to-one correspondence between the conjugacy classes and the irreducible representations. For S_n , however, we can take advantage of the fact that the conjugacy classes admit a very nice combinatorial description to construct irreducible representations corresponding to them.

Recall that any element of S_n can be written as a product of disjoint cycles, and two elements of S_n are conjugate if and only if they have the same number of cycles of any given length. Thus, a conjugacy class of S_n can be specified by giving an unordered partition of the number n.

We will write a partition of n as a tuple $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$, where $\sum_{i=1}^k \lambda_i = n$, and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > 0$. We can also express such a partition pictorially using

a Young diagram, which is a series of rows of boxes, in which the *i*th row contains λ_i boxes.

For example, the partition (5,4,1,1) of 11 can be represented by the following Young diagram.



Before we find a general construction for all n, let's first see how we can use the theory we've developed above to study the character theory of S_3 and S_4 .

Let's start with a few explicit examples of irreducible representations of S_n . In fact, aside from a couple of small-n exceptions, these are the smallest irreducible representations of S_n .

For every $n \geq 2$, there are two one-dimensional representations of S_n . One is the trivial representation $V_{\rm triv}$, in which every $\sigma \in S_n$ acts trivially. The other is the **sign representation** $V_{\rm sgn}$. Recall that every element $\sigma = S_n$ can be written as a product of transpositions, the number of which is unique modulo 2. We the define $\operatorname{sgn} \sigma \in \mathbb{Z}/2$ to be the number of transpositions. In the sign representation, σ acts as multiplication by $(-1)^{\operatorname{sgn} \sigma}$.

The third representation is a bit more complicated to define. Let V_{perm} be an n-dimensional vector space with chosen basis (e_1, \ldots, e_n) . Then S_n acts on V_{perm} by the **permutation representation**, in which $\sigma \in S_n$ acts by sending e_i to $e_{\sigma(i)}$. However, this representation is reducible, because the subspace spanned by $\sum_{i=1}^n e_i$ is invariant.

Proposition 3.1. The (n-1)-dimensional representation $V_{\text{stan}} = V_{\text{perm}}/\mathbb{C}(e_1 + \ldots + e_n)$ is irreducible.

Proof. Let $A \in \operatorname{End}_G(V)$. Then $A = (a_{ij})$ must commute with all permutation matrices. In other words, switching any two rows of A must be the same as switching any two columns. This implies that $a_{ii} = a_{jj}$ for all i and j, and that $a_{ij} = a_{k\ell}$ for all $i \neq j, k \neq \ell$. So the space of such A is two-dimensional, which implies that V is the direct sum of two irreducible representations. Thus, W must be irreducible. \square

We call this irreducible representation the standard representation.

The group S_3 has three conjugacy classes, corresponding to the Young diagrams



Thus, the three irreducible representations defined above must be the only three. We can now construct the **character table** of S_3 . The character table shows the value of each irreducible character on each conjugacy class.

S_3			
$V_{ m triv}$	1	1	1
$V_{ m sgn}$	1	-1	1
$V_{\rm stan}$	2	0	-1

Exercise 3.2. Show that $\chi_{\text{perm}}(\sigma)$ is equal to the number of fixed points of $\sigma \in S_n$ acting as permutations of n points.

Since characters are additive with respect to direct sums, we can now compute $\chi_{\text{stan}} = \chi_{\text{perm}} - \chi_{\text{triv}}$.

The character table of S_4 is a bit more complicated. We know three lines of the table.

S_4					
$\overline{V_{ m triv}}$	1	1	1	1	1
$V_{ m sgn}$	1	-1	1	-1	1
$V_{ m stan}$	3	1	0	-1	-1

We know that the sum of the squares of the dimensions of the representations must be 4! = 24. So far, we have $1^2 + 1^2 + 3^2 = 11$, so we are missing a representation of dimension 2 and one of dimension 3.

We can get one more representation of dimension 3 by taking $V_{\text{stan}} \otimes V_{\text{sgn}}$. Since characters are multiplicative with respect to tensor product, we get the following.

S_4					
$V_{ m triv}$	1	1	1	1	1
$V_{ m sgn}$	1	-1	1	-1	1
$V_{ m stan}$	3	1	0	-1	-1
$V_{ m stan} \otimes V_{ m sgn}$	3	-1	0	1	-1

To get the last irreducible character χ , we can use orthogonality. We know it must be two-dimensional, so $\chi(1)=2$.

