A Combinatorial Introduction to the Representation Theory of S_n

Mark Wildon

January 2001

Contents

1 **Introduction** 2

2 **Construction of Specht modules** 2

3 **The RSK-correspondence and a basis for S^λ** 6

4 **Background on symmetric functions** 9

5 **The Frobenius character formula** 15

6 **Sources** 22

1 Introduction

This text is a much expanded version of a seminar I gave in November 2000 to the Oxford Algebra Kinderseminar. The purpose of the seminar was to give an introduction to the basic representation theory of the symmetric group S_n, stressing its combinatorial aspects. The planned denouement, using the Robinson-Schensted-Knuth correspondence to prove most of the fundamental facts about Specht Modules, which unfortuntately was never reached can be read here. Alas it is not original but I still think it is a nice application of combinatorics to representation theory that deserves to be better known.

I have taken the opportunity to add two final sections connecting the characteristic zero representation theory of Sn with symmetric polynomials. First we develop the necessary properties of the polynomials assuming minimal background knowledge. Then we prove the Frobenius character formula and deduce from it the hook formula (for the number of standard tableau of a given shape). This gives in turn an attractive application of algebra to combinatorics. There are now fairly short combinatorial proofs of the Hook Formula (see sources at end) but it is still nice to see the interplay between the representation theory of the symmetric group and combinatorics producing applications in both directions.

Details of the sources for each section are given at the end.

2 Construction of Specht Modules

In this section we will construct a family of kS_n modules indexed by the partions of n known as Specht modules. For the moment we make no assumptions on our field k but later on we will be mainly concerned with the case when the characteristic of k is zero. In this case we will see that the Specht modules form a complete set of non-isomorphic irreducible representations. The methods used here extend to give results about the positive characteristic case but we will not pursue the subject much in this direction (despite it's considerable interest). Inevitably we have to begin with several definitions.

Let λ be a partition of n, written $\lambda \vdash n$. A **λ-tableau** is a numbering of the Young diagram with distinct numbers from $[1..n]$. More formally a tableau can be regarded as a bijection $t : \lambda = [1..n] \to [1..n]$ (but little is gained by doing so).

Our first observation is that S_n acts on the set of λ-tableau in an obvious way: if m is the number at a certain node of the λ tabloid t then $m\sigma$ is the node in the new tableau. With the definition of tableau as function the new tableau is simply $t\sigma$. Linearizing this action gives a $n!$ dimensional representation of S_n where elements of the group act on the right

Given a λ-tableau t there are associated subroups of S_n:

$$\begin{aligned} R_t &= \{\sigma \in S_n : \sigma \text{ preserves rows of } t\} \text{ and} \\ C_t &= \{\sigma \in S_n : \sigma \text{ preserves columns of } t\} \end{aligned}$$

We will also need the signed column sum b_t defined by $b_t = \sum_{\sigma \in C_t} \epsilon(\sigma)\sigma$. If t_λ is the λ-tableau where entries increase first along the rows and then down the columns let $S_\lambda = R_{t_\lambda}$, the **Young subgroup** associated with the partition λ.

For example if $\lambda = (4, 2, 1)$ then $S_\lambda = S_{\{1,2,3,4\}} \times S_{\{5,6\}} \times S_{\{7\}}$ and:

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

A **λ-tabloid** is a Young tableau where we don't care about the order of entries within each row. More formally a tabloid is an equivalence class of tableaux under the action of their row stabiliser group. If t is a λ-tableau we write $\{t\}$ for the associated λ-tabloid. We draw tabloids by missing out the vertical lines from an associated tableau. For example if $\lambda = (4, 2, 1)$ and

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

then $R_t = S_{\{1,3,6,7\}} \times S_{\{4,5\}} \times S_{\{2\}}$ and the associated tabloid can be written as:

$$\{t\} = \begin{array}{cccc}\hline 3 & 6 & 7 & 1 \\ \hline 5 & 4 \\ \cline{1-2} 2 \\ \cline{1-1}\end{array} = \begin{array}{cccc}\hline 1 & 3 & 6 & 7 \\ \hline 4 & 5 \\ \cline{1-2} 2 \\ \cline{1-1}\end{array} = \dots$$

We note that the action of S_n on tableaux extends naturally to a well-defined action of S_n on the set of λ-tabloids, given by $\{t\}\sigma = \{t\sigma\}$. For if s and t are λ-tableaux representing the same tabloid, with say $s\tau = t$ ($\tau \in R_s$) then $s\sigma\tau^\sigma = s\sigma\,(\sigma^{-1}\tau\sigma) = t\sigma$ so $\{s\sigma\} = \{t\sigma\}$ showing that the action is well defined. That $\tau^\sigma \in R_{s\sigma}$ is easily seen by inspection.

Now let M^λ be the vector space spanned by k-linear combinations of tabloids. Our result shows that M^λ is a representation of S_n in the natural way. It will be important to us later that $M^\lambda = \mathrm{Ind}_{S_\lambda}^{S_n} k$. (The reason for not defining M^λ in this way will emerge shortly.)

Define the **λ-polytabloid** e_t associated with a λ-tableau t by

$$e_t = \{t\}b_t = \{t\} \sum_{\sigma \in C_t} \sigma\epsilon(\sigma)$$

We can now finally define the **Specht module**, S^λ associated with the partition λ to be the subspace of M^λ spanned by all polytabloids. Thus S^λ is defined as a subrepresentation of M^λ. Since $e_t\sigma = e_{t\sigma}$, S^λ is cyclic, generated by any one polytabloid. Although it is easy to show that the dimension of M^λ is $n!/(\lambda_1! \dots \lambda_k!)$, finding the dimension of S^λ is a much harder problem that will be the motivation for some of our later work.

Before proceeding further we give the three standard examples of Specht modules. Notice that in all cases if char $k = 0$ then S^λ is irreducible. Another instructive example which can be analysed by hand is $S^{(3,2)}$: see James, *The representation theory of the symmetric groups* [5], Ch. 5.

1. For any n, $M^{(n)}$ is the trivial representation and $S^{(n)} = M^{(n)}$.

2. $M^{(1^n)}$ is the regular representation of S_n (this can be seen immediately from its description as an induced module) and:

$$S^{(1^n)} = \left\langle t \sum_{\pi \in S_n} \epsilon(\pi)\pi \mid t \text{ is a } (1^n)\text{-tableau} \right\rangle.$$

It is easy to check that this implies $S^{(1^n)}$ is one dimensional and isomorphic to the sign representation, ϵ of S_n.

3. For any n, $M^{(n-1,1)}$ is the (linearized) action of S_n on $[1..n]$ as any $(n-1,1)$ tabloid is determined by the entry in its second row. Accordingly we will omit the redundant first row and so write $\overline{\underline{3}}$ rather than $\begin{array}{l}\overline{\underline{1\ 2\ 4}}\\ \underline{3}\end{array}$. Let t_{ij} be a tableau with first column entries (i,j). We see that:

$$\begin{aligned} S^{(n-1,1)} &= \langle t_{ij}\kappa_{t_{ij}} \mid i,j = 1..n \rangle \\ &= \langle \overline{\underline{j}} - \overline{\underline{k}} \mid i,j = 1..n \rangle \end{aligned}$$

so $S^{(n-1,1)}$ is $(n-1)$ dimensional and is isomorphic to the standard permutation module on n letters, quotiented out by $< 1 + 2 + \ldots + n > \cong k$. We already know (from the theory of permuation representations) or can easily check that it is irreducible in characteristic zero.

We now prove the key lemma about polytabloids. The proof is typical of the kind of combinatorial reasoning we need to employ. (Some accounts prove a more general version but I prefer to leave this until we really need it — see lemma 3.)

Lemma 1 *If u and t are λ-tableaux then either $\{u\}b_t = 0$ or $\{u\}b_t = \pm e_t$.*

Proof: Suppose there are two numbers, i and j that appear in the same row of u and in the same column of t. Then the transposition (ij) appears in C_t. If we write C_t as a union of right cosets for $\{1,(ij)\}$, say $C_t = \cup_k \{1,(ij)\}g_k$ we obtain:

$$\{u\}b_t = \{u\}\sum_k (1-(ij))g_k = \sum_k (\{u\} - \{u\})g_k = 0.$$

(We may assume $\epsilon(g_k) = 1$). So if $\{u\}b_t \neq 0$ any two numbers in the same row of u must lie in different columns of t. By applying some τ in C_t we may rearrange the entries in the first column of t so that they each entry appears in the same row of t as u. Continuing in this way for each column of t we obtain a τ in C_t such that $\{u\} = \{t\}\tau$. Thus:

$$\{u\}b_t = \{t\}\tau b_t = \{t\}b_t\epsilon(\tau) = \pm e_t. \qquad \square$$

The next theorem, due to James, is the key to beginning to understand the structure of S^λ.

