Matrix representations of the symmetric group

Tony Mendes (aamendes@calpoly.edu)

Version Date: December 12, 2023

Abstract

These notes introduce to the representation theory of the symmetric group. Prerequisites include undergraduate courses on linear algebra and abstract algebra.

Contents

1 Representations of finite groups **2**
- 1.1 Matrix representations 2
- 1.2 New representations from known representations 5
- 1.3 Irreducible representations 8
- 1.4 Characters 11
- 1.5 The conjugacy classes of the symmetric group 13
- 1.6 Class functions and an inner product for characters 15
- 1.7 Character tables 19

2 Tableau combinatorics **23**
- 2.1 Tabloids 23
- 2.2 The irreducible representations of the symmetric group 26

3 Symmetric functions **31**
- 3.1 Standard bases for symmetric functions 31
- 3.2 Sums involving the power symmetric functions 33
- 3.3 Rim hook tableaux 36

4 The RSK algorithm **43**
- 4.1 Bumping 43
- 4.2 The Pieri rules 47
- 4.3 The full RSK algorithm 50

5 The Frobenius map **57**

6 Hooks **58**
6.1 The major index for tableaux . . . 58
6.2 Hooks . . . 60

7 A summary of the main results **63**

8 Exercises **64**

1 Representations of finite groups

1.1 Matrix representations

Definition. *Let G be a finite group and let $GL_n(\mathbb{C})$ be the group of $n \times n$ invertible matrices with entries in $\mathbb{C}$. A (matrix) representation of G of degree n is a group homomorphism $X : G \to GL_n(\mathbb{C})$.*

In particular, if X is a degree n representation of G, then

1. $X(\varepsilon) = I_n$ since any homomorphism sends the identity $\varepsilon \in G$ to the identity matrix $I_n \in GL_n(\mathbb{C})$.
2. $X(gh) = X(g)X(h)$ for all $g, h \in G$.

As a consequence, $X(g^{-1}) = X(g)^{-1}$ for all $g \in G$.

Definition. *The trivial representation for G is the degree 1 representation X defined by $X(g) = 1$ for all $g \in G$. The trivial representation is sometimes written as simply 1 so that $1(g) = 1$ for all $g \in G$.*

Example 1. Let $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ be the cyclic group of order 4. A degree 1 representation X is given by

$$X(0) = 1, \qquad X(1) = i, \qquad X(2) = -1, \qquad X(3) = -i.$$

Example 2. A degree 2 representation X of $\mathbb{Z}_4$ is given by

$$X(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad X(1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad X(2) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad X(3) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Definition. *Let S_n be the symmetric group. The sign representation X is the degree 1 representation given by $X(\sigma) = \operatorname{sign} \sigma$.*

It is a good exercise to show that $\operatorname{sign}(\sigma\tau) = \operatorname{sign}\sigma \operatorname{sign}\tau$, proving that the sign function is a homomorphism and thus a representation.

Example 3. By directly verifying the matrix multiplications, it can be seen that a degree 2 representation X of S_3 is given by

$$X\big((1)\big) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad X\big((1\,2)\big) = \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix},$$
$$X\big((1\,3)\big) = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}, \qquad X\big((2\,3)\big) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$
$$X\big((1\,2\,3)\big) = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}, \qquad X\big((1\,3\,2)\big) = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix},$$

where permutations are written in cycle notation.

A group G acts on a set S if there is a function from $G \times S$ to S such that $\varepsilon\, s = s$ for all $s \in S$ and $(gh)\, s = g\,(h\, s)$ for all $g, h \in G$ and $s \in S$. Our next theorem says that there is a degree $|S|$ representation of G every time G acts on a finite set S.

Theorem 4. *Let G act on a set $S = \{s_1, \ldots, s_n\}$ and define $X : G \to GL_n(\mathbb{C})$ such that the i, j entry of $X(g)$ is*

$$X(g)_{i,j} = \begin{cases} 1 & \text{if } gs_j = s_i, \\ 0 & \text{otherwise} \end{cases}$$

for all $g \in G$. Then X is a degree n representation of G.

Proof. Let V be the vector space over $\mathbb{C}$ with basis $s_1, \ldots, s_n$. The vector space V contains all formal sums of the form $c_1 s_1 + \cdots + c_n s_n$ where $c_1, \ldots, c_n \in \mathbb{C}$.

