Essays on Coxeter groups

Representations of the symmetric group

Bill Casselman
University of British Columbia
cass@math.ubc.ca

I propose here a succinct account of the basic facts about the representations of $\mathfrak{S}_n$. Finding all the representations of any interesting group can be very difficult. For the symmetric group, however, there is a natural way to construct them, if a bit difficult to verify that they are irreducible and make up a complete set. I generally follow the presentation of Bill Fulton's book, but rely more on the use of induced representations.

Contents

1. The representation on polynomials

The clue to a natural construction of representations of $\mathfrak{S}_n$ is given by its one-dimensional representations. One of them is the trivial representation. Not much to be said there. But a second is the sign character. This can be most easily characterized in terms of the polynomial

$$\Delta_n(x) = \prod_{1 \le i < j \le n} (x_i - x_j)\,.$$

Any permutation σ of the range $[1, n]$ induces a permutation of the variables, replacing x_i by $x_{\sigma(i)}$. It will change $\Delta_n(x)$ to $\pm\Delta_n(x)$:

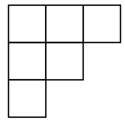
$$\sigma\Delta_n = \mathrm{sgn}(\sigma)\Delta_n\,.$$

It is easy to see that $\mathrm{sgn}(\sigma)\mathrm{sgn}(\tau) = \mathrm{sgn}(\sigma\tau)$ so that sgn is in fact a character—a one-dimensional representation—of $\mathfrak{S}_n$.

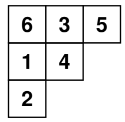
All other irreducible representations of $|gothS_n$ *can also be defined in terms of the permutations of polynomials.* They are parametrized by partitions of $[1, n]$, which is reasonable because the conjugacy classes are also parametrized by partitions, and we know that there are as many irreducible representations as conjugacy classes. Each partition $(\lambda_1, \lambda_2, \ldots, \lambda_k]$ with

$$\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_k$$

corresponds to a **Young diagram**, which has rows of lengths λ_1 etc. arrayed under each other. For example, the partition $6 = 3 + 2 + 1$ matches with



Suppose we are given a **numbering** of this diagram—that is to say, we label each box with a distinct number in the range $[1, n]$, like this:



Such a numbered diagram T—which is called a **(Young) tableau**—with $t_{i,j}$ is matrix position (i, j), corresponds to a **discriminant polynomial**

$$\Delta_T(x) = \prod_k \prod_{i<j} (x_{t_{i,k}} - x_{t_{j,k}}) \,.$$

That is to say, to construct Δ_T you first traverse all pairs in the first column, adding a factor $x_i - x_j$ when i occurs before j in the column; then go on to multiply by similar factors for all the succeeding columns. For example, the numbering above gives rise to

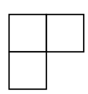
$$(x_6 - x_1)(x_6 - x_2)(x_1 - x_2)(x_3 - x_4) \,.$$

Let $C(T)$ be the subgroup of $\mathfrak{S}_n$ which permutes entries in each of the columns of T. It is a direct product of symmetric groups. If σ lies in $C(T)$ then $\sigma\Delta_T = \mathrm{sgn}(\sigma)\Delta_T$.

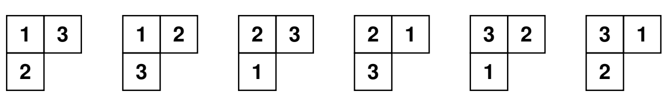
Lemma 1.1. *Any polynomial P in n variables with the property that $\sigma P = \mathrm{sgn}(\sigma)P$ is a multiple of Δ_T.*

This because the polynomial ring is a unique factorization domain. In particular, Δ_T is up to scalar the only eigenfunction of $C(T)$ with character sgn.