Theorem 2 (James' submodule theorem) *Define a non-degenerate symmetric S_n-invariant bilinear form on M^λ by:*

$$\langle \{s\},\{t\} \rangle = \begin{cases} 1 & \text{if} \quad \{s\} = \{t\} \\ 0 & \text{if} \quad \{s\} \neq \{t\} \end{cases}$$

Then if U is a subrepresentation of M, either $U \supseteq S^\lambda$ or $U \subseteq S^{\lambda\perp}$. $\perp$ is, of course, taken with respect to $\langle\ ,\ \rangle$.

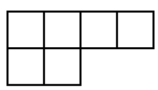
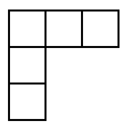
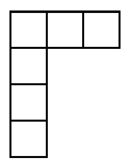
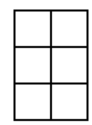
Proof: Suppose there is an element $u \in U$ and a tableau t such that $ub_t \neq 0$. Then by the lemma above, ub_t is a scalar multiple of e_t so $e_t \in U$. But since S^λ is generated by any single polytabloid, this means $U \supseteq S^\lambda$. Otherwise:

$$\langle u, e_t \rangle = \langle u, \{t\}b_t \rangle = \langle ub_t, \{t\} \rangle = \langle 0, \{t\} \rangle = 0 \ \ \forall u \in U, \text{ tableau } t$$

implying $U \subseteq S^{\lambda\perp}$. □

An immediate corollary is that if char $k = 0$ then S^λ is irreducible. For if U is a submodule of S^λ then by James' submodule theorem either $U = S^\lambda$ or $U \subseteq S^\lambda \cap S^{\lambda\perp} = 0$.

We now want to show that we have found all irreducible representations. First of all we show that if $\lambda \neq \mu$ then $S^\lambda \not\cong S^\mu$. To do this we extend lemma 1 to cover the case where u and t are tableau for different partitions. Introduce a partial ordering (often called the dominance or natural order) on the set of partitions of n by defining $\lambda \trianglerighteq \mu$ if $\sum_{i=1}^m \lambda_i \geq \sum_{i=1}^m$ for every m. (If this calls for a non-existent part of λ or μ simply take the size of that part to be zero). For example:

 $\trianglerighteq$  but  is not comparable to .

Lemma 3 *Let u be a μ-tableau and let t be a λ-tableaux. Then either $\{u\}b_t = 0$ or $\lambda \trianglerighteq \mu$ and $\{u\}b_t = \pm e_t$.*

Proof: Just as before we argue that if $\{u\}b_t \neq 0$ then any two numbers appearing in the same row of u must lie in different columns of t. Looking at the first row of u we see that μ_1 numbers must be fitted into the λ_1 columns of t, so $\lambda_1 \geq \mu_1$. Similarly, looking at the second row of u, we see that μ_2 numbers must be fitted into $\lambda_1 - \mu_1 + \lambda_2$ columns, so $\lambda_1 + \lambda_2 \geq \mu_1 + \mu_2$. Continuing in this way (by induction on the number of rows of u) gives the result. □

Theorem 4 *Let* char $k = 0$. *If $\lambda \neq \mu$ then $S^\lambda \not\cong S^\mu$.*

Proof: Suppose $\phi : S^\mu \to S^\lambda$ is an isomorphism. Extend ϕ to a map $\bar{\phi} : M^\mu \to S^\lambda$ by setting it zero on $S^{\mu\perp}$ (recall that in characteristic zero $M^\mu = S^\mu \oplus S^{\mu\perp}$). As $\bar{\phi}$ is non zero, we can find $\{u\} \in M^\mu$ and a λ tableau t such that $\left\langle \{u\}\bar{\phi}, e_t \right\rangle \neq 0$. Therefore $0 \neq \{u\}\bar{\phi}b_t = \{u\}b_t\bar{\phi}$. We conclude that $\{u\}b_t \neq 0$. By the lemma $\lambda \trianglerighteq \mu$. But ϕ is an isomorphism so by symmetry $\mu \trianglerighteq \lambda$. Therefore $\lambda = \mu$. □

Note that we have also shown that if ϕ is a non zero homomorphism from M^μ to S^λ then $\lambda \trianglerighteq \mu$. In characteristic zero we know that complete reducibility holds (Maschke's theorem) so we can write:

$$M^\mu = S^\mu \oplus \bigoplus_{\lambda \triangleright \mu} k_{\lambda\mu} S^\lambda$$

where $k_{\lambda\mu}$ is a non-negative integer. In the final section we will be able to give a combinatorial interpretation of these coefficients.

Most readers will be aware of the general theorem that the number of irreducible representations of a group (over a splitting field of characteristic zero) is equal to the number of conjugacy classes in that group. Using this, together with theorem 4 above, we can immediately deduce that the Specht modules S^λ, $\lambda \vdash n$ form a complete set of irreducible kSn representations for any field k such that char $k = 0$. (We have shown that $\mathbb{Q}$ is a splitting field for S_n so nothing is changed by extending the ground field.) In the next section we will be able to deduce this result from the Wedderburn theorem. One could perhaps argue that proving the Wedderburn theorem involves almost as much work as proving the theorem on conjugacy classes but at least it makes a change!

Although we have proved the Specht modules are irreducible we are still somewhat in the dark about their properties. In the next section we will find a basis for each S^λ which should go some way to addressing this problem.

3 The RSK-correspondence and a basis for S^λ

Next we look for a basis for S^λ. Our first step is to establish the Richardson-Schensted-Knuth (RSK) correspondence which will be the key to our proof. The RSK correspondence will also receive important use in section 4 so we establish it in slightly greater generality than we immediately need. Some proofs will be sketched since I don't think much is gained by long-winded explanation when a few hand examples make everything clear. Greater rigour can be found in Fulton *Young Tableaux* [3] (Ch. 4) and Stanley *Enumerative Combinatorics II* [8] (§7.11).

Call a tableau **standard** if its rows and columns are increasing sequences. Let f^λ be the number of standard tableau. We will also need **semi-standard** tableau. These are tableau where the entries need no longer be distinct. We do however insist that the rows are weakly increasing and the columns are strictly increasing. For example, the first tableau below is semi-standard but not standard. The second is standard (and so semi-standard).

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

The basic operation in the RSK algorithm is row insertion into a semi-standard tableau. To row insert the number m into row j of the tableau t proceed as follows: if no number if row j exceeds m, put m at the end of row j. Otherwise find the smallest number not less than m and replace it with m. (If there are several to choose from pick the leftmost). The replaced number, r say is said to be 'bumped out' of the tableau. We can now continue by inserting r into row $j+1$. If we are required to insert a number into an empty row we simply create a new box at the start of that row and put the number into it. We write $m \leftarrow t$ for the tableau resulting from this operation.

For example, when we row insert 2 into the semi-standard tableau

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

the 2 bumps the 3 from the first row, which in turn bumps the first 5 from the second row, which bumps the 7 from the third row, which is put at the end of the forth row. (It really is a lot easier to do than describe!) The resulting tableau is:

$$\begin{ytableau} 1 & 2 & 2 & 2 & 5 \\ 2 & 3 & 3 & 5 \\ 4 & 4 & 5 \\ 5 & 6 & 7 \end{ytableau}\ .$$

Notice that if we are given a tableau resulting from a row bumping operation *and* told the location of the last box formed (necessarily an outside corner) then we can run the algorithm backwards to recover the original tableau and the inserted number. For example, if we were told that the last box entered into the tableau above was the one containing the 7 we would know that the 7 had bumped the 5 in the row above which had bumped the rightmost 3 in the row above which had bumped the rightmost 2 in the first row. There are no further rows so we know 2 must have been the inserted number.

It is easy to see that if t is standard and m does not lie in t then $m \leftarrow t$ is standard and that if t is semi-standard then $m \leftarrow t$ is semi-standard.

For our purposes, the RSK algorithm is a map (in fact a bijection)

$$\begin{gathered}\big(2 \times n \text{ lexicographically ordered arrays with entries from } [1..n]\big) \\ \longrightarrow \big(\text{pairs } \ (s,t) \ \text{ of semi-standard tableaux}\big)\end{gathered}$$

where the entries of s are the numbers in the bottom row of the array and the entries of t are the numbers in the top row of the array. A two-line array is ordered lexicographically if whenever $\binom{i}{j}$ occurs before $\binom{i'}{j'}$ either $i < i'$ or $i = i'$ and $j \leq j'$. Note that any permutation is represented (in two-line notation) by such an array. In this section we will only need the RSK correspondence for arrays representing permutations.