Since linear maps are completely determined by their action on a basis, for each $g \in G$ we can define a linear map from V to V by sending s_j to gs_j for all $j = 1, \ldots, n$. Then $X(g)$ is the matrix for this linear map in the basis $s_1, \ldots, s_n$.

By definition, group actions satisfy $\varepsilon s_j = s_j$ and $(gh)s_j = g(hs_j)$. This means $X(\varepsilon) = I_n$ and $X(gh) = X(g)X(h)$, as needed. $\square$

Our next definitions and examples show that a number of interesting representations can be found by choosing a specific group G and set S in the statement of Theorem 4.

Definition. *The defining representation D for the symmetric group S_n is the representation found by taking $G = S_n$ and $S = \{1, \ldots, n\}$ in Theorem 4. That is, the i, j entry of $D(\sigma)$ is*

$$D(\sigma)_{i,j} = \begin{cases} 1 & \text{if } \sigma(j) = i, \\ 0 & \text{otherwise.} \end{cases}$$

For instance, when $n = 3$, the defining representation D for S_3 is

$$D\big((1)\big) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad D\big((1\,2)\big) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$D\big((1\,3)\big) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad D\big((2\,3)\big) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$D\big((1\,2\,3)\big) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad D\big((1\,3\,2)\big) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

The matrices given by the defining representation are also known as permutation matrices.

Example 5. A set partition of n is a set of nonempty disjoint sets with union $\{1, \ldots, n\}$. Let S be the set of all set partitions of 3. The 5 elements in S are:

$$\{\{1\}, \{2\}, \{3\}\}, \quad \{\{1, 2\}, \{3\}\}, \quad \{\{1, 3\}, \{2\}\}, \quad \{\{1\}, \{2, 3\}\}, \quad \{\{1, 2, 3\}\}.$$

The symmetric group S_3 acts on S by changing j in a set partition to $\sigma(j)$ for $j = 1, 2, 3$. A degree 5 representation X for S_3 now comes from Theorem 4. For instance,

$$X\big((1\,3)\big) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Definition. *The left regular representation L is the degree n representation given by Theorem 4 when the finite group $G = \{g_1, \ldots, g_n\}$ acts on itself by left multiplication. This means the i, j entry of $L(g)$ is*

$$L(g)_{i,j} = \begin{cases} 1 & \text{if } gg_j = g_i, \\ 0 & \text{otherwise.} \end{cases}$$

The left regular representation will play an important role in our development of the general theory of matrix representations.

Example 6. The left regular representation L for $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ is

$$L(0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad L(1) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$L(2) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \qquad L(3) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Example 7. The left regular representation for

$$S_3 = \{(1), (1\,2), (1\,3), (2\,3), (1\,2\,3), (1\,3\,2)\}$$

satisfies

$$L\big((1\,3)\big) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

The left regular representation turns any finite group into a group of permutation matrices. This means that a finite group of order n is a subgroup of the symmetric group S_n, a result known as Cayley's theorem.

Theorem 8. *Let L be the left regular representation for $G = \{g_1, \ldots, g_n\}$. Then the matrices $L(g_1), \ldots, L(g_n)$ are linearly independent.*

Proof. Take $c_1, \ldots, c_n \in \mathbb{C}$. The top row of

$$c_1 L(g_1) + \cdots + c_n L(g_n)$$

is a rearrangement of the row $\begin{bmatrix} c_1, \ldots, c_n \end{bmatrix}$ because the $g_a g_j = g_1$ and $g_b g_j = g_1$ implies $g_a = g_b$. Thus the only way this linear combination can be the zero matrix is when all coefficients are 0. $\square$

1.2 New representations from known representations

Definition. *If X is a representation of G, then $\overline{X}$ is the representation defined by $\overline{X}(g) = \overline{X(g)}$ for all $g \in G$. Here, $\overline{c}$ denotes the complex conjugate of $c \in \mathbb{C}$.*

The fact that $\overline{X}$ is a representation follows immediately from the fact that X is a representation and from properties of complex conjugation.

Definition. *If A and B are square matrices, then the direct sum $A \oplus B$ is the block matrix*

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}.$$

If A and B are two representations of G, then we define $A \oplus B$ to be the function which sends g to $A(g) \oplus B(g)$.