S_4					
$\overline{V_{ m triv}}$	1	1	1	1	1
$V_{ m sgn}$	1	-1	1	-1	1
$V_{ m stan}$	3	1	0	-1	-1
$V_{ m stan} \otimes V_{ m sgn}$	3	-1	0	1	-1
$V_{?}$	2	a	b	c	d

We now have four linear equations $\langle \chi, \chi_i \rangle = 0$ in four unknowns a, b, c, d, so we can hopefully solve for the character. For example, the equation $\langle \chi, \chi_{\text{triv}} \rangle = 0$ says that

$$2 + 6 \cdot a + 8 \cdot b + 6 \cdot c + 3 \cdot d = 0.$$

Note the multiplicities; we must multiply the value at each conjugacy class by the number of elements in the conjugacy class.

Exercise 3.3. Compute the last irreducible character of S_4 .

4. Specht Modules

Our general approach to the representation theory of S_n will be combinatorial. Recall our construction of the standard representation. We got this representation by looking at the representation corresponding to a certain group action of S_n (namely, its usual action as permutations of n elements). This permutation representation was not irreducible, but it had an interesting irreducible representation sitting inside of it.

Our general approach will follow along the same lines. Namely, we will find some combinatorial objects on which S_n acts, construct a corresponding representation, and then find an interesting irreducible representation sitting inside of it.

Let λ be a partition of n. Recall that we can think of λ as a Young diagram. A **numbering** of the Young diagram λ is a labeling of the boxes by the numbers $1, 2, \ldots, n$. For instance, we can number the diagram \longrightarrow by $\frac{2|5|1}{4|3|}$.

There are, of course, n! such numberings. There is an action of S_n on the set of numberings by permuting the numbers.

A **tabloid** is an equivalence class of numberings of λ , where we consider two numberings T and T' equivalent if the entries of each row of T agree with those in the corresponding row of T'. For instance, we have

We denote by $\{T\}$ the tabloid corresponding to T. There is an action of S_n on the set of tabloids of shape λ by $\sigma \cdot \{T\} = \{\sigma \cdot T\}$.

We now define M^{λ} to be the complex vector space with basis the tabloids of shape λ , and we let S_n act by permuting the basis elements. We note that this is a generalization of V_{perm} , because if we take tabloids of shape (n-1,1), then there are exactly n of them, one for each number that can appear in the second row, and the action of S_n is the usual action on the set $\{1, 2, \ldots, n\}$.

For any numbering T, we define C(T) to be the subgroup of S_n that leaves the columns of T invariant. As an abstract group, $C(T) \cong S_{\mu_1} \times S_{\mu_2} \times \cdots \times S_{\mu_\ell}$, where $\mu = (\mu_1, \mu_2, \dots, \mu_\ell)$ is the partition we get by flipping the Young diagram for λ over the diagonal. This partition μ is called the **conjugate partition** to λ .

We can similarly define R(T) to be the subgroup of S_n that leaves the rows of T invariant. As before, we have $R(T) \cong S_{\lambda_1} \times \cdots \times S_{\lambda_k}$.

Let

$$v_T = \sum_{q \in C(T)} (-1)^{\operatorname{sgn} q} \{ q \cdot T \}.$$

Proposition 4.1. For any $\sigma \in S_n$, $\sigma \cdot v_T = v_{\sigma \cdot T}$.

Proof. First, we note that $C(\sigma \cdot T) = \sigma C(T)\sigma^{-1}$. We have

$$\sigma \cdot v_T = \sigma \cdot \sum_{q \in C(T)} (-1)^{\operatorname{sgn} q} \{ q \cdot T \}$$

$$= \sum_{q \in C(T)} (-1)^{\operatorname{sgn} q} \{ (\sigma \cdot q) \cdot T \}$$

$$= \sum_{q' \in C(\sigma \cdot T)} (-1)^{\operatorname{sgn}(\sigma^{-1}q'\sigma)} \{ (\sigma \cdot \sigma^{-1}q'\sigma) \cdot T \}$$

$$= \sum_{q' \in C(\sigma \cdot T)} (-1)^{\operatorname{sgn} q'} \{ q' \cdot (\sigma \cdot T) \}$$

$$= v_{\sigma \cdot T}.$$

Corollary 4.2. Let S^{λ} be the subspace of M^{λ} spanned by the v_T . Then S^{λ} is a subrepresentation.

The representation S^{λ} is called the **Specht module** associated with λ .

Exercise 4.3. Exhibit the following isomorphisms.

- (1) $S^{(n)} \cong V_{\text{triv}}$.
- (2) $S^{(1,1,1,\ldots,1)} \cong V_{\text{sgn}}$.
- (3) $S^{(n-1,1)} \cong V_{\text{stan}}$.