For example we could start with the array:

$$\begin{pmatrix} 1 & 1 & 2 & 3 & 3 & 3 \\ 3 & 3 & 2 & 1 & 2 & 4 \end{pmatrix}.$$

RSK 1: We start with two empty tableaux, s_0 and t_0 say. If the first entry in our two-line array is $\binom{i}{j}$ then row insert j into the first row of s. This creates a new box somewhere in s. Find the corresponding position in t and put i into that box. Call the resulting tableaux s_1 and t_1.

RSK 2: Repeat step 1 for each pair in the array. We end up with two tableau, $s_n = s$ and $t_n = t$. It is immediate that s is semi-standard since it is the result of a series of row insertions. That t is semi-standard can be proved quite easily. (Basically we argue that if i' is inserted in a box below i in t then j' must have bumped an entry from a higher line. So j' is smaller than a previous entry in the array. As the array is ordered lexicographically this can only happen if $i' > i$.)

For our example array above we form the sequence of tableaux:

$$\left(\begin{ytableau} 3 \end{ytableau}, \begin{ytableau} 1 \end{ytableau}\right) \to \left(\begin{ytableau} 3 & 3 \end{ytableau}, \begin{ytableau} 1 & 1 \end{ytableau}\right) \to \left(\begin{ytableau} 2 & 3 \\ 3 \end{ytableau}, \begin{ytableau} 1 & 1 \\ 2 \end{ytableau}\right) \to$$

$$\left(\begin{ytableau} 1 & 3 \\ 2 \\ 3 \end{ytableau}, \begin{ytableau} 1 & 1 \\ 2 & 3 \end{ytableau}\right) \to \left(\begin{ytableau} 1 & 2 \\ 2 & 3 \\ 3 \end{ytableau}, \begin{ytableau} 1 & 1 \\ 2 & 3 \\ 3 \end{ytableau}\right) \to \left(\begin{ytableau} 1 & 2 & 4 \\ 2 & 3 \\ 3 \end{ytableau}, \begin{ytableau} 1 & 1 & 3 \\ 2 & 3 \\ 3 \end{ytableau}\right).$$

The RSK algorithm is reversible in that given a pair of semi-standard tableaux (s, t) we can recover the corresponding two-row array. In the case where the original array corresponded to a permutation and so the resulting tableaux are standard this is easy to see: t records the location of the i^{th} box drawn in s. Using reverse row-insertion on its entry yields the previous s. For the general case it is not quite so obvious how to reverse the process because we don't immediately know in which order boxes containing the same number were drawn in t. In fact the rightmost entry in t always corresponds to the last number inserted. This point is clearly explained in Stanley [8]. Granted this, it is clear that the reverse algorithm can be applied to any pair of semi-standard tableaux. The theorem below summarises our findings.

Theorem 5 *The RSK correspondence is a bijection between lexicographically ordered $(2 \times n)$ arrays and pairs of semi-standard tableaux (s, t) with shape λ for some partition $\lambda \vdash n$ (with all entries as described above). Restricted to arrays representing a permutation the RSK correspondence gives a bijection between elements of S_n and pairs of standard tableaux with entries from $[1..n]$.* □

Corollary 6 *There are $n!$ elements of S_n which are in bijection with pairs of standard tableaux so $\sum_{\lambda \vdash n} \left(f^\lambda\right)^2 = n!$* □

Thus armed we look for a basis for S^λ. Some experimentation with the examples above (the treatment of $S^{(3,2)}$ in James [5] is also very instructive) may lead us to suspect that whatever char k,

Theorem 7 *The standard polytabloids, i.e. $\{e_t : t \text{ is a standard } \lambda\text{-tableau}\}$ form a basis of S^λ.*

First of all we will prove that the standard polytabloids are linearly independent. Our proof depends on constructing a partial ordering on the set of λ-tabloids and then showing that a greatest element in this ordering cannot be cancelled by adding a linear combination of other tabloids.

If $\{s\}$ and $\{t\}$ are λ-tabloids with their elements ordered so that they increase along the rows, define $\{s\} \trianglerighteq \{t\}$ if the largest number, m say, occuring in different places in $\{s\}$ and $\{t\}$ occurs earlier in the column word of s than t. The column word of a tableau is found by reading the entries from bottom to top in each column starting in the leftmost column and working rightwards. For example the column words of the tableaux below are: 41523, 31524 and 31425:

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

If we take a standard tableau t and apply any column permutation τ then it is clear then $\{t\} \trianglerighteq \{t\}\tau$ (the greatest element lying in a different place in $\{t\}\tau$ than t must have been moved upwards and rightwards by τ so occurs later in the column word). Hence if t is standard tableau and the tabloid $\{s\}$ appears in e_t then $\{t\} \trianglerighteq \{s\}$.

We can now prove our linear independence theorem. Let the standard polytabloids be $e_{t_1} \ldots e_{t_k}$ where we order the t_i so that if $\{t_i\} \trianglerighteq \{t_j\}$ then $i \geq j$.

Thus $\{t_1\}$ is the greatest in our ordering. Suppose we have a linear relation: $x_1 e_{t_1} + \ldots x_k e_{t_k} = 0$. By the remark above $\{t_i\}$ is maximal amongst the tabloids appearing in e_{t_i}. Furthermore if $\{t_j\}$ occurs in e_{t_i} then $\{t_i\} \trianglerighteq \{t_j\}$ so $i \leq j$. Thus the only way $\{t_1\}$ can be cancelled is if $x_1 = 0$. Our result follows by induction on k.

Theorem 8 *Our linear independence result and the RSK-correspondence now combine to gives us a rapid sequence of corollaries:*
i) $\dim S^\lambda = f^\lambda$
ii) $(e_t : t$ *is a standard tableau) is a basis of* S^λ *(theorem 7).*
iii) Assume $\operatorname{char} k = 0$. *Then* $\{S^\lambda : \lambda \vdash n\}$ *is a complete set of non-isomorphic irreducible representations of* S_n*.*

Proof: Wedderburn's theorem implies that if the irreducible representations of a group G are $V_1 \ldots V_r$ then:

$$|G| = \sum_{i=1}^{r} (\dim V_i)^2 \ .$$

We know that each Specht module S^λ has dimension at least f^λ since it contains f^λ linearly independent polytabloids. We also know that two Specht modules for different partitions are non-isomorphic (theorem 4). Hence:

$$n! \geq \sum_{\lambda \vdash n} \left(\dim S^\lambda\right)^2 \geq \sum_{\lambda \vdash n} \left(f^\lambda\right)^2 = n! \ .$$

where the last equality follows, as seen above, from the RSK-correspondence. We conclude that we must have equality everywhere, so $\dim S^\lambda = f^\lambda$. This proves (i) and (ii). (iii) follows immediately because if there were any other irreducible representations we would not get the first equality. $\square$

Note that we still don't have a formula for the dimension of the Specht modules. This is one of the questions that motivates the next two sections where we develop some of the theory of symmetric functions and give a representation-theory proof of the 'hook formula' for the dimension of S^λ.

4 Background on symmetric functions

Our aim in the next two sections is to relate characteristic zero representations of S_n to the theory of symmetric polynomials. In this section we develop necessary parts of this companion theory. In the subsequent section we apply it to deduce the Frobenius character formula (for the character of S^λ) and, perhaps more importantly, a complete description of the 'representation ring' of S_n.

In section two we saw that $\mathbb{Q}$ is a splitting field for S_n. So once we assume $\operatorname{char} k = 0$ we lose nothing by assuming that $k = \mathbb{Q}$. Accordingly we will deal with representations of $\mathbb{Q}S_n$ and polynomial rings over $\mathbb{Q}$ and $\mathbb{Z}$.

