REPRESENTATION THEORY OF $\mathfrak{S}_n$

Steven Byrnes

Abstract

We use Young tableaux and Young symmetrizers to classify the irreducible representations over $\mathbb{C}$ of the symmetric group on n letters, $\mathfrak{S}_n$. We establish an isomorphism between a ring of all representations of finite symmetric groups and the ring of symmetric functions, and, as a corollary, prove the Frobenius formula for the characters of representations of $\mathfrak{S}_n$. We use the theory thus developed to characterize the representations of the Lie algebra $\mathfrak{sl}_n$.

1 Irreducible Representations of $\mathfrak{S}_n$

1.1 Young tableaux

In any finite group, the number of conjugacy classes equals the number of irreps[1]. But the symmetric group $\mathfrak{S}_n$ (i.e. the group of permutations of n objects, under composition) is one of the few where there is a natural one-to-one correspondence between the two. As a first step in defining this correspondence, we define a *partition* of a positive integer n to be a nonincreasing sequence of positive integers summing to n. That is, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)$ is a partition of n iff $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 1$ and $\sum_{i=1}^m \lambda_i = n$. When convenient, we will allow there to be 0s at the end of a partition, with the understanding that $(\lambda_1, \ldots, \lambda_m, 0, \ldots, 0) = (\lambda_1, \ldots, \lambda_m)$. The partitions of n are in one-to-one correspondence with conjugacy classes of $\mathfrak{S}_n$: the partition $(\lambda_1, \cdots, \lambda_m)$ is associated with the conjugacy class of permutations which, written in cycle notation, contain cycles of length $\lambda_1, \lambda_2, \cdots, \lambda_m$. If λ is a partition of n, we write $\lambda \vdash n$ or $|\lambda| = n$.

The next step is to associate each partition with a Young diagram. A *Young diagram* consists of rows of square boxes, with each row left-justified, and with each row no longer than the row above[2]. The Young diagram associated with a partition $\lambda = (\lambda_1, \cdots, \lambda_m)$ is the one with λ_i boxes in the ith row from the top. Abusing notation, we use λ to denote the Young diagram in addition the partition, when there is no risk of confusion.

Recall that the *regular representation* of a group G is the group algebra $\mathbb{C}[G]$, with

[1] For the purposes of this paper, an "irrep" denotes an irreducible representation over $\mathbb{C}$.

[2] This is the *English convention*; diagrams written using the *French convention* have the smallest row on top.

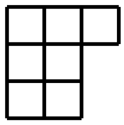

Figure 1: The Young diagram of $(3, 2, 2)$

G-action defined by left-multiplication, and that each irrep of G appears as a subrepresentation of the regular representation. Thus each irrep of $\mathfrak{S}_n$ is associated with an $\mathfrak{S}_n$-invariant subspace of $\mathbb{C}[\mathfrak{S}_n]$, and to define the irrep it suffices to define a map into one of its associated subspaces.

Starting with a Young diagram $\lambda \vdash n$, number the boxes 1 to n, left-to-right top-to-bottom. Let $\mathfrak{S}_n$ permute the numbers in the boxes, and define the subgroups $A_\lambda, B_\lambda \subset \mathfrak{S}_n$ to be the permutations preserving rows and the permutations preserving columns, respectively. Now define two elements of $\mathbb{C}[\mathfrak{S}_n]$ by $a_\lambda = \sum_{g \in A_\lambda} g$ and $b_\lambda = \sum_{g \in B_\lambda} (\text{sign}(g)) g$ (where $\text{sign}(g)$ is the determinant of the associated permutation matrix). The *Young symmetrizer* of λ is defined to be

$$c_\lambda = a_\lambda b_\lambda.$$

1.2 Characterization of the irreps of $\mathfrak{S}_n$

The fundamental theorem is the following:

Theorem 1. *For $\lambda \vdash n$, the image of the map $\varphi_\lambda : \mathbb{C}[\mathfrak{S}_n] \to \mathbb{C}[\mathfrak{S}_n]$, $x \mapsto xc_\lambda$ is an irrep of $\mathfrak{S}_n$. As λ varies over partitions of n, each irrep of $\mathfrak{S}_n$ (up to isomorphism) is generated exactly once.*

Proof. We start by putting an ordering on the set of Young diagrams of size n. For $\lambda = (\lambda_1, \ldots, \lambda_m)$ and $\eta = (\eta_1, \ldots, \eta_m)$, we say $\lambda > \eta$ iff there is a $j \geq 1$ with $\lambda_i = \eta_i$ for $1 \leq i < j$ and $\lambda_j = \eta_j$. Next we define an *Young tableau* as follows: if $\lambda \vdash n$ is a Young diagram, we make a Young tableau with *shape* λ by putting a number in each box from 1 to n with no repeats. Note that $\mathfrak{S}_n$ operates on tableaux of a given size by permuting the numbers. Now we prove a useful lemma about Young tableaux.

Lemma 1. *Suppose that S and T are Young tableaux with shape σ and τ respectively (where $|\sigma| = |\tau| = n$), and suppose $\sigma \geq \tau$. Then* either *we can find two distinct integers which are in the same row of S and the same column of T;* or *$\lambda = \lambda'$ and there is a $p \in \mathfrak{S}_n$ preserving rows of S and a $q \in \mathfrak{S}_n$ preserving columns of T, with $pS = qT$.*

Proof. Suppose that there are no two distinct integers in the same row of S and the same column of T. Then each of the σ_1 numbers in the first row of S is in a different column of T. In particular, T must have at least σ_1 different columns, and since $\sigma > \tau$, T has exactly $\sigma_1 = \tau_1$ columns, and there is a $q_1 \in \mathfrak{S}_n$ preserving columns of T such that the top row of q_1T has the same elements as the top row of S. Likewise, examining the second row of S, we see that $\sigma_2 = \tau_2$ and there is a $q_2 \in \mathfrak{S}_n$ fixing the top row of T and preserving columns of T, such that the second row of q_2q_1T has the same elements as the second row of S. Repeating this process, and noting that $(q_mq_{m-1}\cdots q_1)$ still preserves columns of T, we have proven the lemma. $\square$

We define the *canonical* tableau of a given shape to be the one numbered in order left-to-right, top-to-bottom; this is the tableau we used implicitly when defining A_λ and B_λ. To complete the proof of Theorem 1, we need four more straightforward results.

Lemma 2. *For $p \in A_\lambda$ and $q \in B_\lambda$, $pa_\lambda = a_\lambda p = p$ and $(\mathrm{sign}(q)q)b_\lambda = b_\lambda(\mathrm{sign}(q)q) = b_\lambda$.*

Proof. This follows from writing out the summation definition of a_λ and b_λ. $\square$

Lemma 3. *The set of elements $x \in \mathbb{C}[\mathfrak{S}_n]$ such that $px(\mathrm{sign}(q)q) = x$ for all $p \in A_\lambda$, $q \in B_\lambda$, is $\mathbb{C}c_\lambda$.*

Proof. One inclusion follows from Lemma 2. For the opposite inclusion, suppose that $x = \sum_{g\in\mathfrak{S}_n} \alpha_g g$ satisfies $px(\mathrm{sign}(q)q) = x$. Since the set of such x's is closed under addition and negatives, we may subtract a multiple of c_λ to assume WLOG that $\alpha_e = 0$ (where e denotes the identity in $\mathfrak{S}_n$). Then for $p \in A_\lambda$, $q \in B_\lambda$, it follows from $px(\mathrm{sign}(q)q) = x$ that $\alpha_{pq} = \alpha_e \,\mathrm{sign}(q) = 0$. And for g that cannot be written in the form pq, we have by Lemma 1 that there are two elements in the same row of T and the same column of gT (where T is the canonical tableau with shape λ). Let $p \in A_\lambda$ be the transposition of those two elements and let $q = g^{-1}pg \in B_\lambda$. Then it follows from $px(\mathrm{sign}(q)q) = x$ that $\alpha_g \,\mathrm{sign}\, q = \alpha_{pgq} = \alpha_g$, and since $\mathrm{sign}\, q = \mathrm{sign}\, p = -1$, we have $\alpha_g = 0$. So we have proven that $\alpha_g = 0$ for all g, so indeed only scalar multiples of c_λ satisfy the given identity. $\square$

For any tableau T with shape λ, we can define b_T in a way analogous to b_λ, but starting from the tableau T instead of the canonical tableau T_λ (in particular, $b_\lambda = b_{T_\lambda}$). Note that $xb_Tx^{-1} = b_{xT}$.