Theorem 9. *If A and B are representations of G, then so is $X = A \oplus B$.*

Proof. We see $X(\varepsilon) = I_n$ and

$$\begin{aligned} X(gh) &= \begin{bmatrix} A(gh) & 0 \\ 0 & B(gh) \end{bmatrix} \\ &= \begin{bmatrix} A(g)A(h) & 0 \\ 0 & B(g)B(h) \end{bmatrix} \\ &= \begin{bmatrix} A(g) & 0 \\ 0 & B(g) \end{bmatrix} \begin{bmatrix} A(h) & 0 \\ 0 & B(h) \end{bmatrix} \\ &= X(g)X(h). \end{aligned}$$

□

It can be easily shown that if T is a fixed invertible $n \times n$ matrix and if X is a degree n representation of G, then the function which sends g to $TX(g)T^{-1}$ is also a degree n representation of G. If we think of $X(g)$ as a linear map which sends $\mathbf{v} \in \mathbb{C}^n$ to $X(g)\mathbf{v}$, then changing $X(g)$ to $TX(g)T^{-1}$ has the effect of changing the basis used to represent $\mathbf{v}$.

Definition. *If there is an invertible matrix T such that the representations X and Y satisfy $X(g) = TY(g)T^{-1}$ for all $g \in G$, then X and Y are similar.*

Example 10. Let D be the defining representation for S_3 and take

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Then $X = T\,(\mathrm{sign} \oplus D)\,T^{-1}$ is a new representation. For instance,

$$X\big((1\,2)\big) = \begin{bmatrix} -1 & 0 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ -2 & 0 & 1 & 0 \\ -2 & 0 & 0 & 1 \end{bmatrix} \qquad \text{and} \qquad X\big((1\,2\,3)\big) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Definition. *If $A = (a_{i,j})$ and B are matrices, then the tensor product $A \otimes B$ is the block matrix with i, j block given by $a_{i,j}B$. If A is a representation of G and B is a representation of H, then $A \otimes B$ is defined by*

$$A \otimes B(g, h) = A(g) \otimes B(h)$$

for all $(g, h) \in G \times H$. If A and B are both representations of G, then we define $A \otimes B(g)$ to be $A(g) \otimes B(g)$.

Example 11. If X is the degree 2 representation of S_3 in example 3 and D is the defining representation for S_2, then we have

$$X \otimes D\big((1\,3\,2), (1\,2)\big) = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & -1 \\ -1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

and

$$D \otimes X\big((1\,2),(1\,3\,2)\big) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ -1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Theorem 12. *If A is a representation of G and B a representation of H, then $A \otimes B$ is a representation of $G \times H$.*

Proof. If $X = (x_{i,j})$ and $Y = (y_{i,j})$ are $n \times n$ matrices and Z and W are $m \times m$ matrices, then the i, j block of $(XY) \otimes (ZW)$ is

$$\left(\sum_{k=1}^{n} x_{i,k} y_{k,j}\right) ZW = \sum_{k=1}^{n} x_{i,k} Z\, y_{k,j} W = (x_{i,j} Z)(y_{i,j} W),$$

which is the i, j block of $(X \otimes Z)(Y \otimes W)$. Therefore $(XY) \otimes (ZW) = (X \otimes Z)(Y \otimes W)$ and so

$$\begin{aligned} A \otimes B((g_1, h_1)(g_2, h_2)) &= A(g_1 g_2) \otimes B(h_1 h_2) \\ &= (A(g_1)A(g_2)) \otimes (B(h_1)B(h_2)) \\ &= (A(g_1) \otimes B(h_1))\,(A(g_2) \otimes B(h_2)) \\ &= A \otimes B(g_1, h_1) A \otimes B(g_2, h_2). \end{aligned}$$

This, combined with the observation that if ε and ε' are the identities for G and H, then $A \otimes B(\varepsilon, \varepsilon') = I_n \otimes I_m = I_{nm}$, shows that $A \otimes B$ is indeed a representation of $G \times H$. $\square$

A representation X for G, when restricted to a subgroup $H \leq G$, is also a representation of H. We denote this restriction as $X{\downarrow}_H^G$. Conversely, if X is a representation of a subgroup H of G, then we can use an induced representation to define a representation of G.

Definition. *Let $g_1, \ldots, g_k$ be a left transversal for the subgroup H of G. This means that the cosets of H in G are $g_1 H, \ldots, g_k H$. Let X be a degree n representation of H. The induced representation of X from H to G, denoted $X{\uparrow}_H^G$, has i, j block matrix of size $k \times k$ given by*

$$\left(X{\uparrow}_H^G(g)\right)_{i,j} = \begin{cases} X(g_i^{-1} g g_j) & \text{if } g g_j \in g_i H, \\ 0 & \text{otherwise.} \end{cases}$$

It is a good exercise to verify that the induced representation is indeed a representation.