The reader will probably be familiar with the construction of the ring of symmetric polynomials in m variables as a ring of fixed points (classically invariants) in $\mathbb{Z}(x_1, x_2, \dots, x_m)$ under the action of S_m: if $\sigma \in S_m$ then σ acts on monomials in this ring by $x_1^{a_1} x_2^{a_2} \dots x_m^{a_m}\, \sigma = x_{1\sigma}^{a_1} x_{2\sigma}^{a_2} \dots x_{m\sigma}^{a_m}$. It is convenient to write x rather than $x_1, x_2, \dots, x_m$ for the variables in our symmetric polynomials. If λ is a partition of n write $x^\lambda = x_1^{\lambda_1} x_2^{\lambda_2} \dots x_m^{\lambda_m}$ (provided $m \geq k$, the number of rows of λ). If t is a semi-standard λ-tableau let $x^t = x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$ where a_i is the number of times i occurs in t. (Again this assumes we have enough variables: $n \geq m$.) Finally if μ is a partition of n with a_1 parts of size 1, a_2 parts of size 2, on so on, let $z(\mu) = 1^{a_1} a_1! 2^{a_2} a_2! \dots n^{a_n} a_n!$ It is not hard to show that if $C(\mu)$ is the conjugacy class of S_n determined by μ then $|C(\mu)| = z(\mu)/n!$

We have already seen that it can be rather a bother to have to keep track of the number of variables, and later on we will find that the exact number rarely matters provided that there are enough. With this in mind we introduce the next definition:

Definition 9 *A* ***symmetric function of degree n*** *is a collection of symmetric polynomials, $f^i(x_1, x_2 \dots x_i)$ for $i \geq K$ (some fixed K) written simply as f. We require that if $j > i \geq K$ then $f^j(x_1 \dots x_i, 0 \dots 0) = f^i(x_1 \dots x_i)$. Let Λ_n be the collection of all symmetric functions of degree n.*

Informally the last condition means that it doesn't matter which polynomial we choose, provided it has sufficient variables. We add and multiply symmetric functions in the obvious way, by adding and multiplying their associated symmetric polynomials. It is clear that multiplication is a bilinar map $\Lambda_p \times \Lambda_q \to \Lambda_{p+q}$. So with these definitions we obtain $\Lambda = \bigoplus_{n \geq 0} \Lambda_n$ the (graded) ring of symmetric functions. The masterly exposition of Stanley's book *Enumerative Combinatorics II* (Ch. 7) makes it clear that the study of symmetric polynomials is essentially the study of this ring, together with certain extra structures that can be naturally put upon it. This mode of thinking sheds a great deal of light on what might previously have seemed like a mere collection of clever combinatorial tricks.

Naturally enough we start our study of Λ by establishing a basis. In fact we will establish several, in the process looking at all the usual types of symmetric polynomial. This is necessary because the usefulness of a basis for computation is roughly inversely proportional to its real combinatorial or algebraic significance.

Definition 10 *Let $\lambda \vdash n$, $q \in \mathbb{N}$. We will define five types of symmetric polynomial. In each case the associated symmetric functions should be clear.*

i) The ***monomial symmetric polynomial*** *$m_\lambda(x_1, x_2, \dots, x_m)$ is the sum of all monomials arising from $x^\lambda = x_1^{\lambda_1} x_2^{\lambda_2} \dots x_k^{\lambda_k}$ by the action of S_m:*

$$m_\lambda(x) = \sum_{\sigma \in S_m} x^\lambda \sigma.$$

For example $m_{(2,1)}(x_1, x_2, x_3) = x_1^2 x_2 + x_1 x_2^2 + x_1^2 x_3 + x_1^2 x_3^2 + x_2^2 x_3 + x_2 x_3^2$.

If $f(x_1, x_2, \dots, x_m)$ is a symmetric polynomial and the monomial x^λ occurs in f it is immediate that m_λ occurs in f. Thus we see that $(m_\lambda : \lambda \vdash n)$ is a basis for Λ_n and so $\dim \Lambda_n = p(n)$, the number of partitions of n.

ii) The ***complete symmetric polynomial*** *$h_q(x_1, x_2, \dots, x_m)$ is the sum of all monomials arising from products of $x_1, x_2, \dots, x_m$ such that the total degree* *of each monomial is q. Equivalently:*

$$h_q(x) = \sum_{\mu \vdash q} m_\mu(x).$$

It will be convenient to define h_λ for $\lambda \vdash n$ by $h_\lambda = h_{\lambda_1} h_{\lambda_2} \dots h_{\lambda_k}$. For example $h_{(2,1)}(x_1, x_2) = (x_1^2 + x_2^2 + x_1 x_2)(x_1 + x_2)$.

iii) The ***elementary symmetric polynomial*** *$e_q(x_1, x_2, \dots, x_m)$ is the sum of all monomials arising from products of $x_1, x_2, \dots, x_m$ of total degree q where each x_i is used at most once. Thus $e_q(x_1, x_2, \dots, x_m)$ vanishes if $m < q$. Again it will be convenient to define e_λ for $\lambda \vdash n$ by $e_\lambda = e_{\lambda_1} e_{\lambda_2} \dots e_{\lambda_k}$. For example $e_{(2,1)}(x_1, x_2) = (x_1 x_2)(x_1 + x_2)$*

iv) The ***power sum symmetric polynomial*** *$p_q(x_1, x_2, \dots, x_m)$ is simply the sum of the q^{th} powers of its variables:*

$$p_q(x_1, x_2, \dots, x_m) = x_1^q + x_2^q + \dots x_m^q.$$

Once more it will be convenient to define p_λ for $\lambda \vdash n$ by $p_\lambda = p_{\lambda_1} p_{\lambda_2} \dots p_{\lambda_k}$.

v) Finally we define the ***Schur polynomial****, s_λ. Many definitions are possible — naturally we choose the most combinatorial:*

$$s_\lambda(x_1, x_2, \dots, x_m) = \sum x^t$$

where the sum is taken over all semi-standard λ-tableau t with entries from $[1..m]$. Thus $s_\lambda(x)$ vanishes if $m < k$.

Although the Schur polynomials are surely the most fundamental type of symmetric polynomial they may at first seem hard to work with. It is not even obvious that they are symmetric! For example, to calculate $s_{(2,1)}(x_1, x_2, x_3)$ we draw all semi-standard $(2,1)$-tableaux with entries from $[1..3]$:

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

and find the associated monomials: $x_1^2 x_2, x_1^2 x_3, x_1 x_3^2, x_1 x_2^2, x_1 x_2 x_3, x_2^2 x_3$ and $x_2 x_3^2$. Adding them together we obtain:

$$s_{(2,1)}(x_1, x_2, x_3) = x_1^2 x_2 + x_1^2 x_3 + x_1 x_3^2 + x_1 x_2^2 + x_1 x_2 x_3, + x_2^2 x_3 + x_2 x_3^2$$

It is a useful exercise to show that $s_{(n)}(x) = h_n(x)$ and that $s_{(1^n)}(x) = e_n(x)$.

Theorem 11 *The Schur polynomial s_λ is symmetric.*

Proof: Since S_m is generated by the transpositions $(12), (23), \dots, (m-1\, m)$ it is sufficient to prove that $s_\lambda(x_1, x_2, \dots, x_m)(i\, i+1) = s_\lambda(x_1, x_2, \dots, x_m)$ for each $i \in [1..m-1]$. Suppose row j is of the form:

$$\cdots\ \overbrace{\boxed{i}\ \cdots\ \boxed{i}}^{r}\overbrace{\boxed{i+1}\ \cdots\ \boxed{i+1}}^{s}\ \ldots$$

The obvious try would be to replace it with:

$$\cdots\ \overbrace{\boxed{i}\ \cdots\ \boxed{i}}^{s}\overbrace{\boxed{i+1}\ \cdots\ \boxed{i+1}}^{r}\ \ldots$$

but of course this leads to problems when an i occurs above an $i+1$ in the original tableau. Suppose however that we ignore all such columns (so we only look at columns containing either an i or an $i+1$ but not both). Then this swapping process, applied to each row in turn will yield another semi-standard tableau. Applying it twice gives back the original tableau so the process is bijective. For example if $i = 2$ then:

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

Thus $s_\lambda(x)(i\,i+1) = \sum x^t(i\,i+1) = \sum x^{t(i\,i+1)} = \sum x^t = s_\lambda(x)$. □

Now we prove that the symmetric functions we have defined give bases for Λ_n. We have already seen that $(m_\lambda : \lambda \vdash n)$ is a basis for Λ_n. I assume that the reader will be familiar with the analagous fact for the elementary symmetric functions (although perhaps stated in slightly different language). If not, proofs can be found in every basic algebra textbook. There are closely related generating functions for $e_n(x)$ and $h_n(x)$: $\prod_{i=1}^m 1/(1 - x_i t) = \sum_{n\geq 0} h_n(x)t^n$ and $\prod_{i=1}^m (1 - x_i t) = \sum_{n=0}^m (-1)^n e_n(x)t^n$. Multiplying these generating functions and finding the coefficient of t^n gives the identity:

$$\sum_{i=1}^n (-1)^{n-i} h_i(x) e_{n-i}(x) = 0$$

which shows that $\langle h_1(x) \ldots h_q(x)\rangle = \langle e_1(x) \ldots e_q(x)\rangle$. Hence $(h_\lambda : \lambda \vdash n)$ is a basis for Λ_n, also.