Lemma 4. *If $\lambda > \eta$, then $c_\lambda x c_\eta = 0$ for all $x \in \mathbb{C}[\mathfrak{S}_n]$.*

Proof. First we note that for any S of shape λ and T of shape η, by Lemma 1, there are two elements in the same row of S and the same column of T. Letting $s \in \mathfrak{S}_n$ switch those two elements, we have $a_Sb_T = (a_Ss)(sb_T) = (a_S)(-b_T)$ by Lemma 2, so $a_Sb_T = 0$. For x in the subgroup $\mathfrak{S}_n$ of $\mathbb{C}[\mathfrak{S}_n]$, we have $a_\lambda x b_\eta = a_{T_\lambda} b_{xT_\eta} x = 0x = 0$. By taking linear combinations of these, we get $a_\lambda x b_\eta = 0$ for all $x \in \mathbb{C}[\mathfrak{S}_n]$. In particular, $a_\lambda(b_\lambda x a_\eta)b_\eta = 0$, proving the lemma. $\square$

Lemma 5. *Suppose a finite group G acts on $\mathbb{C}[G]$ by left-multiplication, and suppose $\nu : \mathbb{C}[G] \to \mathbb{C}[G]$ is a G-invariant $\mathbb{C}$-endomorphism. Then ν is equivalent to right-multiplication by an element of $\mathbb{C}[G]$.*

Proof. For $g \in G$, we have $\nu(g) = \nu(g(1)) = g(\nu(1))$. Then by $\mathbb{C}$-linearity of ν, we have $\nu(x) = x\nu(1)$ for general $x \in \mathbb{C}[G]$. □

Now we can finish the proof of Theorem 1. For $\lambda \vdash n$, let $S_\lambda = \mathbb{C}[\mathfrak{S}_n]c_\lambda$ be the image of φ_λ. Recall that $\mathfrak{S}_n$ acts on $\mathbb{C}[\mathfrak{S}_n]$ on the left, so S_λ is certainly an $\mathfrak{S}_n$-invariant subspace. By Lemmas 2 and 3, we have $c_\lambda x c_\lambda \in \mathbb{C}c_\lambda$ for all $x \in \mathbb{C}[\mathfrak{S}_n]$. It follows from this that $c_\lambda S_\lambda \subset \mathbb{C}c_\lambda$. Now suppose for contradiction that $c_\lambda S_\lambda = 0$. Then $S_\lambda S_\lambda = \mathbb{C}[\mathfrak{S}_n]c_\lambda S_\lambda = 0$. By basic representation theory, there is an $\mathfrak{S}_n$-invariant projection from $\mathbb{C}[\mathfrak{S}_n]$ to S_λ; by Lemma 5 suppose the projection is right-multiplication by $x \in \mathbb{C}[\mathfrak{S}_n]$. Then $x = 1x \in S_\lambda$ and by definition of a projection, $x = x^2 \in S_\lambda \cdot S_\lambda = \{0\}$, so $S_\lambda = 0$, a contradiction. Thus $c_\lambda S_\lambda \neq 0$, and since $c_\lambda S_\lambda \subset \mathbb{C}c_\lambda$, a 1-dimensional space, we get $c_\lambda S_\lambda = \mathbb{C}c_\lambda$. Using the same reasoning, if W is a nonzero subrepresentation of S_λ, then $c_\lambda W = \mathbb{C}c_\lambda$, so $S_\lambda = \mathbb{C}[\mathfrak{S}_n]c_\lambda = \mathbb{C}[\mathfrak{S}_n]c_\lambda W \subset W$, since W is an $\mathfrak{S}_n$-invariant subspace, and thus $W = S_\lambda$. So S_λ is irreducible. Finally, assume $\lambda \neq \eta$, and we will show $S_\lambda \ncong S_\eta$. WLOG $\lambda > \eta$, so by Lemma 4, $c_\lambda S_\eta = \{0\}$. But we saw above that $c_\lambda S_\lambda \neq \{0\}$, so S_λ and S_η have distinct $\mathbb{C}[\mathfrak{S}_n]$-actions, and thus they are not isomorphic as representations. Now we are done – we have found a distinct irrep for each conjugacy class of $\mathfrak{S}_n$ (see Section 1.1), so this exhausts the irreps (up to isomorphism). □

1.3 Corollaries of the characterization theorem

As practice working with these representations, we can look at the 1-dimensional representations of $\mathfrak{S}_n$ and their tensor products with other irreps.

For $n \geq 2$, there are exactly two 1-dimensional irreps of $\mathfrak{S}_n$. To see this, note that all of $\mathfrak{S}_n$ is generated by the conjugacy class of two-letter transpositions, which in a 1-dimensional irrep is mapped to a square-root of 1, i.e. to ± 1. If it is mapped to 1, we get the trivial representation $\rho(g) = 1$; and if it is mapped to -1, we get the "alternating representation" A, defined by $\rho(g) = \text{sign}(g)$. It follows immediately from the construction of S_λ that $S_{(n)}$ is the trivial representation, and $S_{(1,1,\ldots,1)}$ is the alternating representation.

We define the *conjugate partition* λ' of $\lambda \vdash n$ by flipping the associated Young diagram about a 45° line. In terms of the partition $\lambda = (\lambda_1, \ldots, \lambda_m)$, we have $\lambda'_i = j$ iff $\lambda_j \geq i$ and $\lambda_{j+1} < i$.

Corollary 1. *For $n \geq 2$, let A be the alternating representation of $\mathfrak{S}_n$, and suppose $\lambda \vdash n$. Then $S_\lambda \cong A \otimes S_{\lambda'}$.*

Proof. We have $S_\lambda = \mathbb{C}[\mathfrak{S}_n]a_\lambda b_\lambda$, and we let $V = \mathbb{C}[\mathfrak{S}_n]b_\lambda a_\lambda$. First I will prove $V \cong A \otimes S_{\lambda'}$. We have that $\mathbb{C}[\mathfrak{S}_n] \cong A \otimes \mathbb{C}[\mathfrak{S}_n]$ as $\mathfrak{S}_n$-modules, by the automorphism taking $g \in \mathfrak{S}_n$ to $(g\,\text{sign}(g))$, and this automorphism maps b_λ to $a_{\lambda'}$ and a_λ to $b_{\lambda'}$, and thus V to $A \otimes S_{\lambda'}$. Hence, $V \cong A \otimes S_{\lambda'}$. So it suffices to show that $S_\lambda \cong V$. First note that both

S_λ and V are irreps – if V were reducible, then tensoring each component with A would give a decomposition of $S_{\lambda'}$, which is impossible. Now right-multiplication by $a_\lambda b_\lambda$ gives an endomorphism of S_λ, which by Schur's Lemma is multiplication by a scalar. I claim that the scalar is not 0; in fact,

Lemma 6. $c_\lambda c_\lambda = \frac{n!}{\dim S_\lambda} c_\lambda$.