Constructing the Specht modules was fairly easy, but showing that they are irreducible and pairwise nonisomorphic is a little trickier.

We will need to do a bit of combinatorics. First, we define a total ordering on the set of all partitions called the **lexicographic ordering**. Given partitions λ and μ , we say that $\mu < \lambda$ if $\mu_i < \lambda_i$ for the first i where $\mu_i \neq \lambda_i$. (Here we use the convention that $\mu_i = 0$ if μ has fewer than i parts.)

The set of all partitions has a partial ordering called the **dominance ordering**. We write $\mu \leq \lambda$, and we say that λ dominates μ , if for all i,

$$\mu_1 + \mu_2 + \dots + \mu_i \le \lambda_1 + \lambda_2 + \dots + \lambda_i$$
.

We say that λ strictly dominates μ if $\mu \leq \lambda$ and $\mu \neq \lambda$.

Note that $\mu \leq \lambda$ implies $\mu \leq \lambda$, but not conversely.

Lemma 4.4. Let T and T' be numberings of λ and λ' , where λ does not strictly dominate λ' . Then either

- (1) there exist two distinct integers that occur in the same row of T' and the same column of T; or
- (2) $\lambda' = \lambda$, and there exist $p' \in R(T')$ and $q' \in C(T')$ such that $p' \cdot T' = q \cdot T$.

Proof. Suppose (1) does not hold. Then all of the entries in the first row of T' must appear in different columns of T, so we can find a $q_1 \in C(T)$ such that the first row of $q_1 \cdot T$ has all of the entries of the first row of T'. All of the entries in the second row of T' must also appear in different columns of T, and hence also those of $q_1 \cdot T$. So we can find a $q_2 \in C(T)$ such that the second row of $q_2 \cdot (q_1 \cdot T)$ has all the entries of the second row of T', and the first row still has the entries of the first row of T'. Iterating this procedure gives, for any $i, q_1, q_2, \ldots, q_i \in C(T)$ such that the first i rows of $q_i \cdot q_{i-1} \cdot \cdots \cdot q_1 \cdot T$ contain the entries of the first i rows of T'. This implies that $\lambda'_1 + \cdots + \lambda'_i \leq \lambda_1 + \cdots + \lambda_i$, and so $\lambda' \leq \lambda$.

Since we assumed λ did not strictly dominate λ' , it follows that $\lambda = \lambda'$. If k is the number of rows of λ , and $q = q_k \cdot q_{k-1} \cdot \dots \cdot q_1$, it follows that $q \cdot T$ and T' have the same entries in each row, by construction. Thus, there exists $p' \in R(T')$ such that $p' \cdot T' = q \cdot T$.

We define $b_T = \sum_{q \in C(T)} (-1)^{\operatorname{sgn}(q)} q \in \mathbb{C}G$. Then $v_T = b_T \cdot \{T\}$.

Exercise 4.5. (1) If
$$q \in C(T)$$
, then $b_T \cdot q = (-1)^{\operatorname{sgn} q} \cdot b_T$. (2) $b_T \cdot b_T = |C(T)|b_T$.

Lemma 4.6. Let T and T' be numberings of shapes λ and λ' , and assume that λ does not strictly dominate λ' . If there is a pair of integers in the same row of T' and the same column of T, then $b_T \cdot \{T'\} = 0$. Otherwise, $b_T \cdot \{T'\} = \pm v_T$.

Proof. Suppose such a pair exists. Let $t \in C(T) \cap R(T')$ be the transposition permuting them. Then

$$b_T t = \sum_{q \in C(T)} (-1)^{\operatorname{sgn}(q)} qt$$

$$= \sum_{q' \in C(T)} (-1)^{\operatorname{sgn}(q't^{-1})} q'$$

$$= -\sum_{q' \in C(T)} (-1)^{\operatorname{sgn}(q')} q'$$

$$= -b_T.$$

But $t \in R(T')$ implies $t \cdot \{T'\} = \{T'\}$, so

$$b_T \cdot \{T'\} = b_T \cdot (t \cdot \{T'\}) = (b_T \cdot t) \cdot \{T'\} = -b_T \cdot \{T'\},$$

so $b_T \cdot \{T'\} = 0$.