We deal with the power sum symmetric polynomials in a similar way. Notice that:

$$\log \sum_{n\geq 0} h_n(x)t^n = \log \prod_{i=1}^m \frac{1}{1 - x_i t} = \sum_{i=1}^m \sum_{j\geq 0} (x_i t)^j = \sum_{j\geq 0} p_j(x)t^j$$

Taking exponentials of each side gives:

$$\begin{aligned} \sum_{n\geq 0} h_n(x)t^n &= \prod_{j\geq 0} \exp\left(p_j(x)t^j\right) \\ &= \prod_{j\geq 0} \sum_{r\geq 0} p_j(x)^r t^{ir}/r! \\ &= \sum_{n\geq 0} \left(\sum_{\mu\vdash n} \frac{p_\mu(x)}{z(\mu)} \right) t^n \end{aligned}$$

where $z(\mu)$ is as defined earlier. Comparing coefficients of t^n shows that:

$$h_n(x) = \sum_{\mu \vdash n} \frac{1}{z(\mu)} p_\mu(x)$$

hence $\Lambda_n = \langle p_\mu : \mu \vdash n \rangle_{\mathbb{Q}}$. Since each side has the same dimension we have proved that $(p_\mu : \mu \vdash n)$ is a $\mathbb{Q}$-basis for Λ_n, or more formally, a $\mathbb{Q}$-basis for the vector space $\Lambda_n \otimes \mathbb{Q}$.

Finally we want to show that the Schur functions give a basis of Λ_n. We will prove this by first giving a rule for multiplying a Schur polynomial by a complete symmetric polynomial.

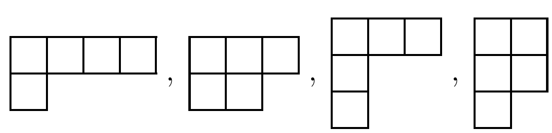
Theorem 12 (Pieri's rule) *Let λ be a partition of n, $q \in \mathbb{N}$. Then:*

$$s_\lambda(x) \, . \, h_q(x) = \sum_\mu s_\mu(x)$$

where the sum is taken over all partitions μ of $n+q$ obtained by adding q boxes to λ, no two in the same column.

Sketch Proof: We know that the product is symmetric. Further it is possible to see that if x^t is a monomial in s_λ and ν is a partition of q then $x^t . x_1^{\nu_1} x_2^{\nu_2} \dots x_k^{\nu_k}$ comes from the tableau s obtained by row inserting first ν_k k's into t, then ν_{k-1} $k-1$'s into the new tableau, on so on until we insert ν_1 1's. Close examination of the results of these row insertions shows that no two boxes are entered in the same column and that every μ tableau figuring in the right hand side above is obtained in this way. (A complete proof is given in Fulton [3] Ch. 2 where the necessary lemma on row bumping is proved.)

For example, to calculate $s_{(2,1)}(x).h_2(x)$ we find all partitions μ obtained from $(2,1)$ by adding two boxes, no two in the same column:



so $s_{(2,1)}(x).h_2(x) = s_{(4,1)}(x) + s_{(3,2)}(x) + s_{(3,1^2)}(x) + s_{(2^2,1)}(x)$.

Anyone familiar with Young's rule (or even just the Branching rule) for representations of S_n will not be surprised to learn that Pieri's rule is essentially the echo of these rules in Λ.

It is now possible to prove that the Schur functions give a basis for Λ_n. Since $h_q(x) = s_{(q)}(x)$ we can expand the product $h_\lambda(x) = h_{\lambda_1}(x) h_{\lambda_2}(x) \dots h_{\lambda_k}(x)$ as a sum of Schur polynomials by first regarding h_{λ_1} as the Schur polynomial $s_{(\lambda_1)}$ and using Pieri's rule to multiply it by h_{λ_2}, then multiplying each of the resulting Schur polynomials by h_{λ_3} on so on. So we can write

$$h_\mu(x) = \sum_\lambda K_{\lambda\mu} s_\lambda(x).$$

We can give a combinatorial interpretation to the coefficients in this expression: $K_{\lambda\mu}$ is the number of semi-standard λ-tabloids t with μ_1 1's, μ_2 2's, on so on. If this is the case we say λ is of **type** μ. It is a simple exercise (uncoincidentally like the proof of theorem 4 above) to show that if $K_{\lambda\mu} \neq 0$ then $\lambda \trianglerighteq \mu$. The coefficients $K_{\lambda\mu}$ are known as the **Kostka numbers**. In the final section we will prove that $K_{\lambda\mu} = k_{\lambda\mu}$, the multiplicity of S^λ in M^μ defined earlier.

We now introduce an additional structure on Λ. Define an inner product $\langle\ ,\ \rangle$ on each Λ_n by insisting that the Schur functions s_λ for $\lambda \vdash n$ form an orthonormal basis: $\langle s_\lambda, s_\mu\rangle = \delta_{\lambda\mu}$. Note that we need to know that the Schur functions form a basis for this definition to make sense. This definition makes Λ a graded inner product space.

The theorem below records how our other symmetric polynomials behave with respect to $\langle\ ,\ \rangle$.

Theorem 13 *Let $\lambda, \mu \vdash n$. Then*
i) $\langle s_\lambda, s_\mu\rangle = \delta_{\lambda\mu}$
ii) $\langle m_\lambda, h_\mu\rangle = \delta_{\lambda\mu}$
iii) $\langle p_\lambda, p_\mu\rangle = \delta_{\lambda\mu}/z(\lambda)$

Proof: Our proof depends on establishing a polynomial identity that is of some independent interest. We work with variables $x_1, x_2, \dots, x_m$ and $y_1, y_2, \dots, y_l$, so for instance $h_n(x_iy_j)$ is the complete symmetric polynomial of degree n in the ml variables x_iy_j. We claim:

$$\sum_{n\geq 0} h_n(x_iy_j) = \prod_{i=1}^{m}\prod_{j=1}^{l}\frac{1}{1-x_iy_j} = \sum_\lambda s_\lambda(x_1,\dots,x_m)s_\lambda(y_1,\dots,y_l) \tag{1}$$

$$= \sum_\lambda m_\lambda(x_1,\dots,x_m)h_\lambda(y_1,\dots,y_l) \tag{2}$$

$$= \sum_\lambda p_\lambda(x_1,\dots,x_m)p_\lambda(y_1,\dots,y_l)/z(\lambda). \tag{3}$$

The inequality equality follows by setting $t=1$ in the generating function for the h_n. (2) and (3) are left as straight-forward exercises. (1) which was originally due to Cauchy and Littlewood can be proved by an attractive application of the RSK-correspondence: on the right hand side we have the sum of all monomials x^ty^s where t and s are semi-standard λ-tableau with entries from $[1..m]$ and $[1..l]$ respectively. By the RSK-correspondence such pairs are in bijective correspondence with lexicographically ordered two row arrays whose top row has entries from $[1..l]$ and whose bottom row has entries from $[1..m]$. Given given any such an array we get a product $\prod(x_iy_j)^{a_{ij}}$ where a_{ij} is the number of times $\binom{i}{j}$ occurs in the array. Since the left hand side is the sum of all such products, over all $(l\times m)$ matrices A with non-negative entries, the two sides agree.

Each of the results in the theorem now follows by expressing the symmetric polynomials involved in terms of Schur polynomials and seeing what the identities imply for the change of basis matrix. For example to prove (iii) let $\sqrt{z(\lambda)}p_\lambda = \sum_\mu a_{\lambda\mu}s_\mu$. Substituting we obtain:

$$\begin{aligned}\sum_\lambda s_\lambda(x)s_\lambda(y) &= \sum_\lambda p_\lambda(x)p_\lambda(y)/z(\lambda)\\ &= \sum_\lambda\left(\sum_\mu a_{\lambda\mu}s_\mu(x)\right)\left(\sum_\nu a_{\lambda\nu}s_\nu(x)\right)\end{aligned}$$

$$= \sum_{\mu\,\nu} (A^{tr}A)_{\mu\nu} s_\mu(x) x_\nu(x).$$

We conclude that A is orthogonal and so:

$$\langle p_\lambda, p_\mu \rangle = \langle s_\lambda, s_\mu \rangle / \sqrt{z(\lambda)z(\mu)}$$

which implies (iii). $\square$

Finally we define two further change of basis matrices that will be of special importance to us. It will transpire that their coefficients are certain character values, hence the choice of notation.

Definition 14 *Let λ, μ be partitions of n Define coefficients $\xi_{\lambda\mu}$ and $\chi_{\lambda\mu}$ by:*

$$\begin{aligned} p_\mu &= \sum_{\lambda \vdash n} \xi_{\lambda\mu} m_\lambda \\ p_\mu &= \sum_{\lambda \vdash n} \chi_{\lambda\mu} s_\lambda. \end{aligned}$$

It is a simple exercise in using our orthogonality relations to deduce the equivalent forms:

$$\begin{aligned} h_\lambda &= \sum_{\mu \vdash n} \frac{\xi_{\lambda\mu}}{z(\mu)}\, p_\mu \\ s_\lambda &= \sum_{\mu \vdash n} \frac{\chi_{\lambda\mu}}{z(\mu)}\, p_\mu. \end{aligned}$$

Well this section has perhaps been rather a slog. Rest assured the final section offers ample payback!