Example 13. Let $H = \{(1), (2\,3)\}$ be a subgroup of S_3. A transversal for H in G is $(1), (1\,2), (1\,3)$ since $(1)H = \{(1), (2\,3)\}$, $(1\,2)H = \{(1\,2), (1\,2\,3)\}$ and $(1\,3)H = \{(1\,3), (1\,3\,2)\}$. Let X be the representation of H defined by

$$X\big((1)\big) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad X\big((2\,3)\big) = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

Then the induced representation $X\uparrow_H^{S_3}$ of S_3 satisfies

$$
\begin{aligned}
&X\uparrow_H^{S_3}\big((1\,2)\big) \\
&\quad= \begin{bmatrix} 0 & X\big((1)^{-1}(1\,2)(1\,2)\big) & 0 \\ X\big((1\,2)^{-1}(1\,2)(1)\big) & 0 & 0 \\ 0 & 0 & X\big((1\,3)^{-1}(1\,2)(1\,3)\big) \end{bmatrix} \\
&\quad= \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}.
\end{aligned}
$$

Example 14. The left regular representation L is equal to $1\uparrow_{\{\varepsilon\}}^G$.

1.3 Irreducible representations

Definition. *A matrix representation X is reducible if there is a constant matrix T and two matrix representations A and B such that $X(g) = T(A(g) \oplus B(g))T^{-1}$ for all $g \in G$. A representation that is not reducible is irreducible.*

Example 15. Any degree 1 representation is irreducible.

Example 16. Let D be the defining representation for S_3 and let

$$T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}.$$

Then it can be checked that

$$
\begin{aligned}
T^{-1}D\big((1)\big)T &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, & T^{-1}D\big((1\,2)\big)T &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & -1 \end{bmatrix}, \\
T^{-1}D\big((1\,3)\big)T &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & 1 \end{bmatrix}, & T^{-1}D\big((2\,3)\big)T &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \\
T^{-1}D\big((1\,2\,3)\big)T &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & -1 \end{bmatrix}, & T^{-1}D\big((1\,3\,2)\big)T &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 1 & 0 \end{bmatrix}.
\end{aligned}
$$

From this we can see $T^{-1}DT = 1 \oplus X$ where 1 is the trivial representation and X is the degree 2 representation seen in Example 3. This means the defining representation D is reducible with $D = T(1 \oplus X)T^{-1}$.

Theorem 17. *Every representation X is irreducible or there is a fixed matrix T and irreducible representations $X^{(1)}, \ldots, X^{(k)}$ such that*

$$X = T\left(X^{(1)} \oplus \cdots \oplus X^{(k)}\right)T^{-1}.$$

Proof. We use induction on the degree of X. When the degree of X is 1, X must be irreducible.

If X is degree larger than 1 and reducible, then there is a matrix T and a representations A and B with $X = T(A \oplus B)T^{-1}$. Write T and T^{-1} as block matrices of the form

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \qquad \text{and} \qquad T^{-1} = \begin{bmatrix} T'_{11} & T'_{12} \\ T'_{21} & T'_{22} \end{bmatrix}$$

where T_{11} and T'_{11} are the same dimensions as A. This way we have

$$X = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} T'_{11} & T'_{12} \\ T'_{21} & T'_{22} \end{bmatrix}.$$

By the induction hypothesis we can write $A = U \left(X^{(1)} \oplus \cdots \oplus X^{(i)} \right) U^{-1}$ and $B = V \left(X^{(i+1)} \oplus \cdots \oplus X^{(k)} \right) V^{-1}$ where $X^{(1)}, \ldots, X^{(k)}$ are irreducible representations and U, V are fixed matrices.