A horizontal diagram gives rise to the constant 1, and a vertical diagram gives rise, up to sign, to the polynomial that defines sgn. In general, every diagram λ gives rise to the set of polynomials of all those corresponding to various tableaux with the shape determined by the diagram. For example, if $n = 3$, the diagram



gives rise to the tableaux and polynomials



$$x_1 - x_2 \quad x_1 - x_3 \quad x_2 - x_1 \quad x_2 - x_3 \quad x_3 - x_1 \quad x_3 - x_2$$

Of course permutations of $[1, n]$ permute labels in such a diagram, giving a new diagram and a new polynomial. The linear span $\mathfrak{P}_\lambda$ of these polynomials is therefore the space of a representation of $\mathfrak{S}_n$, which I call π_λ.

These polynomials are not linearly independent. In the example above, we have relations

$$\begin{aligned} x_2 - x_1 &= -(x_1 - x_2) \\ x_3 - x_1 &= -(x_1 - x_3) \\ x_3 - x_2 &= \quad (x_1 - x_2) - (x_1 - x_3) \\ x_2 - x_3 &= -(x_1 - x_2) + (x_1 - x_3)\,. \end{aligned}$$

The polynomials $x_1 - x_2$ and $x_1 - x_3$ make up a basis of this representation space. They correspond to the labeled diagrams

1	3
2	

$x_1 - x_2$

1	2
3	

$x_1 - x_3$

What characterizes these somewhat special tableaux? An **ordered tableau** is one with the property that number labels increase from left to right and top to bottom. (In the literature, these are usually called **standard tableaux**, but there are far too many things in the mathematical world called standard.)

The main result in this business is this:

Theorem. *The polynomials associated to ordered tableaux of a given partition λ form a basis of the representation π_λ of $\mathfrak{S}_n$ on the space of all discriminant polynomials associated to it.*

Proof. In the rest of this section I'll show that they span the space $\mathfrak{P}_\lambda$ and explain an algorithm laying out how to write down explicitly the representation of $\mathfrak{S}_n$ in terms of ordered tableaux. In the next, I'll prove independence, which is best accounted for in terms of induced representations.

I take the algorithm from Chapter 7 of Bill Fulton's book on Young tableaux, but the approach I have taken is a bit different from the exposition there. The description of the representation associated to a partition in terms of polynomials is Exercise 17 in that chapter, and I'll use earlier parts to justify its solution. Using polynomials as the basis of the construction of representations is a more natural starting point than the somewhat abstract one Fulton's book takes, and fits better with the more general environment of other finite Coxeter groups.

Put an order on pairs (i, j):

$$(i_1, j_1) < (i_2, j_2)$$

if $j_1 < j_2$ or $j_1 = j_2$ and $i_1 < i_2$. Thus left columns are smaller than right columsn, and higher rows are smaller than lower ones. If S and T are two tableaux, I say $S < T$ if $S_{i,j} < T_{i,j}$ where (i, j) is the least location where they differ. Another way to see this is to associate to each tableau a sequence of numbers, listing the entries in each column from the bottom up, proceeding through the columns right to left. For example, from the following figure we get the string $5\,4\,3\,2\,1\,6$.

6	3	5
1	4	
2		

Then $S < T$ if the word associated to S comes first in dictionary order.

Now, suppose we want to express the polynomial Δ_T associated to a tableau T as a linear combination of polynomials associated to ordered tableaux.

Step 1. In an ordered tableaux we arrange the columns of T so that the entries are increasing in each column, and in doing this we only change the sign of Δ_T.

Step 2. Suppose that the labeling we now have is not an ordered tableau. That means that for some j, k we have $T_{k,j} > T_{k,j+1}$. Let $\pi_{k,j}(T)$ be the collection of all tableaux we get by exchanging all top k elements of column $j+1$ with k entries in column j, preserving order in the column. One of these exchanges amounts neither more nor less than choosing from which rows we take the new entries in column j, so the number of new tableaux we get is $n_j!/k!(n_j-k)!$ if column j has n_j entries. The basic fact now is that

$$\Delta_T = \sum_{S \in \pi_{k,j}(T)} \Delta_S \,.$$

I postpone verifying this. If we do not now have all ordered tableaux we loop back to Step 1.

We must assure ourselves that sooner or later this process will halt. But each step we are left with tableaux that come earlier in this ordering, so eventually we must come to a stop.