Proof. By Lemmas 2 and 3, $c_\lambda c_\lambda = \alpha c_\lambda$ for some $\alpha \in \mathbb{C}$. As before, let $\varphi_\lambda : x \mapsto x c_\lambda$ act on $\mathbb{C}[\mathfrak{S}_n]$. Using the basis $\mathfrak{S}_n$ of $\mathbb{C}[\mathfrak{S}_n]$, we have that $\mathrm{Tr}(\varphi_\lambda)$ is $n!$, since it has only 1s on the diagonal (the g-component of $g c_\lambda$ is 1). Now we compute $\mathrm{Tr}(\varphi_\lambda)$ in a different basis, consisting of a basis for $\ker \varphi_\lambda$, unioned with the preimage of a basis for $\mathrm{Img}(\varphi_\lambda) = S_\lambda$. For the latter, we can just pick a basis of S_λ, since $\alpha \neq 0$ (if $c_\lambda^2 = 0$ then its eigenvalues would all be 0, so φ_λ would have trace 0 which we saw above is false). Then φ_λ is multiplication by 0 on the basis elements in $\ker \varphi_\lambda$, and multiplication by α on the basis elements in S_λ, so $\mathrm{Tr}(\varphi_\lambda) = \alpha \dim(S_\lambda)$. We thus have $\alpha \dim S_\lambda = \mathrm{Tr}(\varphi_\lambda) = n!$, proving the lemma. □

So indeed, right-multiplication by $a_\lambda b_\lambda$ is an isomorphism of S_λ. Then right-multiplication by a_λ is an injective (hence nontrivial) homomorphism $S_\lambda \to V$, so by Schur's Lemma $S_\lambda \cong V$ proving the corollary. □

2 Symmetric Functions and Irreps of $\mathfrak{S}_n$

We will define two $\mathbb{Z}$-graded rings – the ring of homogeneous symmetric functions (a concept closely related to symmetric polynomials), and a ring of representations of all finite symmetric groups – and show that they are isomorphic. The first step will be proving some basic facts about symmetric polynomials.

2.1 Symmetric polynomials

We define the *elementary symmetric polynomial* e_n in m variables by

$$e_n(x_1, \ldots, x_m) = \prod_{1 \leq i_1 < \cdots < i_n \leq m} x_{i_1} \cdots x_{i_n}$$

(in particular, if $m > n$, $e_n(x_1, \ldots, x_m) = 0$), and we define the *complete symmetric polynomial* h_n in m variables by

$$h_n(x_1, \ldots, x_m) = \prod_{1 \leq i_1 \leq \cdots \leq i_n \leq m} x_{i_1} \cdots x_{i_n}.$$

For a partition $\lambda = (\lambda_1, \dots, \lambda_j)$, we define e_λ and h_λ by

$$e_\lambda = e_{\lambda_1} \cdots e_{\lambda_j}; \quad h_\lambda = h_{\lambda_1} \cdots h_{\lambda_j}.$$

Next we define the *Schur function* $s_\lambda(x_1, \dots, x_m)$. We will give two definitions, both of which we will need later (for example, cf. (11) and (7)).

The first definition is in terms of a *Young semi-standard tableau* (abbreviated sst, plural ssts). An sst T of shape $\lambda \vdash n$ is obtained by numbering the boxes of the Young diagram λ with positive integers in such a way that the numbers are non-strictly increasing along each row and strictly increasing along each column. For an sst T with entries $t_1, \dots, t_n$ (in some order), define $(x_1, \dots, x_m)^T = x_{t_1} \cdots x_{t_n}$, or $(x_1, \dots, x_m)^T = 0$ in the case that some $t_i > m$. Now we can give our first definition of Schur functions: for $\lambda \vdash n$,

$$s_\lambda(x_1, \dots, x_m) = \sum_{T \text{ an sst of shape } \lambda} (x_1, \dots, x_m)^T. \tag{1}$$

Our second definition is simpler to state: for $\lambda \vdash n$, we say $s_\lambda(x_1, \dots, x_m) = 0$ if λ has more than m rows; otherwise we write $\lambda = (\lambda_1, \dots, \lambda_m)$ (possibly with ending 0s), and then

$$s_\lambda(x_1, \dots, x_m) = \frac{\begin{vmatrix} x_1^{\lambda_1+m-1} & \cdots & x_m^{\lambda_1+m-1} \\ \vdots & & \vdots \\ x_1^{\lambda_m+m-m} & \cdots & x_m^{\lambda_m+m-m} \end{vmatrix}}{\begin{vmatrix} x_1^{m-1} & \cdots & x_m^{m-1} \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{vmatrix}}. \tag{2}$$

We will prove the equivalence by induction, by proving for both definitions the "Pieri formula":

$$s_\lambda(x_1, \dots, x_m) s_{(p)}(x_1, \dots, x_m) = \sum_\eta s_\eta(x_1, \dots, x_m), \tag{3}$$

where the sum is over all ssts η that are obtained from λ by adding p boxes, each in a different column.

Lemma 7. *Define s_λ as in (1). Then the Pieri formula (3) holds.*

Proof. We set up a one-to-one correspondence between a pair of ssts for λ and (p), and an sst for some such η containing the same total collection of box-labels. We do this by a process called "row-bumping". Given an sst μ and a positive integer x, we *join* x to μ as follows: If x is at least as large as the right-most entry of the top row, we add a box onto the right of the top row and label it with x. Otherwise, we put x into its correct lexicographic place in the top row, shifting all the later entries one box to the right and "bumping" the right-most entry into the second row, where we repeat the process.

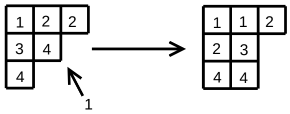

Figure 2: An example of row-bumping.

Given ssts in shapes λ and (p), we join the entry of the leftmost box of (p) to λ, then the second-to-leftmost, and so on. Since the entries of (p) are non-strictly increasing, it is clear from the algorithm that each added box will be strictly to the right of the previous added box (in particular, the added boxes will all be in different columns), and the entries of the resulting ssts are clearly the union of the entries of the two we started with. Conversely, starting with an sst with shape η as in the lemma, we can take the rightmost added box and run the row-bumping algorithm in reverse, and repeat that process to get two ssts with shapes λ and (p). The definition of Schur polynomials (1) gives the desired result. $\square$

Lemma 8. *Define s_λ as in (2). Then the Pieri formula (3) holds.*

Proof. Let $f(n_1, \ldots, n_m)$ be the determinant of the $m \times m$ matrix with (i,j)th entry $x_i^{n_j}$. So we need to prove that $f(l_1 + m - 1, \ldots, l_m)h_p = \sum f(n_1 + m - 1, \ldots, n_m)$, with the sum over $n_1 \geq l_1 \geq n_2 \geq l_2 \geq \cdots \geq n_m \geq l_m$ with $\sum n_i = p + \sum l_i$ (this expresses the condition that p squares were added, no two in the same column). Expressing this for all ps at once in a formal power series (and changing variables $n_i \mapsto n_i + m - i$, $l_i \mapsto l_i + m - i$), it suffices to show that

$$f(l_1, \ldots, l_m) \prod_{i=1}^{m} \frac{1}{1 - x_i} = \sum f(n_1, \ldots, n_m), \tag{4}$$

with the sum over all $n_1 \geq l_1 > n_2 \geq l_2 > \cdots n_m \geq l_m$. We prove this by induction on m. The base case $m = 1$ is trivial. And expanding the determinant along the top row, we see that (4) follows immediately from the induction hypothesis. $\square$

So we have proven that Pieri's formula (3) holds, and now we can easily prove that the two definitions of Schur polynomials (1) and (2) are equivalent. Since $s_{(p)} = h_p$ (under either definition), we can repeatedly apply Pieri's formula to get, for $\lambda \vdash n$,

$$h_\lambda = \sum_{\eta \vdash n} K_{\eta\lambda} s_\eta, \tag{5}$$

where $K_{\eta\lambda}$ counts the number of ways to build up η by adding successively $\lambda_1, \ldots, \lambda_j$ boxes in such a way that no two are added to the same column in a given stage, and such that the result of a given stage is a valid Young diagram. Equivalently, $K_{\eta\lambda}$ is the number of ssts of shape η with λ_1 1s, λ_2 2s, etc. (the equivalence of the definitions follows from labeling the λ_1 boxes added at the first stage 1, at the second stage 2, etc.). The

$K_{\eta\lambda}$ are known as *Kostka numbers.* If $\eta > \lambda$ in the lexicographic ordering defined above, it is clear from the first definition that $K_{\eta\lambda} = 0$, and it is also clear that $K_{\eta\eta} = 1$. Thus,

$$(K_{\eta\lambda})_{\eta,\lambda\vdash n} \text{ is an upper triangular matrix with 1s on the diagonal.} \tag{6}$$

In particular, we can invert the matrix K to get an expression for s_λ in terms of h_η. It follows that the two definitions for Schur functions are indeed equivalent. And it is clear from (2) that the Schur polynomials are symmetric.