Therefore X is equal to

$$\begin{bmatrix} T_{11}U & T_{12}V \\ T_{21}U & T_{22}V \end{bmatrix} \begin{bmatrix} X^{(1)} \oplus \cdots \oplus X^{(i)} & 0 \\ 0 & X^{(i+1)} \oplus \cdots \oplus X^{(k)} \end{bmatrix} \begin{bmatrix} U^{-1}T'_{11} & U^{-1}T'_{12} \\ V^{-1}T'_{21} & V^{-1}T'_{22} \end{bmatrix}.$$

This is the desired form because we see

$$\begin{bmatrix} U^{-1}T'_{11} & U^{-1}T'_{12} \\ V^{-1}T'_{21} & V^{-1}T'_{22} \end{bmatrix} \begin{bmatrix} T_{11}U & T_{12}V \\ T_{21}U & T_{22}V \end{bmatrix} = I,$$

meaning that these two matrices are inverses, as needed. □

Our development of the representation theory of the symmetric group will give nice answers to these two fundamental questions about representations:

1. How can we identify if a representation is irreducible?
2. How can we determine how a reducible representation breaks up into a direct sum of irreducible representations?

Definition. *Let X be a representation of G. A subspace $V \subseteq \mathbb{C}^n$ is X-invariant if $X(g)\mathbf{v} \in V$ for all $\mathbf{v} \in V$ and $g \in G$.*

Example 18. Let X be the representation found when G acts on a set S as given in Theorem 4. Each row in $X(g)$ contains exactly one 1 and the rest of the entries are 0. Therefore, if 1 denotes the column vector of all 1's, then $X(g)1 = 1$ for all $g \in G$. This shows $\operatorname{span}\{1\}$ is an X-invariant subspace.

Theorem 19. *A representation X is reducible if and only if there is an X-invariant subspace other than the trivial subspaces $\{0\}$ and $\mathbb{C}^n$.*

Proof. If X is reducible, then $X = T(A \oplus B)T^{-1}$ where A and B are representations and T is a fixed matrix. If A is degree k, let V be the span of the first k columns of T. Then every $\mathbf{v} \in V$ can be written as

$$\mathbf{v} = T \begin{bmatrix} \mathbf{w} \\ 0 \end{bmatrix}$$

where $\mathbf{w}$ is some vector in $\mathbb{C}^k$. Therefore

$$X(g)\mathbf{v} = T \begin{bmatrix} A(g) & 0 \\ 0 & B(g) \end{bmatrix} T^{-1} T \begin{bmatrix} \mathbf{w} \\ 0 \end{bmatrix} = T \begin{bmatrix} A(g)\mathbf{w} \\ 0 \end{bmatrix}$$

for all $g \in G$ and so V is X-invariant.

Now suppose V is a nontrivial X-invariant subspace. Let M be the matrix

$$M = \sum_{g \in G} X(g)^* X(g)$$

where $X(g)^*$ denotes the conjugate transpose of $X(g)$. Not only is this matrix Hermitian, meaning $M = M^*$, but it is also positive definite because

$$\mathbf{v}^* M \mathbf{v} = \sum_{g \in G} \|X(g)\mathbf{v}\|^2 > 0$$

for all nonzero $\mathbf{v} \in \mathbb{C}^n$. This means that we can define a valid inner product on $\mathbb{C}^n$ by setting $\langle \mathbf{v}, \mathbf{u} \rangle = \mathbf{v}^* M \mathbf{u}$. This inner product satisfies

$$\begin{aligned}
\langle \mathbf{v}, X(h)\mathbf{w} \rangle &= \mathbf{v}^* \Big(\sum_{g \in G} X(g)^* X(g) \Big) X(h)\mathbf{w} \\
&= \mathbf{v}^* \Big(\sum_{g' \in G} X(g'h^{-1})^* X(g') \Big) \mathbf{w} \\
&= \mathbf{v}^* X(h^{-1})^* \Big(\sum_{g' \in G} X(g')^* X(g') \Big) \mathbf{w} \\
&= \big\langle X(h^{-1})\mathbf{v}, \mathbf{w} \big\rangle
\end{aligned}$$

for all $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$ and $h \in G$.

The orthogonal complement of V with respect to this inner product is

$$V^{\perp} = \{\mathbf{w} \in \mathbb{C}^n : \langle \mathbf{v}, \mathbf{w} \rangle = 0 \text{ for all } \mathbf{v} \in V\}.$$

The subspace $V^{\perp}$ is also X invariant because $\langle \mathbf{v}, X(h)\mathbf{w} \rangle = \big\langle X(h^{-1})\mathbf{v}, \mathbf{w} \big\rangle = 0$ for all $\mathbf{v} \in V$, $\mathbf{w} \in V^{\perp}$, and $h \in G$.