5 The Frobenius character formula

Our first task in this section is to define the representation ring of S_n. Next we will prove this ring is isomorphic (in fact isometric) with Λ, the ring of symmetric functions defined in the previous section. Using this result we will deduce the Frobenius character formula, and after a bit more work the much heralded hook formula for the dimension of S^λ.

Given that we constructed S^λ as a submodule of M^λ it is natural to look first at the character of $M^\lambda = \mathrm{Ind}_{S_\lambda}^{S_n} k$. We will write $C(\mu)$ for the conjugacy class of S_n given by the partition $\mu \vdash n$. Let g be an element in $C(\mu)$. The formula for induced characters gives:

$$\chi_{M^\lambda}(C(\mu)) \;=\; \frac{1}{|S_\lambda|} \sum_{\substack{x \in G \\ x^{-1}gx \in S_\lambda}} 1 \;=\; \frac{|G : C(\mu)|}{|S_\lambda|}\, |C(\mu) \cap S_\lambda|$$

as each member of $C(\mu) \cap S_\lambda$ will occur $|G : C(\mu)|$ times.

Finding $|C(\mu) \cap S_\lambda|$ is slightly harder. Suppose g has a_i cycles of length i. Recall that:

$$|C(\mu)| = n!/z(\mu) = n!/1^{a_1}a_1!2^{a_2}a_2!\dots n^{a_n}a_n!$$

If $x \in C(\mu) \cap S_\lambda$ each of its cycles must lie entirely within one of the factors of $S_\lambda = S_{\{1,\dots\lambda_1\}} \times \dots \times S_{\{\lambda_{k-1}+1\dots n\}}$. The number of ways we can fit x_{ij} cycles of length a_i into λ_j is:

$$\sum \frac{\lambda_j!}{1^{x_{1j}}x_{1j}!2^{x_{2j}}x_{2j}!\dots n^{x_{nj}}x_{nj}!}$$

where the summation is taken over all x_{ij} satisfying $\sum_{i=1}^n ix_{ij} = \lambda_j$. Looking at all parts of λ together we obtain:

$$|C(\mu) \cap S_\lambda| = \sum \prod_{j=1}^k \frac{\lambda_j!}{1^{x_{1j}}x_{1j}!2^{x_{2j}}x_{2j}!\dots n^{x_{nj}}x_{nj}!}$$

where the sum is over all sequences x_{ij} satisfying $\sum_{i=1}^n ix_{ij} = \lambda_j$ for $j = 1\dots k$ and $\sum_{j=1}^k x_{ij} = a_i$ for $i = 1\dots n$. Substituting and tidying up we arrive at the final answer:

$$\chi_{M^\lambda}(C(\mu)) = a_1!a_2!\dots a_n! \sum \prod_{j=1}^k \frac{1}{x_{1j}!x_{2j}!\dots x_{nj}!} = \sum \prod_{i=1}^n \frac{a_i!}{x_{i1}!x_{i2}!\dots x_{ik}!}$$

with the same conditions on the sum. Notice that the right hand side number is the coefficient of $x^\lambda = x_1^{\lambda_1}x_2^{\lambda_2}\dots x_k^{\lambda_k}$ in the power sum symmetric polynomial $p_\mu(x_1, x_2, \dots, x_k)$ which we defined earlier to be $\xi_{\lambda\mu}$. Thus:

$$\chi_{M^\lambda}(C(\mu)) = \xi_{\lambda\mu} = \left\langle m^\lambda, p^\mu \right\rangle .$$

Next we look closely at M^λ and how S^λ sits inside it. To do this we use a tool well adapted to the task: let $R_n = K_0(\mathbb{Q}S_n\text{–}\mathbf{mod})$ be the Grothendieck group of $\mathbb{Q}S_n$. We will write the isomorphism class of the $\mathbb{Q}S_n$ module V as $[V]$. Paralleling our construction of Λ we take R_n to be the n^{th} part of a graded ring $R = \bigoplus_{n\geq 0} R_n$.

Given our previous results we know that R_n has a basis consisting of the isomorphism classes of each Specht module, $[S^\lambda]$. The next theorem gives us an alternative basis:

Theorem 15 *$([M^\lambda] : \lambda \vdash n)$ is a basis for R_n.*

Proof: We proved earlier that $M^\mu = S^\mu \oplus \bigoplus_{\lambda \rhd \mu} k_{\lambda\mu} S^\lambda$. The corresponding relation in R_n is:

$$[M^\mu] = [S^\mu] + \sum_{\lambda \rhd \mu} k_{\lambda\mu}[S^\lambda]$$

Thus the transition matrix from $([S^\lambda] : \lambda \vdash n])$ to $([M^\lambda] : \lambda \vdash n)$ is just $k_{\lambda\mu}$ which because $k_{\lambda\mu}$ is lower triangular with 1's on the diagonal, has an inverse with integer coefficients. The theorem follows. $\square$

We use this theorem to define a product $\circ : R_p \times R_q \to R_{p+q}$ by:

$$[M^\lambda] \circ [M^\mu] = \left[\mathrm{Ind}_{S_p \times S_q}^{S_{p+q}} M^\lambda \otimes M^\mu\right]$$

where $M^\lambda \otimes M^\mu$ is regarded as a $\mathbb{Q}(S_p \times S_q)$ module in the obvious way. It is a straightforward exercise in properties of Ind and $\otimes$ to show that $\circ$ is well defined, commutative and associative. For example, to prove associativity, use the description of M^λ as an induced module to show that if $\lambda \vdash p$, $\mu \vdash q$ and $\nu \vdash r$ then:

$$\big([M^\lambda] \otimes [M^\mu]\big) \otimes [M^\nu] = \left[\mathrm{Ind}_{S_p \times S_q \times S_r}^{S_{p+q+r}} \mathrm{Ind}_{S_\lambda \times S_\mu \times S_\nu}^{S_p \times S_q \times S_r} k\right] = [M^\lambda] \otimes ([M^\mu] \otimes [M^\nu]).$$

Finally we define an inner product on R_n by

$$\left\langle [S^\lambda], [S^\mu] \right\rangle = \left\{ \begin{array}{lll} 1 & \text{if} & [S^\lambda] = [S^\mu] \\ 0 & \text{if} & [S^\lambda] \neq [S^\mu] \end{array} \right. .$$

Our notation is suggestive of the close link between Λ and R and between s^λ and S^λ (although the latter may be entirely due to the fortunate coincidence that both 'Specht' and 'Schur' start with an 'S'). The next theorem makes this explicit.

Theorem 16 *Define $\phi : \Lambda \to R$ by $h_\lambda \phi = [M^\lambda]$. Then*
i) ϕ is a well defined ring isometry.
ii) $s_\lambda \phi = [S^\lambda]$.

Proof: i) By our results on bases for Λ and R we know that ϕ is a well defined, additive and bijective. To show $(h_\lambda h_\mu)\phi = h_\lambda \phi \circ h_\mu \phi$ it is sufficient to look at the case where $\lambda = (p)$ and $\mu = (q)$ since any complete symmetric polynomial is a product of such polynomials. Without lose of generality take $p \geq q$. Then:

$$(h_\lambda \phi)(h_\mu \phi) = \mathrm{Ind}_{S_p \times S_q}^{S_{p+q}} \left(M^{(p)} \otimes M^{(q)}\right) = \mathrm{Ind}_{S_p \times S_q}^{S_{p+q}} k = M^{(p,q)} = (h_\lambda h_\mu)\phi.$$

Thus ϕ is also a homomorphism.