We define two more symmetric polynomials. For $\lambda = (\lambda_1, \ldots, \lambda_j) \vdash n$, the *monomial symmetric polynomial*

$$m_\lambda(x_1, \ldots, x_m) = \sum x_{i_1}^{\lambda_1} \cdots x_{i_j}^{\lambda_j},$$

with the sum over sets $\{i_1, \ldots, i_j\}$ of distinct elements in $\{1, \ldots, m\}$. And the *Newton power sum*

$$p_\lambda(x_1, \ldots, x_m) = \prod_{i=1}^{j} p_{\lambda_i}(x_1, \ldots, x_m) = \prod_{i=1}^{j} (x_1^{\lambda_i} + \cdots + x_m^{\lambda_i}).$$

We prove another lemma relating these functions:

Lemma 9. *The following equality holds for formal power series:*

$$\begin{aligned}
\prod_{i=1}^{m}\prod_{j=1}^{m} \frac{1}{1 - x_i y_j} &= \sum_\lambda m_\lambda(x_1, \ldots, x_m) h_\lambda(y_1, \ldots, y_m) \\
&= \sum_\lambda s_\lambda(x_1, \ldots, x_m) s_\lambda(y_1, \ldots, y_m) \\
&= \sum_\lambda \frac{1}{z(\lambda)} p_\lambda(x_1, \ldots, x_m) p_\lambda(y_1, \ldots, y_m),
\end{aligned}$$

with the sum over all partitions λ of any integer n, where if $(\lambda_1, \ldots, \lambda_j)$ contains the integer n a_n times, $z(\lambda) := \prod_n n^{a_n} n!$.

Proof. The first line follows from a simple expansion:

$$\begin{aligned}
\prod_{i=1}^{m}\prod_{j=1}^{l} (1 + x_i y_j + x_i^2 y_j^2 + \cdots) &= \prod_{i=1}^{m}\sum_{n=1}^{\infty} x_i^n h_n(y_1, \ldots, y_m) \\
&= \sum_\lambda m_\lambda(x_1, \ldots, x_m) h_\lambda(y_1, \ldots, y_m).
\end{aligned}$$

The second line follows from the definition (2) of Schur polynomials. We first have

$$\sum s_\lambda(x) s_\lambda(y) \Delta(x) \Delta(y) = \det\left((1 - x_i y_j)^{-1}\right)_{1\le i,j\le m}, \tag{7}$$

where $\Delta(x) := \det |x_i^{j-1}|_{1\le i,j\le m}$: expanding the RHS, the coefficient of $x_1^{l_1}\cdots x_m^{l_m}$ for $l_1 > \cdots > l_m$ is $\det |y_i^{l_j}|_{1\le i,j\le n}$, and vice versa, so this equation holds. On the other hand,

$$\det\left((1-x_iy_j)^{-1}\right)_{1\le i,j\le m} = \Delta(x)\Delta(y)\prod_{i=1}^m\prod_{j=1}^m (1-x_iy_j)^{-1},$$

which follows from induction on m: subtract the first row from each other row, then factor out $(x_i - x_1)$ from the ith row ($2 \le i \le m$) and $\frac{1}{1-x_1y_j}$ from the jth column ($1 \le j \le m$). Next, subtract the first column from each other column, then factor out $(y_j - y_1)$ from the jth column ($2 \le j \le m$) and $\frac{1}{1-x_iy_1}$ from the ith row ($2 \le i \le m$). The determinant of the new matrix, by expansion in the first row, equals a smaller matrix of the original form, so we can use the induction hypothesis to prove this formula. So we have proven the second line of the lemma.

For the third line, we have a proof using formal power series:

$$\begin{aligned}\prod_{i=1}^m \frac{1}{1-x_it} &= \prod_{i=1}^m \exp(-\log(1-x_it)) = \prod_{i=1}^m \exp\left(\sum_{n=1}^\infty \frac{(x_it)^n}{n}\right) = \prod_{n=1}^\infty \exp\left(\sum_{i=1}^m \frac{(x_it)^n}{n}\right)\\ &= \prod_{n=1}^\infty \exp\left(\frac{p_n(x)t^n}{n}\right) = \prod_{n=1}^\infty\sum_{a_n=0}^\infty \left(\frac{p_n(x)t^n}{n}\right)^{a_n}\frac{1}{a_n!} = \sum_\lambda \frac{1}{z(\lambda)}p_\lambda(x)t^{|\lambda|}.\end{aligned}$$

If instead of $(x_1,\dots,x_m)$, we plug in $(x_1y_1, x_1y_2,\dots,x_my_m)$, and plug in $t=1$, we get the third line of the lemma. □

2.2 The ring Λ of symmetric functions

We define a *symmetric function p of degree d* to be a family of polynomials $p = (p_n)_{n\in\mathbb{Z}^+}$ such that each p_n is a symmetric homogeneous polynomial of degree d in n variables with integer coefficients, and such that for $m > n$, $p_m(x_1,\dots,x_n,0,\dots,0) = p_n(x_1,\dots,x_n)$. All the symmetric polynomials we have defined so far can be grouped into such families, so we can regard them all as symmetric functions – for example, instead of having the Schur polynomial s_λ in m variables for some $m\in\mathbb{Z}$, we can speak simply of the Schur function s_λ. Let Λ_d be the set of all symmetric functions of degree d. We define

$$\Lambda = \oplus_{d=1}^\infty \Lambda_d$$

(with $\Lambda_0 := \mathbb{Z}$). Λ is a graded ring, under ordinary polynomial addition and multiplication.

Lemma 10. *$\{s_\lambda|\lambda\vdash d\}$ is a basis of Λ_d.*

Proof. In a nontrivial sum $\sum_\lambda a_\lambda s_\lambda = 0$, let η be the first (lexicographically) λ with $a_\lambda \ne 0$, and then the coefficient of $x_1^{\eta_1}\cdots x_j^{\eta_j}$ is a_λ, a contradiction proving linear independence.

And for an arbitrary element $f \in \Lambda_d$, write f in d variables, let η be the first lexicographic λ with a nonzero coefficient of $x_1^{\lambda_1} \cdots x_j^{\lambda_j}$, and subtract s_λ times that coefficient from f. Repeating this process will express f as a linear combination of s_λ (both written in d variables). For more than d variables, the same relation will hold, since each term of f (being degree d) is uniquely determined by its terms that contain just the variables x_1 through x_d. So $\{s_\lambda\}$ is a spanning set. □

It follows that $\{s_\lambda | \lambda \vdash d,\ d \in \mathbb{Z}^+\}$ is a basis for Λ. Thus we can define an inner product $\langle \cdot, \cdot \rangle$ on Λ by making $\{s_\lambda\}$ orthonormal.

A lemma about this inner product will be useful later.