Therefore if we let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be a basis for V, $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$ be a basis for $V^{\perp}$, and T be the the matrix with columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$, then T is the change of basis matrix from the standard basis to $\mathbf{v}_1, \ldots, \mathbf{v}_n$ for $\mathbb{C}^n$.

In the basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$, the matrix $X(g)$ is block diagonal, and so

$$T^{-1} X(g) T = \begin{bmatrix} A(g) & 0 \\ 0 & B(g) \end{bmatrix}$$

for some $A(g)$ and $B(g)$ for all $g \in G$. This shows $X = T(A \oplus B)T^{-1}$ and that X is reducible. $\square$

Example 20. In the case of the defining representation D for S_3, the matrix M in the proof of Theorem 19 is $6I_3$ and so the inner product in the proof is the standard inner product for $\mathbb{C}^3$ multiplied by 6.

As shown in Example 18, the vector 1 of all 1's is a basis for a D invariant subspace. Thus we can take a basis for the orthogonal complement of $\operatorname{span}\{1\}$ to be $\begin{bmatrix}1, 0, -1\end{bmatrix}^\top$ and $\begin{bmatrix}1, -1, 0\end{bmatrix}^\top$ in order to find the matrix T as given in Example 16.

Theorem 21 (Schur)**.** *Suppose X and Y are irreducible representations of G and T is a fixed matrix such that $TX(g) = Y(g)T$ for all $g \in G$. Then either T is invertible or T is the zero matrix.*

Proof. Suppose T is an $m \times n$ matrix. The kernel of T, denoted $\ker T$, is an X invariant subspace of $\mathbb{C}^n$. Indeed, if $\mathbf{v} \in \ker T$, then

$$T(X(g)\mathbf{v}) = Y(g)T\mathbf{v} = 0,$$

which implies $X(g)\mathbf{v} \in \ker T$ for all $g \in G$. By Theorem 19, either $\ker T = \{0\}$ or $\ker T = \mathbb{C}^n$.

The image of T, denoted $\operatorname{Im} T$, is a Y invariant subspace. Indeed, if $\mathbf{w} \in \operatorname{Im} T$, then there is a $\mathbf{v} \in \mathbb{C}^n$ with $T\mathbf{v} = \mathbf{w}$ and so

$$Y(g)\mathbf{w} = Y(g)T\mathbf{v} = TX(g)\mathbf{v} \in \operatorname{Im} T,$$

for all $g \in G$. By Theorem 19, either $\operatorname{Im} T = \{0\}$ or $\operatorname{Im} T = \mathbb{C}^m$.

If $\operatorname{Im} T = \{0\}$ or $\ker T = \mathbb{C}^n$, then T is the zero matrix. If $\operatorname{Im} T = \mathbb{C}^m$ and $\ker T = \{0\}$, then $m = n$ by the rank nullity theorem and it follows that T is invertible. $\square$

Theorem 22. *Let X be an irreducible representation and T be a matrix such that $TX(g) = X(g)T$ for all $g \in G$. Then T is a multiple of the identity matrix.*

Proof. If λ is an eigenvalue of T, then $T - \lambda I$ is not invertible and

$$(T - \lambda I)X(g) = X(g)(T - \lambda I)$$

for all $g \in G$. Theorem 21 says $T - \lambda I$ is the zero matrix, and so $T = \lambda I$. $\square$

1.4 Characters

Definition. *The character of a representation X is the function $\chi^X : G \to \mathbb{C}$ defined by $\chi^X(g) = \operatorname{trace} X(g)$ for all $g \in G$.*

Example 23. Let D be the defining representation for S_n. Then for any $\sigma \in S_n$,

$$\begin{aligned}\chi^D(\sigma) &= \operatorname{trace} D(\sigma)\\ &= (\text{the number of } i = 1, \ldots, n \text{ with } \sigma(i) = i)\\ &= (\text{the number of fixed points of } \sigma).\end{aligned}$$

More generally, if G acts on a set S and X is the representation given in Theorem 4, then $\chi^X(g)$ is the number of elements in S which are fixed by g. As an important special case, the character of the left regular representation L for any group G satisfies

$$\chi^L(g) = \begin{cases} |G| & \text{if } g = \varepsilon, \\ 0 & \text{otherwise.} \end{cases}$$

Example 24. If X is a degree 1 representation of G, then $\chi^X = X$.