It remains to show that ϕ is an isometry. To do this we look at the inverse of ϕ, ψ say. Since we will work with power sum symmetric polynomials we consider ψ as a map: $R \otimes \mathbb{Q} \to \Lambda \otimes \mathbb{Q}$, so that $\phi\psi$ is the inclusion $\Lambda \hookrightarrow \Lambda \otimes \mathbb{Q}$. Since $h_\lambda \phi = [M^\lambda]$ we know that $[M^\lambda]\psi = h_\lambda$. Expressing h_λ in terms of the p_μ (see definition 14) we get:

$$[M^\lambda]\psi = \sum_{\mu \vdash n} \frac{\xi_{\lambda\mu}}{z(\mu)} p_\mu.$$

But as $\xi_{\lambda\mu} = \chi_{M^\lambda}(C(\mu))$ this implies

$$[V]\psi = \sum_{\mu \vdash n} \frac{1}{z(\mu)} \chi_V(C(\mu)) p_\mu$$

for any $\mathbb{Q}S_n$ module V. Now we calculate $\langle [V]\psi, [W]\psi\rangle$:

$$\begin{aligned}
\langle [V]\psi, [W]\psi\rangle &= \sum_{\mu,\nu\,\vdash n} \frac{1}{z(\mu)z(\nu)} \chi_V(C(\mu))\,\chi_W(C(\nu))\,\langle p_\mu, p_\nu\rangle \\
&= \sum_{\xi\vdash n} \chi_V(C(\xi))\,\chi_W(C(\xi)) \\
&= \langle [V], [W]\rangle
\end{aligned}$$

where we have used $\langle p_\lambda, p_\mu\rangle = \delta_{\lambda\mu}/z(\lambda)$ (theorem 14) and orthogonality of characters. Thus ψ is an isometry and so is ϕ.

ii) Suppose $s_\mu\phi = \sum_\nu c_{\mu\nu}[S^\nu]$ for some integers $c_{\mu\nu}$. Then as ϕ is an isometry, $\sum_\nu c_{\mu\nu}^2 = 1$ so $s_\mu\phi = [S^\nu]$ for some ν. We now show this ν has to be μ.

We start with the partition $\lambda = (n)$ which is greatest under $\trianglerighteq$. In this case $s_\mu\phi = [S^\mu]$ is certainly true because $s_{(n)} = h_{(n)}$ and $S^{(n)} \cong M^{(n)}$.

Suppose we know that $s_\lambda\phi = [S^\mu]$ for all partitions $\lambda \triangleright \mu$. Applying ϕ to the relation $h_\mu = s_\mu + \sum_{\lambda\triangleright\mu} K_{\lambda\mu}S_\lambda$ (from section 4) gives:

$$h_\mu\phi = [M^\mu] = s_\mu\phi + \sum_{\lambda\triangleright\mu} K_{\lambda\mu}[S^\lambda].$$

But we already know (from section one) that:

$$[M^\mu] = [S^\mu] + \sum_{\lambda\triangleright\mu} k_{\lambda\mu}[S^\lambda].$$

so the only way we can get $[S^\mu]$ in $h_\mu\phi$ is if it features in $s_\mu\phi$. We have already remarked that $s_\mu\phi = [S^\nu]$ for some ν. So we must have $s_\mu\phi = [S^\mu]$.

We have shown that if $s_\lambda\phi = [S^\mu]$ for all partitions $\lambda \triangleright \mu$ then $s_\lambda\phi = [S^\mu]$ for all partitions $\lambda \trianglerighteq \mu$. So we prove the theorem by inductively working our way down the lattice diagram for $\trianglerighteq$. $\square$

Using ϕ to apply our work on symmetric polynomials to the (characteristic zero) representation theory of S_n gives several quick corollaries:

Corollary 17 (Young's Rule) *The multiplicity of S^λ in M^μ is the number of semi-standard tableaux of shape λ and type μ. (Use ϕ to show $k_{\lambda\mu}$ is the Kostka number $K_{\lambda\mu}$.)* $\square$

Corollary 18 (Branching Rule) *Let λ be a partition of n. Then*

$$\mathrm{Ind}_{S_n}^{S_{n+1}} S^\lambda = \bigoplus_\mu S^\mu$$

where the sum is over all partitions μ obtained from λ by adding a single box. (Apply ϕ to Pieri's rule for $q = 1$). $\square$

By transitivity of induction the branching rule has an obvious generalisation to induction from S_n to S_{n+q} for any $q \in \mathbb{N}$. Also, by Frobenius reciprocity an analagous result in terms of the removal of boxes can be obtained for the restriction of a Specht module from S_n to S_m for any $m < n$.

One possible next step would be to deduce the Littlewood-Richardson rule for characteristic zero representations of S_n. (This rule includes the branching rule and Pieri's rule as special cases). One of the reasons for not continuing to cover it (apart from the fact this account is already too long) is that all these results are true with at most minor modifications for representations in positive characteristic, and in the end, it is the module theoretic proofs that are of greater significance. The reader wishing to pursue the subject further in this direction without using the modular theory is referred to Fulton [3] Ch. 5 where the corresponding results for symmetric polynomials are proved.

Obviously the same cannot be said about the Frobenius character formula since by definition it is a characteristic zero result. We have in fact already found $\chi_{S^\lambda}(C(\mu))$ it is the integer $\chi_{\lambda\mu} = \langle s_\lambda, p_\mu \rangle$ defined earlier as part of the transition matrix between s_λ and p_μ (see definition 14). For applying ψ (the inverse map to ϕ) to $[S^\lambda]$ shows that $[S^\lambda]$ corresponds both to the element $\sum_\mu \chi_{S^\lambda}(C(\mu))/z(\mu) p_\mu$ in Λ and to $s_\lambda = \sum_\mu \chi_{\lambda\mu}/z(\mu)\, p_\mu$.

To give a more convenient description of $\chi_{S^\lambda}(C(\mu))$ we need an alternative description of the Schur polynomial $s_\lambda(x)$:

Lemma 19 (Jacobi-Trudi Formula) *Let λ be a partition of n with k rows. We will work with k variables, $x_1, x_2, \ldots, x_k$. Define $\Delta_\lambda(x)$ to be the determinant of the $k \times k$ matrix A where $A_{ij} = x_i^{\lambda_j + k - j}$. Let $\Delta_0(x)$ be the Vandermonde determinant* $\det\left(x_i^{j-1}\right) = \prod_{i<j}(x_i - x_j)$. *Then:*

$$s_\lambda(x_1, x_2, \ldots, x_k) = \frac{\Delta_\lambda(x)}{\Delta_0(x)}.$$

We will not prove this result here. An elegant combinatorial proof is given in Stanley [8] p335 that uses slightly more of the theory of symmetric polynomials than we have developed here.

Theorem 20 (Frobenius Character Formula) *Let λ be a partition of n with k rows. Let $l_i = \lambda_i + k - i$. Then the character of S^λ on the conjugacy class $C(\mu)$, $\chi_{\lambda\mu}$ is the coefficient of $x_1^{l_1} x_2^{l_2} \ldots x_k^{l_k}$ in the polynomial:*

$$\prod_{1 \le i < j \le k} (x_i - x_j) \,.\, p_\mu(x_1, x_2, \ldots, x_k)$$

Proof: Applying the Jacobi-Trudi formula to the equation $p_\mu(x) = \sum_\lambda \chi_{\lambda\mu} s_\lambda(x)$ gives:

$$\prod_{1 \le i < j \le k} (x_i - x_j) p_\mu(x) = \sum_\lambda \chi_{\lambda\mu} \Delta_\lambda(x)$$

The theorem follows by looking at the coefficient of $x_1^{l_1} x_2^{l_2} \ldots x_k^{l_k}$ on each side. □

We now find $\dim S^\lambda$ by evaluating the character of S^λ on the identity. By the character formula $\dim S^\lambda$ is the coefficient of x^l in:

$$\prod_{1 \le i < j \le n} (x_i - x_j)\, (x_1 + \ldots + x_k)^n$$

We expand the Vandermonde determinant as

$$\Delta_0(x) = \sum_{\sigma \in S_k} \epsilon(\sigma) x_1^{1\sigma - 1} x_2^{2\sigma - 1} \ldots x_k^{k\sigma - 1}.$$

The other term is:

$$(x_1 + \ldots + x_k)^n = \sum \frac{n!}{r_1! r_2! \ldots r_k!} x_1^{r_1} x_2^{r_2} \ldots x_k^{r_k}$$

where the sum is over all k-tuples $r_1, r_2, \ldots, r_k$ such that $\sum_{i=1}^k r_i = n$. To find the coefficient of x^l in the product we pair off terms in these two sums obtaining:

$$\sum \epsilon(\sigma) \frac{n!}{(l_1 - (1\sigma - 1))! \ldots (l_k - (k\sigma - 1))!}$$

where the sum is taken over all σ in S_k such that $l_i - (k - (i-1))\sigma + 1 \geq 0$ for $1 \leq i \leq k$. We can sum instead over *all* permuations in S_k provided we take care to cancel all terms not appearing in the denominator above. This leads to:

$$\frac{n!}{l_1! l_2! \ldots l_k!} \sum_{\sigma \in S_k} \epsilon(\sigma) \prod_{j=1}^{k} l_j(l_j - 1) \ldots (l_j - (k - j + 1)\sigma + 2)$$

in which we can recognise the summation as the expansion of the determinant:

$$\begin{vmatrix} 1 & l_1 & l_1(l_1 - 1) & \ldots \\ 1 & l_2 & l_2(l_2 - 1) & \ldots \\ \vdots & \vdots & \vdots & \ddots \\ 1 & l_k & l_k(l_k - 1) & \ldots \end{vmatrix}.$$

By column reduction this determinant reduces to the Vandermonde determinant $\Delta_0(x)$ so we obtain:

$$\dim S^\lambda = \frac{n!}{l_1! l_2! \ldots l_k! \prod_{1 \leq i < j \leq k} (l_i - l_j)}$$

where as usual $l_i = \lambda_i + k - i$.