Lemma 11. *$\langle h_\lambda, m_\eta \rangle = \delta_{\lambda\eta}$ (i.e. 1 if $\lambda = \eta$, 0 otherwise). Also, $\langle p_\lambda, p_\eta \rangle = z(\lambda)\delta_{\lambda\eta}$.*

Proof. Write $h_\lambda = \sum_\nu x_{\lambda\nu} s_\nu$ and $m_\eta = \sum_\nu y_{\lambda\nu} s_\nu$. By Lemma 9, we get

$$\sum_\nu \left(s_\nu(x_1, \ldots, x_m)\right) \left(\sum_\mu s_\mu(y_1, \ldots, y_m) \left(\sum_\lambda x_{\lambda\nu} y_{\lambda\mu} - \delta_{\nu\mu} \right) \right) = 0$$

for any $m \geq 1$. Since the $s_\lambda(x_1, x_2, \ldots)$ are linearly independent over $\mathbb{Z}$, they are certainly linearly independent over $\mathbb{Z}[y_1, y_2, \ldots]$, so we get $\sum_\mu s_\mu(y_1, \ldots, y_m) \left(\sum_\lambda x_{\lambda\nu} y_{\lambda\mu} - \delta_{\nu\mu}\right) = 0$ for any $m \geq 1$. Again, by the linear independence of s_μ, we have $\sum_\lambda x_{\lambda\nu} y_{\lambda\mu} = \delta_{\nu\mu}$, and by definition of the inner product, we get $\langle h_\lambda, m_\eta \rangle = \delta_{\lambda\eta}$. The identity $\langle p_\lambda, p_\eta \rangle = z(\lambda)\delta_{\lambda\eta}$ follows from Lemma 9 in an analogous way. □

2.3 The ring R of representations of symmetric groups

For $n \geq 1$, we let R_n be the free (additive) abelian group generated by the irreps of $\mathfrak{S}_n$ (or more precisely, the isomorphism classes of the irreps of $\mathfrak{S}_n$). Any representation of $\mathfrak{S}_n$ is embedded in this group, by identifying $V \oplus W$ with $V + W \in R_n$. We further define a "multiplication" map $R_n \times R_m \to R_{m+n}$ defined by

$$V \circ W = \mathrm{Ind}_{\mathfrak{S}_n \times \mathfrak{S}_m}^{\mathfrak{S}_{n+m}} (V \otimes W) \tag{8}$$

where V and W are irreps. For general elements, $\circ$ is defined to distribute over addition – i.e.

$$\left(\sum_{\lambda \vdash n} a_\lambda S_\lambda \right) \circ \left(\sum_{\eta \vdash m} b_\eta S_\eta \right) = \sum \sum a_\lambda b_\eta (S_\lambda \circ S_\eta).$$

Elementary properties of tensor products and induced representations imply that equation (8) holds for any representations V, W. We now define

$$R = \oplus_{n=1}^{\infty} R_n.$$

R is a graded ring. The irreps of all $\mathfrak{S}_n$ for $n \geq 1$ are certainly a basis for R, so we define an inner product $\langle \cdot, \cdot \rangle$ on R by making the irreps of all $\mathfrak{S}_n$ an orthonormal set. Note that, when restricted to R_n, this inner product agrees with the regular inner product of representations, defined by $\langle V, W \rangle = \frac{1}{n!} \sum_{g \in \mathfrak{S}_n} \chi_V(g)\chi_W(g)$, where χ_V denotes the character of V.

For $\lambda \vdash n$, we define H_λ to be the representation of $\mathfrak{S}_n$ induced from the trivial representation of the subgroup A_λ (where A_λ is defined as in section 1.1). An equivalent definition is $H_\lambda = \mathbb{C}[\mathfrak{S}_n]a_\lambda$, where a_λ is again defined as in section 1.1.

2.4 Isomorphism theorem

We will show in this section that Λ and R are isomorphic and isometric. The following lemma is the first step in connecting these two rings:

Lemma 12. *For partitions $\lambda, \eta \vdash n$, define ξ^λ_η so that $p_\eta = \sum_\lambda \xi^\lambda_\eta m_\lambda$ (i.e. ξ^λ_η is the coefficient of $x_1^{\lambda_1} \cdots x_n^{\lambda_n}$ in the expansion of p_η). Then* (I) *$h_\lambda = \sum_\eta \frac{1}{z(\eta)} \xi^\lambda_\eta p_\eta$, and* (II) *$\chi_{H_\lambda}(C(\eta)) = \xi^\lambda_\eta$, where $C(\eta)$ is the conjugacy class of elements of $\mathfrak{S}_n$ with cycle structure η.*

Proof. (I) follows from Lemma 11: Taking the inner product with h_λ gives $\langle h_\lambda, p_\eta \rangle = \xi^\lambda_\eta$. Thus $\langle h_\lambda, p_\eta \rangle = \langle \sum_\nu \frac{1}{z(\nu)} \xi^\lambda_\nu p_\nu, p_\eta \rangle$ for all p_η, and since $\{p_\eta\}$ span Λ (by Lemma 9) and the inner product is nondegenerate, (I) must hold.

For (II), by the character formula for induced representations,

$$\chi_{H_\lambda}(C(\mu)) = \frac{[\mathfrak{S}_n : A_\lambda]}{|C(\mu)|} |C(\mu) \cap A_\lambda|.$$

Let a_k be the number of cycles of length k in μ. We have, by combinatorics, $|C(\mu)| = \frac{n!}{\prod_i i^{a_i} a_i!}$ (write an element of $C(\mu)$ in cycle notation, then remove the parentheses to get one of $n!$ permutations; but this overcounts by a factor of the expression in the denominator). Moreover, $|A_\lambda| = \lambda_1! \cdots \lambda_n!$. Finally, we compute $|C(\mu) \cap A_\lambda|$, or the number of permutations with cycle structure μ that preserve the rows of T_λ. Given such a permutation, let r_{pq} be the number of q-cycles that permute elements of the pth row of T_λ. We have

$$r_{p1} + 2r_{p2} + \cdots + nr_{pn} = \lambda_p, \qquad r_{1q} + \cdots + r_{nq} = a_q, \tag{9}$$

and given a set of numbers r_{pq} satisfying (9), there are $\prod_{p=1}^n \frac{\lambda_p!}{1^{r_{p1}} r_{p1}! \cdots n^{r_{pn}} r_{pn}!}$ different permutations characterized by r_{pq} (by the same counting method as above). Combining these, we get

$$\begin{aligned}\chi_{H_\lambda}(C(\mu)) &= \frac{n!}{\lambda_1!\cdots\lambda_n!}\frac{\prod_i i^{a_i}a_i!}{n!}\sum\prod_{p=1}^{n}\frac{\lambda_p!}{1^{r_{p1}}r_{p1}!\cdots n^{r_{pn}}r_{pn}!}\\ &= \sum\prod_{q=1}^{n}\frac{a_q!}{r_{1q}!\cdots r_{nq}!},\end{aligned}$$

with the sum over all sets of nonnegative integers r_{pq} satisfying (9). Meanwhile, for $1 \leq q \leq n$, we have

$$(x_1^q + \cdots + x_n^q)^{m_q} = \sum \frac{a_q!}{r_{1q}!\cdots r_{nq}!}x_1^{qr_{1q}}\cdots x_n^{qr_{nq}}$$

(the sum again over r_{pq} satisfying (9)). Taking the product from $q = 1$ to n, we get that $\chi_{H_\lambda}(C(\mu))$ is the coefficient of $x_1^{\lambda_1}\cdots x_n^{\lambda_n}$ in p_λ, as desired. □

We can now prove the isomorphism theorem.

Theorem 2. *Define $\varphi : \Lambda \to R$ to be an additive homomorphism satisfying $\varphi(h_\lambda) = H_\lambda$. Then* (I) *$\varphi$ is an isomorphism of graded rings respecting the inner product,* (II) *$\varphi(s_\lambda) = S_\lambda$, and* (III) *$\varphi$ has inverse ψ, satisfying, for V a representation of $\mathfrak{S}_n$,*

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

(Note that we need to pass to $\Lambda\otimes\mathbb{Q}$ to define ψ, but its image is nevertheless in $\Lambda \subset \Lambda\otimes\mathbb{Q}$.)