Theorem 25. *If X and Y are similar representations, then $\chi^X = \chi^Y$.*

Proof. Matrix trace satisfies $\mathrm{trace}(AB) = \mathrm{trace}(BA)$ for any matrices A and B for which both the matrix multiplications AB and BA are defined. If X and Y are similar, then there is a fixed matrix T for which $X(g) = TY(g)T^{-1}$ for all $g \in G$. Therefore

$$\chi^X(g) = \mathrm{trace}\, X(g) = \mathrm{trace}(TY(g)T^{-1}) = \mathrm{trace}\, Y(g) = \chi^Y(g)$$

for all $g \in G$. $\square$

We will soon see that the converse to Theorem 25 is also true. This will be a great help when we want to identify all irreducible representations of G up to similarity.

Theorem 26. *If X and Y are representations of G and Z a representation of H, then $\chi^{X \oplus Y} = \chi^X + \chi^Y$ and $\chi^{X \otimes Z} = \chi^X \chi^Z$.*

Proof. As for the direct sum, we have

$$\chi^{X \oplus Y}(g) = \mathrm{trace} \begin{bmatrix} X(g) & 0 \\ 0 & Y(g) \end{bmatrix} = \mathrm{trace}\, X(g) + \mathrm{trace}\, Y(g),$$

which is $\chi^X(g) + \chi^Y(g)$ for all $g \in G$. Similarly, for the tensor product,

$$\chi^{X \otimes Z}(g, h) = \mathrm{trace}\, X(g)\, \mathrm{trace}\, Z(h) = \chi^X(g)\chi^Z(h)$$

for all $g \in G$ and $h \in H$. $\square$

Theorem 27. *If H is a subgroup of G and X is a representation of H, then*

$$\chi^{X\uparrow_H^G}(g) = \frac{1}{|H|} \sum_{h \in G} \chi^X(h^{-1}gh)$$

where $\chi^X(g)$ is taken to be 0 *if $g \notin H$.*

Proof. Let $g_1, \ldots, g_k$ be a left transversal for $H \leq G$. If we say that $X(g)$ is the zero matrix whenever $g \notin H$, then we have

$$\begin{aligned}
\chi^{X\uparrow_H^G}(g) &= \sum_{i=1}^{k} \operatorname{trace} X(g_i^{-1} g g_i) \\
&= \frac{1}{|H|} \sum_{h \in H} \sum_{i=1}^{k} \operatorname{trace} X(h^{-1} g_i^{-1} g g_i h) \\
&= \frac{1}{|H|} \sum_{h \in G} \chi^X(h^{-1} g h). \qquad \square
\end{aligned}$$

Example 28. Let $H = \{(1), (2\,3)\}$ be a subgroup of S_3 and X be the same representation as in Example 13. Then $\chi^X\big((1)\big) = 2$ and $\chi^X\big((2\,3)\big) = 0$ and $\chi^X(\sigma) = 0$ if $\sigma \notin H$, and so, for instance, we have

$$\chi^{X\uparrow_H^{S_3}}\big((1)\big) = \frac{1}{2} \sum_{\sigma \in S_3} \chi^X\big(\sigma^{-1}(1)\sigma\big) = \frac{1}{2}(2 + \cdots + 2) = 6,$$

$$\chi^{X\uparrow_H^{S_3}}\big((1\,2)\big) = \frac{1}{2} \sum_{\sigma \in S_3} \chi^X\big(\sigma^{-1}(1\,2)\sigma\big) = \frac{1}{2}(0 + \cdots + 0) = 0,$$

and

$$\chi^{X\uparrow_H^{S_3}}\big((1\,2\,3)\big) = \frac{1}{2} \sum_{\sigma \in S_3} \chi^X\big(\sigma^{-1}(1\,2\,3)\sigma\big) = \frac{1}{2}(0 + \cdots + 0) = 0.$$

It can also be seen that $\chi^{X\uparrow_H^{S_3}}(\sigma) = 0$ when $\sigma = (1\,3), (2\,3)$, and $(1\,3\,2)$. Therefore $\chi^{X\uparrow_H^{S_3}} = \chi^L$ where L is the left regular representation for S_3. Later we will prove that this implies $X\uparrow_H^{S_3}$ and L are similar.

Definition. *The conjugacy class containing $g \in G$ is $C_g = \{h^{-1} g h : h \in G\}$.*

Example 29. If G is an abelian group, then $C_g = \{g\}$ for all $g \in G$.