We now introduce the concept of a **hook** in a partition λ. The hook on a given node consists of that node, together with all boxes directly to the right and directly below it. The **hook-length** of a hook is simply the number of boxes it involves. In the diagram below each box is labelled with its hook-length:

9	7	5	4	2	1
6	4	2	1		
3	1				
1					

Notice that the numbers l_i defined above are precisely the hook-lengths of nodes in the first column of λ. If α is a node of λ write h_α for the hook-length of the hook on α. We will deduce the remarkable:

Corollary 21 (Hook formula)

$$\dim S^\lambda = \frac{n!}{\prod_{\alpha\in\lambda} h_\alpha}$$

Proof: We need to show that:

$$\frac{n!}{l_1!l_2!\dots l_k!\prod_{1\le i<j\le k}} = \frac{n!}{\prod_{\alpha\in\lambda} h_\alpha}$$

The original and elegant proof (dating from 1954) can be found in [2]. Alternatively one can induct on the number of column of λ. If λ has only one column then $\lambda = (1^n)$ and we obtain:

$$\dim S^{(1^n)} = \frac{n!}{n!(n-1)!\dots 1!}\prod_{i=1}^{n-1}(n-i)! = 1$$

which agrees with the hook formula (and with the fact that $S^{(1^n)} \cong \epsilon$).

In general let $\lambda^\star$ be the partition of $n-k$ obtained by deleting the first column of λ. Let $k^\star$ be the number of rows of $\lambda^\star$ and let $l_i^\star$ be the hook length of the i^{th} in the first column of $\lambda^\star$ (for $i \in [1..k^\star]$). It is easy to see that $l_i^\star - l_{i+1}^\star = l_i - l_{i+1}$ and that $l^\star + i = l_i - (k-k^\star) - 1$. So by induction we obtain:

$$\begin{aligned}\Rightarrow \dim S^\lambda &= \frac{n!}{l_1\dots l_k}\,\frac{\prod(l_i-l_j)}{(l_i-1)!\dots(l_k-1)!}\\ &= \frac{n!}{\prod_{\alpha\in\lambda}h_\alpha}\,\frac{l_1^\star!\dots l_k^\star!}{(l_i-1)!\dots(l_k-1)!}\prod_{\substack{1\le i<j\le k\\ j>k^\star}}(l_i-l_j)\end{aligned}$$

where the last equality follows by the inductive hypothesis. So it is sufficient to prove that the final terms cancel. Notice that if $i > k^\star$, $l_i = k-i+1$. Thus:

$$\begin{aligned}\prod_{\substack{1\le i<j\le k\\ j>k^\star}}(l_i-l_j) &= \prod_{i=1}^{k-1}\prod_{j>\max(i,k^\star)}(l_i-l_j)\\ &= \prod_{i=1}^{k-1}\prod_{j>\max(i,k^\star)}\frac{(l_i-1)!}{(l_i-1-k+\max(i,k^\star))!}\\ &= \prod_{i=1}^{k^\star}\frac{(l_i-1)!}{l_i^\star!}\prod_{i=k^\star+1}^{k}(l_i-1)!\end{aligned}$$

The hook formula follows. □

The hook formula is an immensely powerful combinatorial tool that subsumes many apparently unrelated combinatorial results. For example the number of standard (n,n) tableaux is the Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$. One nice way to see this is to observe that f^λ is the number of ways to count the votes in an election between k candidates with candidate A getting λ_1 votes, B getting λ_2 votes on so on, *such that*, at any time once counting has begun (except possibly the end) A leads B, B leads C on so on. For given a standard tableau, if the number

i appears in the j^{th} row, awarding vote i to candidate j leads to such a ballot sequence. Conversely any ballot sequence satisfying the condition above leads to a standard tableau by putting the number i in row j if the i^{th} vote counted went to candidate j. As one of the many combinatorial interpretations of the Catalan numbers is as the number of such ballot sequences for two candidates getting equally many votes we have shown that $C_n = f^{(n,n)}$. Attempting to generalise other combinatorial interpretations of the Catalan numbers to give other combinatorial interpretations of f^{λ} might well be profitable. Certainly there will be no shortage of examples to try: Stanley alone gives more than 60 combinatorial interpretations of the Catalan numbers in his book [8].

I conclude by mentioning a few points about R and Λ that weren't needed for our work above. The reader may be familiar with the theorem that if char $k = 0$ then $S^{\lambda'} \cong S^{\lambda} \otimes \epsilon$ (where λ' is the conjugate partition to λ). As a consequence, 'tensor with ϵ' is an involution on R. Its analogue in Λ, usually written ω turns out to swap the complete and elementary symmetric functions: $h_\lambda \omega = e_\lambda$ and $e_\lambda \omega = h_\lambda$. The power sum symmetric functions are its eigenvectors.

The restriction functor gives a map $R_{m+n} \to R_m \otimes R_n$ that makes R (and so Λ) into a Hopf algebra. Another remark in this line is that all our work on R and Λ can be interpreted in a more general way via the machinery of λ-rings. *λ-Rings and the Representation Theory of the Symmetric Groups* by Donald Knuton [6] continues the story in this direction.

6 Sources

The first two sections are based on my Part III essay *Young Tableaux and Modular Representations of the Symmetric Group* (2000, unpublished). It in turn drew heavily from Gordon James' monograph *The Representation Theory of the Symmetric Groups* [5]. An easier and more combinatorial introduction, to the characteristic zero theory only, can be found in Sagan *The Symmetric Group* [7].

I'm fairly sure the idea of using the RSK correspondence to prove the standard polytabloids spanned S^{λ} occurred to me independently but I've since seen it mentioned, at least briefly, in Cameron *Permutation Groups* [1] and Fulton *Young Tableaux* [3] and seen references tracing it back to at least 1981. Neither of these books mention that it can be used to deduce that the Specht modules form a complete set of ordinary irreducible representations of S_n but perhaps both authors thought this was obvious from the earlier application. Fulton's book is an excellent introduction to young tableaux but given it's focus the reader wlll have to work through a lot of material to get to the applications to representation theory. The account given here of the Frobenius character formula is based on its chapter 7. Material on symmetric polynomials comes both from it and from Stanley's recent book *Enumerative Combinatorics II* [8] (which also gives a good description of the RSK correspondence and an entertaining discussion of the Catalan numbers).

The original proof of the hook formula was from the Frobenius character formula in a 1954 paper by Frame, Robinson and Thrall [2]. Where we have proceeded by brute force the authors use a clever combinatorial argument. The world had to wait till 1979 for the first truly combinatorial of the hook formula, which proceeded by a clever probability argument, in a paper by C. Greene, A. Nijenhuis, H. S. Wilf, *A probabilistic proof of a formula for the number of Young Tableaux of a given shape* [4]. Since then purely bijective proofs have been discovered although all the ones I have seen are somewhat longer than the one in this paper. There is also a very uninspiring inductive proof not using representation theory but I'm sure none of us will be content with that!

References

[1] Peter Cameron *Permutation Groups*, London Mathematical Society Student Texts 45, 1995.

[2] J. S. Frame, G. de B. Robinson, R. M. Thrall *The Hook Graphs of the Symmetric Group*, Canad. J. Math, 6, 316-324, 1954.

[3] W. Fulton *Young Tableaux*, London Mathematical Society Student Texts 35, 1997.

[4] C. Greene, A. Nijenhuis, H. S. Wilf *A probabilistic proof of a formula for the number of Young Tableaux of a given shape*. Adv. Math. 31, 104-109, 1979.

[5] G. D. James *The Representation Theory of the Symmetric Groups*, Springer-Verlag Lecture Notes in Mathematics, 1978.

[6] Donald Knutson *λ-rings and the Representation Theory of the Symmetric Group*, Springer-Verlag Lecture Notes in Mathematics, 1970.

[7] B. E. Sagan *The symmetric group*, Wadsworth & Brooks/Cole, 1991.

[8] Richard P. Stanley *Enumeriative Combinatorics II*, Cambridge Studies in Advanced Mathematics 62, 1999.