Proof. First we note that φ is well-defined, since h_λ is a basis for Λ by Lemma 10: the h_λ span Λ by Lemma 9, and they are linearly independent since $\{h_\lambda\}$ and $\{s_\lambda\}$ have the same number of terms with any given degree. Next, we note that $H_{\lambda_1} \circ \cdots \circ H_{\lambda_j} = H_\lambda$, as follows from transitivity of induced representations and the definitions of H_k and $\circ$. Any element of Λ can be written as a polynomial in $\{h_n\}$, so φ is a homomorphism of rings. It trivially respects the $\mathbb{Z}$-grading.

It follows from the definition of ψ that it is an additive homomorphism into $\Lambda\otimes\mathbb{Q}$ (although we will soon see that its image is in $\Lambda \subset \Lambda\otimes\mathbb{Q}$). I will show that $\psi\circ\varphi$ is the identity map. For $\lambda \vdash n$, we have

$$\psi(\varphi(h_\lambda)) = \psi(H_\lambda) = \sum_{\mu\vdash n}\frac{1}{z(\mu)}\chi_{H_\lambda}(C(\mu))p_\mu.$$

By parts (I) and (II) of Lemma 12, this sum is indeed h_λ. So indeed $\psi\circ\varphi$ is the identity on h_λ, and thus on all of Λ, and it follows that ψ is the inverse of φ on the image of φ, so that φ is an isomorphism into its image.

Next we show that ψ is an isometry. For V and W representations of $\mathfrak{S}_n$, we have

$$\begin{aligned}\langle \psi(V), \psi(W)\rangle &= \sum_{\lambda,\eta\vdash n} \frac{\chi_V(C(\lambda))\chi_W(C(\eta))}{z(\lambda)z(\eta)}\langle p_\lambda, p_\eta\rangle \\ &= \sum_{\lambda\vdash n} \frac{\chi_V(C(\lambda))\chi_W(C(\lambda))}{z(\lambda)} \\ &= \frac{1}{n!}\sum_{\lambda\vdash n} |C(\lambda)|\chi_V(C(\lambda))\chi_W(C(\lambda)) \\ &= \langle V, W\rangle.\end{aligned}$$

Since the representations of $\mathfrak{S}_n$ linearly span R_n, we have, for $x, y \in R_n$, $\langle \psi(x), \psi(y)\rangle = \langle x, y\rangle$. Moreover, ψ respects the $\mathbb{Z}$-grading, so if $x \in R_n, y \in R_m, m \neq n$, then $\langle \psi(x), \psi(y)\rangle = 0 = \langle x, y\rangle$. This proves that ψ is an isometry. In particular, it is injective, so φ is surjective, so φ is an isomorphism. This completes the proof of (I) and (III).

Finally, we prove (II). Fix an n. For $\lambda \vdash n$, we can write $\varphi(s_\lambda) = \sum_{\mu\vdash n} m_\mu S_\mu$ for some $m_\mu \in \mathbb{Z}$. Since $1 = \langle s_\lambda, s_\lambda\rangle = \langle \varphi(s_\lambda), \varphi(s_\lambda)\rangle = \sum m_\mu^2$, there is an η such that $\varphi(s_\lambda) = \pm S_\eta$.

By equations (5) and (6), we can write $h_\lambda = s_\lambda + \sum_{\eta<\lambda} K_{\eta\lambda} s_\eta$. Now suppose inductively that $\varphi(s_\eta) = S_\nu$ for all $\eta < \lambda$. Since φ is a homomorphism, we have

$$\langle H_\lambda, S_\lambda\rangle = \langle \varphi(s_\lambda) + \sum_{\eta<\lambda} K_{\eta\lambda} S_\eta, S_\lambda\rangle = \langle \varphi(s_\lambda), S_\lambda\rangle.$$

Now $\langle H_\lambda, S_\lambda\rangle \geq 1$, since we have a nonzero homomorphism from H_λ to S_λ, by writing $H_\lambda = \mathbb{C}[\mathfrak{S}_n]a_\lambda$ and $S_\lambda = \mathbb{C}[\mathfrak{S}_n]a_\lambda b_\lambda$, and using the surjection $f : x \mapsto xb_\lambda$. So $\langle \varphi(s_\lambda), S_\lambda\rangle \geq 1$. But we showed above that $\varphi(s_\lambda) = \pm S_\eta$ for some η, and it follows that $\varphi(s_\lambda) = S_\lambda$. By induction, (II) holds. □

2.5 Frobenius character formula

A particularly important corollary of this theorem is the Frobenius character formula, which allows us to directly compute the character of S_λ:

Corollary 2. (Frobenius character formula.) *Let $\chi_\lambda(\eta)$ be the character of S_λ on the conjugacy class $C(\eta)$. Then:* (I) $s_\eta = \sum_\lambda \frac{1}{z(\lambda)}\chi_\eta(\lambda)p_\lambda$*;* (II) $p_\eta = \sum_\lambda \chi_\lambda(\eta)s_\lambda$*; and* (III) *$\chi_\lambda(\eta)$ is the coefficient of $x_1^{\lambda_1+m-1}x_2^{\lambda_2+m-2}\cdots x_m^{\lambda_m}$ in the expansion of $(\Delta(x)p_\eta(x))$, where $x = (x_1, \ldots, x_m)$ and $\Delta(x) = \det|x_i^{j-1}|_{1\leq i,j\leq m} = \prod_{i<j}(x_i - x_j)$.*

Proof. (I) follows from the fact that $\psi(S_\lambda) = s_\lambda$. (II) follows from (I) and Lemma 11, in an analogous way to part (I) of Lemma 12. Finally, (III) follows from (II) and the definition (2) of Schur polynomials. □

3 Representations of $\mathfrak{sl}_m$

A direct application of what we have proven above is the classification of the irreps of the Lie algebra $\mathfrak{sl}_m$, defined to be the $\mathbb{C}$-algebra of traceless complex $m \times m$ matrices, with the Lie bracket being matrix commutation. We will be using some elementary facts and terminology for simple Lie algebras and their representations – for definitions and proofs, see [1], especially chapter 14. We will first use the representation theory of $\mathfrak{S}_n$ to characterize the irreps of $\mathfrak{sl}_m$, and then we will relate the representation ring of $\mathfrak{sl}_m$ to the rings Λ and R we defined in the previous section.

3.1 Characterization of the irreps of $\mathfrak{sl}_m$

From the definition of $\mathfrak{sl}_m$, it has an m-dimensional defining representation V. Any tensor power $V^{\otimes n}$ of V is also a representation of $\mathfrak{sl}_m$, by the usual Lie algebra construction:

$$g(v_1 \otimes \cdots \otimes v_n) = (g(v_1) \otimes v_2 \otimes \cdots \otimes v_n) + \cdots + (v_1 \otimes \cdots \otimes v_{n-1} \otimes g(v_n)). \tag{10}$$

But $\mathfrak{S}_n$ acts on $V^{\otimes n}$ as well, on the right, by permuting tensor components:

$$(v_1 \otimes \cdots \otimes v_n)\sigma = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}.$$

These two actions clearly commute (justifying our notation), so we can break down $V^{\otimes n}$ into irreps with respect to $\mathfrak{S}_n$, i.e.

$$V_\lambda = V^{\otimes n} c_\lambda,$$

and we have that V_λ is a representation of $\mathfrak{sl}_m$. We will prove that $V_\lambda = 0$ if λ has more than m rows, and otherwise V_λ is an irrep; and moreover all irreps of $\mathfrak{sl}_m$ are isomorphic to V_λ for some λ.

Take the basis $\{e_1, \ldots, e_m\}$ of V such that the representation V maps matrices in $\mathfrak{sl}_m$ to themselves. A numbering T of the λ with integers in $[1, m]$ corresponds to an element of the space V_λ as follows: if $a_1, \ldots, a_n$ are the numbers in the boxes, from left to right, top to bottom, T gets mapped to

$$e_T := (e_{a_1} \otimes \cdots \otimes e_{a_n})c_\lambda.$$

$\{e_T\}$ is certainly a spanning set for V_λ, but not linearly independent for $n, m > 1$. However, it turns out that $\{e_T \mid T \text{ is an sst with entries in } [1, m]\}$ is a basis. The first step in proving this is:

Lemma 13. *$\{e_T \mid T$ is an sst with entries in $[1, m]\}$ spans V_λ.*

Proof. We want show that if T is not an sst, we can write e_T as a linear combination of $e_{T'}$ where T' are ssts. We have two relations between the e_T:

(I)e_T is alternating under permutations within each column; (II)Fixing a j and k, we have $e_T = \sum e_{T'}$, with the sum over all numberings which result from switching the top k entries of the $(j+1)$st row with some set of k entries in the jth row, while fixing the vertical order of the entries in each set.