On the other hand, if there is no such pair, Lemma 4.4 gives us $p' \in R(T')$, $q \in C(T)$ such that $p' \cdot \{T'\} = q \cdot \{T\}$. We may then conclude

$$b_T \cdot \{T'\} = b_T \cdot \{p' \cdot T'\}$$

$$= b_T \cdot \{q \cdot T\}$$

$$= b_T \cdot q \cdot \{T\}$$

$$= (-1)^{\operatorname{sgn}(q)} b_T \cdot \{T\}$$

$$= (-1)^{\operatorname{sgn}(q)} \cdot v_T.$$

Theorem 4.7. The Specht module S^{λ} is irreducible, and $S^{\lambda} \ncong S^{\lambda'}$ for $\lambda \neq \lambda'$.

Proof. Let T be any numbering of λ . Then by Lemma 4.6 and Exercise 4.5.

$$b_T \cdot S^{\lambda} = b_T \cdot M^{\lambda} = \mathbb{C} \cdot v_T.$$

On the other hand, if $\lambda < \lambda'$ (in the lexicographic ordering), then λ does not dominate λ' , so by Lemmas 4.4 and 4.6, it follows that

$$b_T \cdot S^{\lambda'} = b_T \cdot M^{\lambda'} = 0.$$

Suppose $S^{\lambda} = V \oplus W$. Then

$$\mathbb{C} \cdot v_T = b_T \cdot S^{\lambda} = b_T \cdot V \oplus b_T \cdot W.$$

This implies that either V or W contains v_T , but v_T generates S^{λ} as a $\mathbb{C}S_n$ -module, so S^{λ} will be equal to whichever of V and W contains v_{λ} . Thus, S^{λ} is irreducible.

Suppose $\lambda \neq \lambda'$. Without loss of generality, we may assume $\lambda < \lambda'$. Then $b_T \cdot S^{\lambda'} = 0$, but $b_T \cdot S^{\lambda} \neq 0$, so we conclude that S^{λ} and $S^{\lambda'}$ are distinct.

We have placed the Specht modules in one-to-one correspondence with the conjugacy classes of S_n , so we have indeed constructed all irreducible representations of S_n .

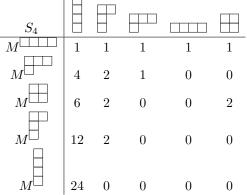
Of course, we are far from finished studying the representation theory of S_n . We would like, for instance, a nice formula for the character of S^{λ} . The usual procedure for producing such a formula is somewhat involved, but we will instead exhibit a simple algorithm.

The starting point for this algorithm is to note that if $\lambda < \lambda'$, then $M^{\lambda'}$ does not contain any copies of S^{λ} ; indeed, this follows directly from the fact that $b_T \cdot M^{\lambda'} = 0$ if T is a numbering of λ . We therefore know that M^{λ} is already irreducible when λ is highest in the lexicographic ordering (i.e., $\lambda = (n)$), and that for any lower λ , M^{λ} only contains Specht modules corresponding to higher partitions λ' .

We can thus proceed inductively. If we know the characters of the M^{λ} , and the characters of the S^{λ} for all $\lambda > \mu$, then we can deduce the character for S^{μ} by subtracting an appropriate number of copies of the characters of S^{λ} .

Exercise 4.8. Investigate the characters of the M^{λ} .

As an example, we recompute the character table for S_4 . First, we write down the characters for the M^{λ} in reverse lexicographic order.



As we said earlier, M^{\square} is irreducible, so $M^{\square} = S^{\square}$.

S_4					
S	1	1	1	1	1
M	4	2	1	0	0
M^{\square}	6	2	0	0	2
M	12	2	0	0	0
M	24	0	0	0	0

Taking the inner product of the first two characters gives 1, so we know that there is one copy of S in M. So we subtract the character of S from that of M to the character of S.

S_4					
S	1	1	1	1	1
S	3	1	0	-1	-1
M^{\square}	6	2	0	0	2
M	12	2	0	0	0
M	24	0	0	0	0

We continue in this fashion. M contains one copy of S , and one copy of S , so we subtract each character once.

S_4					
S	1	1	1	1	1
S	3	1	0	-1	-1
S \Box	2	0	-1	0	2
M^{\boxminus}	12	2	0	0	0
M^{\square}	24	0	0	0	0
S_4					
S_4	1	1	1	1	1
	1 3	1 1	1 0	1 -1	1 -1
	3	1	0	-1	-1

S_4					
S	1	1	1	1	1
S	3	1	0	-1	-1
S	2	0	-1	0	2
S	3	-1	0	1	-1
S	1	-1	1	-1	1

References

[Ful97] William Fulton, Young tableaux, London Mathematical Society Student Texts, no. 35, Cambridge University Press, 1997.

Department of Mathematics, University of Chicago, Chicago, IL 60637 $E\text{-}mail\ address:$ ejenkins@math.uchicago.edu