I take the following example from Fulton's book (page 98). We start out with the tableau

$$\underset{T_0}{\begin{array}{|c|c|}\hline 4 & 3 \\ \hline 1 & 6 \\ \hline 5 & 2 \\ \hline \end{array}}$$

In Step 1 we rearrange its columns to get

$$\underset{T_0}{\begin{array}{|c|c|}\hline 4 & 3 \\ \hline 1 & 6 \\ \hline 5 & 2 \\ \hline \end{array}} = -\underset{T_1}{\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 4 & 3 \\ \hline 5 & 6 \\ \hline \end{array}}$$

The rearrangement of columns involves 3 transpositions, so $\Delta_{T_0} = -\Delta_{T_1}$ (which I shall henceforth write in a more succinct fashion). The tableau T_1 is not ordered, because in the second row $4 > 3$. There are three possible pairs to choose from column 1, so the collection $\pi_{2,1}$ is made up of three tableaux, giving

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 4 & 3 \\ \hline 5 & 6 \\ \hline \end{array} = \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 & 4 \\ \hline 5 & 6 \\ \hline \end{array} + \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 4 & 5 \\ \hline 3 & 6 \\ \hline \end{array} + \begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 & 5 \\ \hline 3 & 6 \\ \hline \end{array}$$

We now loop back to Step 1, and rearrange the first column of the second tableau to bring it into order. This gives us

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

The third tableau is ordered, but the first two are not, so we perform Step 2 on these. This gives us

$$\begin{aligned} \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 & 4 \\ \hline 5 & 6 \\ \hline \end{array} &= \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & 6 \\ \hline \end{array} + \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & 4 \\ \hline 5 & 6 \\ \hline \end{array} + \begin{array}{|c|c|}\hline 2 & 5 \\ \hline 3 & 4 \\ \hline 1 & 6 \\ \hline \end{array} \\ &= \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & 6 \\ \hline \end{array} - \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 4 \\ \hline 5 & 6 \\ \hline \end{array} - \begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 & 5 \\ \hline 3 & 6 \\ \hline \end{array} \end{aligned}$$

$$\begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 & 5 \\ \hline 4 & 6 \\ \hline\end{array} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 5 \\ \hline 4 & 6 \\ \hline\end{array} + \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & 5 \\ \hline 4 & 6 \\ \hline\end{array} + \begin{array}{|c|c|}\hline 2 & 4 \\ \hline 3 & 5 \\ \hline 1 & 6 \\ \hline\end{array}$$

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

and we may halt with the equation

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

$$= - \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & 6 \\ \hline\end{array} + \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 4 \\ \hline 5 & 6 \\ \hline\end{array} + \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 5 \\ \hline 4 & 6 \\ \hline\end{array} - \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 5 \\ \hline 4 & 6 \\ \hline\end{array} + \begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 & 5 \\ \hline 3 & 6 \\ \hline\end{array} ,$$

or more precisely

$$\Delta_{\begin{array}{|c|c|}\hline 4 & 3 \\ \hline 1 & 6 \\ \hline 5 & 2 \\ \hline\end{array}} = -\Delta_{\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline 5 & 6 \\ \hline\end{array}} + \Delta_{\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 4 \\ \hline 5 & 6 \\ \hline\end{array}} + \Delta_{\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 5 \\ \hline 4 & 6 \\ \hline\end{array}} - \Delta_{\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 5 \\ \hline 4 & 6 \\ \hline\end{array}} + \Delta_{\begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 & 5 \\ \hline 3 & 6 \\ \hline\end{array}} .$$

This argument proves that the polynomials associated to ordered tableaux span the space, except that it remains to justify the identity

$$\Delta_T = \sum_{S \in \pi_{k,j}(T)} \Delta_S .$$

For this, I follow the answer to Exercise 7.17 of Fulton's book, who in turn follows [Peel:1975].

I'll prove in the next section that the discriminants of ordered tableaux are linearly independent.

2. Relations with induced representations

In this section let $G = \mathfrak{S}_n$.