(I) follows directly from Lemma 2, but the proof to (II) is more involved. For a nonempty subset Y of the $(j+1)$st column of T, we define $\gamma_Y(T) = \sum e(S,T)e_S$, with the sum over all S that result from exchanging a subset of Y with a subset of the jth column, preserving the vertical order of each subset, and with $e(S,T)$ being (-1) to the power of the number of entries in each subset. I claim $\gamma_Y(T) = 0$. We let X be the set of entries exchanged with Y, let H be the subset of $\mathfrak{S}_n$ preserving the elements outside $X \cup Y$, and let K be the subset of H taking X to itself (and Y to itself). Note that if $\sigma, \sigma' \in H$ differ by an element of K, then by (I), $\text{sign}(\sigma)e_{\sigma T} = \text{sign}(\sigma_2)e_{\sigma_2 T}$. Thus

$$\gamma_Y(T) = \sum_{\sigma \in H/K} \text{sign}(\sigma)e_{\sigma T} = \frac{1}{|X|!|Y|!}\sum_{h \in H} \text{sign}(h)e_{hT},$$

so it suffices to show that

$$0 = \sum_{h \in H} \text{sign}(h)e_{hT} = \sum_{q \in B_T} \text{sign}(q)\left(\sum_{p \in A_T}\sum_{h \in H} \text{sign}(h)(e_{t_1} \otimes \cdots \otimes e_{t_n})hp\right) q.$$

We fix a q, and pick an $t \in H$ that transposes two elements in the same row of qT. Then, since $(ht)p = hp$ and $\text{sign}(ht) = -\text{sign}(h)$, we can pair up terms to prove that the sum is in fact 0.

To finish the proof of (II), we note that $\sum_Y (-1)^{|Y|}\gamma_Y(T) = \sum e_{T'} - e_T$, by the inclusion-exclusion principle, so indeed $\sum e_{T'} = e_T$.

Now we define an ordering on tableaux of shape λ with entries in $[1,m]$, by $S > T$ if S has some entry bigger than T, but S and T agree in all entries directly below and all entries (non-directly) to the right. Take an arbitrary numbering T, and use (I) to assume WLOG that it is strictly increasing along columns. Suppose that T is not an sst. Then take the rightmost column that contains an entry greater than the one to its left, and the lowest such entry in that column – say this is the kth from the top in the $(j+1)$st column. We perform the exchange (II) with the top k entries of the jth column, and we get a sum of e_S, with each $S > T$ in the above ordering. We then repeat this process with each non-sst in our expression for T, and so on. Since there are only finitely many possible numberings, the process must terminate, with e_T expressed as a linear combination of e_S for ssts S. □

Next we can prove

Lemma 14. *$\{e_T \mid T$ is an sst with entries in $[1,m]\}$ is a basis for V_λ.*

Proof. We count dimensions. We have $\mathbb{C}[\mathfrak{S}_n] = \oplus_{\lambda\vdash n}(S_\lambda)^{\oplus \dim S_\lambda}$, so $V^{\oplus n} = \oplus_{\lambda\vdash n}(V_\lambda)^{\oplus \dim S_\lambda}$. Taking the dimension of each side and using Lemma 13 and part (II) of Corollary 2,

$$\begin{aligned} m^n &= \sum_{\lambda\vdash n}(\dim S_\lambda)(\dim V_\lambda) \\ &\leq \sum_{\lambda\vdash n}(\chi_\lambda(1^n))(\#\text{ of ssts of shape }\lambda\text{ with entries in }[1,m]) \\ &= \sum_{\lambda\vdash n}(\chi_\lambda((1^n))(s_\lambda(1^m)) \qquad (11) \\ &= p_{(1^n)}(1^m) \\ &= m^n \end{aligned}$$

(where the notation (1^k) denotes the k-tuple $(1,\ldots,1)$). Hence the inequality is an equality, so the number of ssts of shape λ with entries in $[1,m]$ is equal to $\dim V_\lambda$. Combining this with Lemma 13, we are done. □

If λ has more than m rows, there are no ssts of shape λ with entries in $[1,m]$, so $V_\lambda = 0$. But if V_λ has at most m rows, V_λ is nonzero and irreducible, and we prove this next.

We choose a basis for $\mathfrak{sl}_m$ as follows: Let M_{ij} be the $m\times m$ matrix with 1 at (i,j), 0 elsewhere. Now $\mathfrak{sl}_m$ is the linear span of:
(1) $H_i = M_{i,i} - M_{i+1,i+1}$ for $1 \leq i \leq m-1$;
(2) $E_{ij} = M_{ij}$ for $1 \leq i < j \leq m$;
(3) $F_{ij} = M_{ij}$ for $1 \leq j < i \leq m$.
It is easy to show that $\{H_i\}$ spans a Cartan subalgebra and that we can say each E_{ij} is a positive root vector and each F_{ij} is a negative root vector.

For any representation V, we say $v \in V$ is a *weight vector* iff it is an eigenvector of each H_i; since the H_i are already diagonal, the weight vectors are exactly the scalar multiples of e_T for ssts T. We say a nonzero weight vector v is a *highest weight vector* iff $E_{ij}(v) = 0$ for all $1 \leq i < j \leq m$. A basic fact, which we will not prove, is that a representation is irreducible iff it has a unique highest weight vector.

Lemma 15. *Suppose a partition λ has at most m rows. Then e_T is a highest weight vector of V_λ iff T is the sst of shape λ such that each entry in row i is i.*

Proof. Let T^* be the specified sst. First we write a formula for the action of $\mathfrak{sl}_m$ on a weight vector e_T. Let $T_{t_1\ldots t_n}$ be the numbering of λ with entries $t_1,\ldots,t_n$, in left-to-right, top-to-bottom order, and let g_{ij} be the (i,j)th entry of $g \in \mathfrak{sl}_m$. I claim

$$g e_{T_{t_1\ldots t_n}} = \sum_{i=1}^{n}\sum_{j=1}^{m} g_{j,t_i} e_{T_{t_1\ldots t_{i-1} j t_{i+1}\ldots t_n}}. \qquad (12)$$

Since the actions of $\mathfrak{sl}_m$ and $\mathfrak{S}_n$ commute, it suffices to prove that

$$g(e_{t_1} \otimes \cdots \otimes e_{t_n}) = \sum_{i=1}^{n} \sum_{j=1}^{m} g_{j,t_i}(e_{t_1} \otimes \cdots e_{t_{i-1}} \otimes e_j \otimes e_{t_{i+1}} \otimes e_{t_n}),$$

which follows directly from equation (10).

Now it follows easily that $E_{pq}e_{T^*} = 0$: If there were a nonzero term in (12), it would have $j = p$, $t_i = q$, $p < q$. Hence the modified tableau $e_{T'} = e_{T_{t_1 \cdots t_{i-1} j t_{i+1} \cdots t_n}}$ will have a p entry in the qth row. Looking directly above that entry in the pth row, there will be another p entry, so by antisymmetry in the columns, $e_{T'} = 0$. Hence we have proven that $E_{pq}e_{T^*} = 0$.

Next, suppose $T \neq T^*$ is an sst with entries in $[1, m]$, and write $T = T_{t_1 \dots t_n}$. Let i be the lowest number such that the ith box is in the jth row, but $j \neq t_i$. By definition of an sst, $j < t_i$, so we can let $g = E_{j,t_i}$. Then $e_{T_{t_1 \cdots t_{i-1} j t_{i+1} \cdots t_n}}$ is an sst with entries in $[1, m]$, so looking at (12), we see $ge_T \neq 0$. We have thus proven the lemma. □

It follows that each V_λ is an irrep of $\mathfrak{sl}_m$. We now prove that each irrep of $\mathfrak{sl}_m$ is of the form V_λ. The proof is based purely on Lie algebra theory, so we will be brief.

Lemma 16. *Each representation of $\mathfrak{sl}_m$ is isomorphic to V_λ for some partition λ.*

Proof. One computes the simple roots $(0, \dots, 0, 1, -1, 0, \dots, 0)$, and the fundamental weights $(1, \dots, 1, 0, \dots, 0)$. The pth fundamental weight w_p is the sum of the p highest weights of V, and thus is the highest weight of $\wedge^p V = V_{(1^p)}$. Thus, for any $a_1, \dots, a_m \geq 0$, we know that $(\wedge^1 V)^{\otimes a_1} \otimes \cdots \otimes (\wedge^m V)^{\otimes a_m} \subset V^{\otimes(a_1 + \cdots + a_m)}$ has highest weight $a_1 w_1 + \cdots + a_m w_m$. Hence every irrep W is in a tensor power of V. We decomposed this tensor power into spaces of the form V_λ, so $W \cong V_\lambda$ for some λ. □

The highest weight of V_λ is the weight of e_{T^*} (as defined in Lemma 15), which is $(\lambda_1 - \lambda_2, \dots, \lambda_{m-1} - \lambda_m)$, so $V_\lambda \cong V_{\lambda'}$ iff $\lambda_i - \lambda'_i$ is constant for $1 \leq i \leq m$. Combining this with the above Lemmas, we have proven:

Theorem 3. *If λ has more than m rows, $V_\lambda = 0$, and otherwise V_λ is an irrep of $\mathfrak{sl}_m$. All irreps of $\mathfrak{sl}_m$ are isomorphic to some V_λ. If λ and λ' have at most m rows, then $V_\lambda \cong V_{\lambda'}$ iff $\lambda_1 - \lambda'_1 = \lambda_2 - \lambda'_2 = \cdots = \lambda_m - \lambda'_m$.*

3.2 The representation ring of $\mathfrak{sl}_m$

We define the *representation ring* $\mathfrak{R}_m$ of $\mathfrak{sl}_m$ to be the free abelian group on the irreps of $\mathfrak{sl}_m$ (more precisely, on the isomorphism classes of irreps of $\mathfrak{sl}_m$). We embed all representations of $\mathfrak{sl}_m$ into $\mathfrak{R}_m$ by writing $U \oplus W = U + W$, and then define multiplication so that $U \cdot W = U \otimes W$.

We define the *character* of a representation W of $\mathfrak{sl}_m$ to be a function $\mathrm{Char}_W(x_1, \ldots, x_m)$ on the nonzero complex variables $x_1, \ldots, x_m$, defined as follows: Let $\gamma : W \to U$ be the exponentiation map, so that U is a representation of $SL_m\mathbb{C}$. Then define $\mathrm{Char}_W(x_1, \ldots, x_m)$, for $x_1 \cdots x_m = 1$, to be the trace of the element $D(x_1, \ldots, x_m) \in SL_m\mathbb{C}$ on U, where

$$D(x_1, \ldots, x_m) = \begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_m \end{pmatrix}.$$

The map γ takes the Cartan subalgebra to the set of diagonal matrices, and thus takes weight vectors to eigenvectors of $D(x_1, \ldots, x_m)$. Moreover, we can compute

$$\begin{aligned} D(x_1, \ldots, x_m)\gamma(e_T) &= \gamma(D(\ln x_1, \ldots, \ln x_m)(e_T)) \\ &= \gamma((\sum_i \ln x_{t_i})e_T) \\ &= \gamma((\ln x_1, \ldots, \ln x_m)^T e_T) \\ &= (x_1, \ldots, x_m)^T \gamma(e_T), \end{aligned}$$

and it follows from Lemma 14 that $\mathrm{Char}_{V_\lambda}(x_1, \ldots, x_m) = s_\lambda(x_1, \ldots, x_m)$. Note that the restriction $x_1 \cdots x_m = 1$ is important – for example, if $\lambda_i + 1 = \lambda'_i$ for $1 \le i \le m$, then $V_\lambda \cong V_{\lambda'}$ but

$$\mathrm{Char}_{V_{\lambda'}} = s_{\lambda'}(x_1, \ldots, x_m) = x_1 \cdots x_m s_\lambda(x_1, \ldots, x_m) = x_1 \cdots x_m \,\mathrm{Char}_{V_\lambda}.$$

It is easy to verify that $\mathrm{Char}_{U \oplus W} = \mathrm{Char}_U + \mathrm{Char}_W$ and $\mathrm{Char}_{U \otimes W} = \mathrm{Char}_U \cdot \mathrm{Char}_W$. Char is thus a homomorphism from $\mathfrak{R}_m$ to $\Lambda(m)/(x_1 \cdots x_m - 1)$, where $\Lambda(m) \subset \Lambda$ is the ring of symmetric polynomials in m variables.

We also have an surjective additive homomorphism $\alpha : R \to \mathfrak{R}_m$, taking λ to V_λ (which might be 0). Putting together Char, α, and the map ψ defined in Theorem 2, we have a composite map

$$\Lambda \to R \to \mathfrak{R}_m \to \Lambda(m).$$

This map takes the symmetric function $s_\lambda \in \Lambda$ to the symmetric polynomial $s_\lambda(x_1, \ldots, x_m) \in \Lambda(m)/(x_1 \cdots x_m - 1)$, and since the s_λ are a basis for Λ, it takes any symmetric function $f(x_1, x_2, \ldots)$ to $f(x_1, \ldots, x_m, 0, 0, \ldots)$. By Theorem 3, if we let I be the ideal generated by s_λ with λ having more than m rows, and $s_\lambda - s_{\lambda'}$, where λ and λ' differ by one or more columns of m squares, then $V_\lambda \cong \Lambda/I$. Put another way,

Theorem 4. *Let I be the ideal of Λ generated by the symmetric functions e_{m+1} and $(e_m - 1)$. Then the $\mathbb{Z}/m$-graded ring Λ/I is isomorphic to $\mathfrak{R}_m$.*

Proof. Note that both $e_{m+1} = 0$ and $e_m = 1$ respect the $\mathbb{Z}/m$-grading. The ideal generated by e_{m+1} consists of all symmetric functions f with $f(x_1, \ldots, x_m) = 0$, so $\Lambda/e_{m+1} \cong \Lambda(m)$. And in $\Lambda(m)$, $(e_m - 1) = (x_1 \cdots x_m - 1)$. So this theorem follows from the above discussion. □

Thus we can, for example, decompose tensor products of representations of $\mathfrak{sl}_m$ by doing the analogous operation on Schur polynomials. The $\mathbb{Z}/m$-grading on $\mathfrak{R}_m$ is sometimes called *m-ality*, or *triality* in the case $m = 3$.

As a closing note, the Weyl group of $\mathfrak{sl}_m$ is $\mathfrak{S}_m$, and $\mathfrak{S}_m$ is a symmetry of $\mathfrak{R}_m = \Lambda(m)/I$. This property holds in general.

References

[1] W. Fulton, J. Harris, *Representation Theory*, New York: Springer-Verlag, 1991. [I used the material in Chapters 4, 6, and Appendix A throughout the paper. Section 1 was based especially closely on Chapter 4.]

[2] W. Fulton, *Young Tableaux*, New York: Cambridge University Press, 1997. [I used material throughout this book, in particular Chapters 2, 6, 7, and 8. Section 2 in particular was based mostly on this book.]

[3] D. Knutson, *λ-Rings and the Representation Theory of the Symmetric Group*, New York: Springer-Verlag, 1973. [I did not use this explicitly, but read it through for reference.]