I recall that if T is a tableau corresponding to the partition λ_i of n, $C(T)$ is the subgroup of G permuting only the entries within its columns. Let $R(T)$ be that permuting entries within its rows. The size of $R(T)$, for example, is

$$|R(T)| = \prod \lambda_i!$$

Lemma 2.1. *The intersection of $C(T)$ and $R(T)$ is $\{1\}$.*

Proof. Because if g in $R(T)$ is not 1, it must change the content of at least two columns. $\square$

There are two types of induced representations occurring in the theory of representations of $\mathfrak{S}_n$: the first is the representation of G on

$$\mathrm{I}_T(1) = \mathbb{C}[R(T)\backslash G] = \mathrm{Ind}(1 \,|\, R(T), G)\,,$$

which is induced from the trivial representation of $R(T)$ to G. The second is that on the space

$$\mathrm{I}_T(\mathrm{sgn}) = \mathrm{Ind}(\mathrm{sgn} \,|\, C(T), G)$$

Lemma 2.2. *If $S = gT$ then $C(S) = gC(T)g^{-1}$ and $R(S) = gR(T)g^{-1}$.*

Proof. Let $\mathrm{col}_k(T)$ be the set of entries in the k-th column of T. The subgroup $C(S)$ is that of all c in G such that

$$\mathrm{col}_k(c\,T) = c\,\mathrm{col}_k(T) = \mathrm{col}_k(T)$$

for all k. If $S = gT$ then

$$\mathrm{col}_k(xS) = \mathrm{col}_k(xgT) = xg\,\mathrm{col}_k(T), \quad \mathrm{col}_k(S) = \mathrm{col}_k(gT) = g\,\mathrm{col}_k(T)$$

so $\mathrm{col}_k(xS) = \mathrm{col}_k(S)$ if and only if $g^{-1}xg$ lies in $C(T)$. $\square$

If S and T have the same shape—are associated to the same partition—then $S = \sigma T$ for some unique permutation σ, and $\mathrm{I}_S(\chi)$ is canonically isomorphic to $\mathrm{I}_T(\chi)$ for $\chi = 1$ or sgn. Therefore these representations depend only on the shape of the tableau. In fact, one can identify them with a single space that does not depend on any choices.

The set of tableaux is a principal homogeneous set for G—it acts by permuting entries in tableaux, and the isotropy subgroup of any tableau is trivial. Therefore if we fix a tableau T, any other tableau S of the same shape is gT for some g in G. To f in $\mathrm{I}_T(\mathrm{sgn})$ assign the function F on the set of all tableaux of the same shape according to the rule

$$F(S) = f(g^{-1}) \text{ if } S = gT\,.$$

This is certainly well defined, since g is uniquely determined. But then for $c_S = gc_Tg^{-1}$ in $C(S)$ we have $c_SS = c_SgT = gc_T$ and

$$F(c_SS) = f(c_T^{-1}g^{-1}) = \mathrm{sgn}(c_T)f(g^{-1}) = \mathrm{sgn}(c_S)F(S)\,,$$

since sgn is a real character of G itself, and c_S is conjugate to c_T. Let $\mathfrak{T}_\lambda$ be the set of tableau of shape λ. From this equation it is straightforward to deduce:

Proposition 2.3. *The map taking f in $\mathrm{I}_T(\mathrm{sgn})$ to the function F such that*

$$F(gT) = f(g^{-1})$$

is a G-covariant isomorphism of $\mathrm{I}_T(\mathrm{sgn})$ with the space I_λ of all functions on $\mathfrak{T}$ such that $F(c_T T) = \mathrm{sgn}(T)F(T)$ for all T.

This identification is implicit in Fulton's book and indeed much of the treatment of $\mathfrak{S}_n$ in the literature. A similar identification can be made for $\mathrm{I}_T(1)$.

What do these induced representations have to do with the representation of G on $\mathfrak{P}_\lambda$? There is an obvious and simple map from $\mathfrak{T}$ to $\mathfrak{P}_\lambda$, taking T to Δ_T. This gives rise to a map backwards from the *dual* of $\mathfrak{P}_\lambda$ to the space of functions F on $\mathfrak{T}$:

$$F(T) = f(\Delta_T)\,.$$

Proposition 2.4. *This map from $\widehat{\mathfrak{P}}_\lambda$ to the space of functions on $\mathfrak{T}$ has its image in $\mathrm{I}_\lambda(\mathrm{sgn})$, and is G-covariant.*

Proof. The G-covariance is immediate. The second, because

$$F(c_T T) = f(\Delta_{c_T T}) = f(\mathrm{sgn}(c_T)T) = \mathrm{sgn}(c_T)f(T) = \mathrm{sgn}(c_T)F(T)\,. \square$$

What is the image?

Let ρ_T be the element in the group algebra of $R(T)$

$$\rho_T = \frac{1}{|R(T)|}\sum_{g\in R(T)} g$$

and let σ_T be this one in the group algebra of $C(T)$:

$$\sigma_T = \frac{1}{|C(T)|}\sum_{g\in C(T)} \mathrm{sgn}(g)\, g$$

These are both idempotent. The first amounts to projection onto $R(T)$-fixed vectors, the second onto the sgn-eigenspace of a representation of $C(T)$.

NOTE: the group $C(T)$ is contained in a larger group $\mathfrak{C}(T)$ that permutes columns of the same length wholesale. Any one of these changes Δ_T by at most a sign, since it effectively interchanges the discriminants of columns. The group $C(T)$ is normal in this? Quotient is another product of symmetric groups. There is a splitting, into the intersection of this group with $R(T)$, whose actions do not change Δ_T. There is also a similar group $\mathfrak{R}(T)$ containing $R(T)$. What role do these larger groups play?

Explicit effect of σ_T, ρ_T via Fulton's Lemma.

Order the ordered tableaux, and apply the sign idempotent (Fulton: Corollary in the middle of page 87). Same argument says that the representation is irreducible, since (a) any Δ_T spans the space, and (b) in any decomposition we must have the unique eigenfunction of S_T in one or the other component. Also implies no two are the same.

We only have to look at columns j and $j+1$, and in column $j+1$ only the top k entries. Let A be the group that permutes the entries in column j among themselves, likewise the top k entries of column $j+1$, and let B be the group of all permutations of those entries. Frobenius reciprocity is applied to the representation $\mathrm{sgn}\otimes I$ of A induced to B . . . The point is that the sgn representation of B does not occur in that induced representation. (See Fulton's book, Claims 1 and 2 on page 100.)

3. Characters

The final thing to do is to find the character of the representation associated to a given Young diagram. There is in fact no simple formula, but merely a recursive algorithm . . .

4. Relations with symmetric polynomials

$R = \mathbb{C}[x]$ is free over $I = \mathbb{C}[x]^{\mathfrak{S}_n}$. If $I = (x)^{\mathfrak{S}_n}$, R/I is isomorphic to $\mathbb{C}[\mathfrak{S}_n]$ as a $\mathfrak{S}_n$ module.

Symmetric polynomials?

5. Relations with conjugacy classes of nilpotent matrices

The dominance order and unipotent classes in $\mathrm{GL}_n(\mathbb{C})$. The classes are parametrized by partitions of n, each set of Jordan blocks $n_1 \geq n_2 \geq \ldots n_k$ determines a class. Thus 1 corresponds to $1 + 1 + \cdots + 1$ and the regular unipotent to a single large block. The class of λ lies in the closure of that of μ if and only if $\lambda \leq \mu$. All this is not an accident, there is a natural way to associate a representation of $\mathfrak{S}_n$ to every class . . .

6. References

1. W. Fulton, **Young tableaux**, Cambridge University Press, 1997.

2. M. H. Peel, 'Specht modules and symmetric groups', *Journal of Algebra* **36** (1975), 88–97.

3. A. Young, 'On quantitative substitutional analysis (third paper)', *Proceedings of the London Mathematical Society* **28** (1928), 255–292.