Theorem 30. *Characters are constant on conjugacy classes.*

Proof. If $g, h \in G$ are in the same conjugacy class, then there is a $k \in G$ with $g = k^{-1} h k$. Then for any representation X we have

$$\chi^X(g) = \operatorname{trace} X(g) = \operatorname{trace}(X(k)^{-1} X(h) X(k)) = \operatorname{trace} X(h) = \chi^X(h). \qquad \square$$

1.5 The conjugacy classes of the symmetric group

Definition. *An integer partition of n, written $\lambda \vdash n$, is a finite sequence of weakly decreasing nonnegative integers. If $\lambda = (\lambda_1, \ldots, \lambda_k) \vdash n$ with $\lambda_k \neq 0$, then we write $|\lambda| = n$, $\ell(\lambda) = k$, and $\max(\lambda) = \lambda_1$.*

Example 31. The 7 integer partitions of 5 are

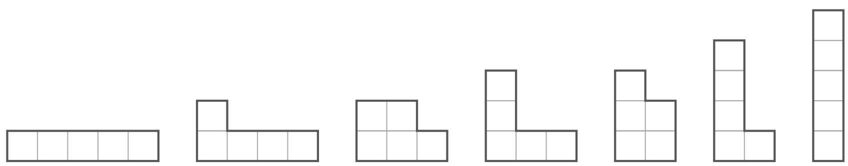
$$(5),\ (4,1),\ (3,2),\ (3,1,1),\ (2,2,1),\ (2,1,1,1),\ (1,1,1,1,1).$$

In this order, the lengths $\ell(\lambda)$ are 1, 2, 2, 3, 3, 4, 5 while the maximum parts are 5, 4, 3, 3, 2, 2, 1.

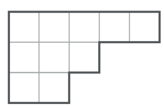
Occasionally it is convenient to denote λ as $1^{m_1}2^{m_2}3^{m_3}\cdots$ if λ has m_i parts of size i. Using this notation, the integer partitions of 5 are

$$5^1,\ 1^1 4^1,\ 2^1 3^1,\ 1^2 3^1,\ 1^1 2^2,\ 1^3 2^1,\ 1^5.$$

Integer partitions can be identified by the corresponding Young diagram; this is a collection of left-justified rows of boxes where row i has λ_i boxes reading from bottom to top. The Young diagrams for all $\lambda \vdash 5$ are



In many places Young diagrams are drawn with the largest row on top; in this way the integer partition $(5,3,2)$ would be drawn as



Definition. *The cycle type of a permutation $\sigma \in S_n$ is the integer partition which records the lengths of the cycles of σ.*

Example 32. The permutation $(1\,3)(2)(4\,6)(5\,7\,9\,8)$ has cycle type $(4,2,2,1)$.

Theorem 33. *The conjugacy class containing $\sigma \in S_n$ is the set of permutations with the same cycle type as σ.*

Proof. It is a good exercise to show that conjugation by τ rearranges the integers in the cycles of σ but preserves the overall cycle structure of σ. □

Example 34. We can index the three distinct conjugacy classes for S_3 by cycle type. They are

$$\begin{aligned} C_{(3)} &= \{(1\,2\,3),(1\,3\,2)\}, \\ C_{(2,1)} &= \{(1\,2),(1\,3),(2\,3)\}, \\ C_{(1,1,1)} &= \{(1)\}. \end{aligned}$$

Definition. *If $\lambda = 1^{m_1}2^{m_2}\cdots$ is an integer partition, we define*

$$z_\lambda = 1^{m_1}2^{m_2}\cdots m_1!m_2!\cdots.$$

Theorem 35. *We have* $|C_\lambda| = n!/z_\lambda$.

Proof. There are $n!$ ways to select a permutation in S_n in one line notation and then place parentheses around the integers to create a permutation of cycle type λ.

Any one of i cyclic rearrangements of a cycle of length i leaves the permutation unchanged; divide by $1^{m_1}2^{m_2}\cdots$ to account for this. Any permutation of the m_i cycles of length i will also not change the permutation; division by $m_1!m_2!\cdots$ will resolve this. Therefore the number of permutations with cycle type λ is

$$\frac{n!}{1^{m_1}2^{m_2}\cdots m_1!m_2!\cdots} = \frac{n!}{z_\lambda},$$

as desired. $\square$

Example 36. The sizes of the conjugacy classes for S_4 are

$$|C_{(4)}| = 6, \quad |C_{(3,1)}| = 8, \quad |C_{(2,2)}| = 3, \quad |C_{(2,1,1)}| = 6, \quad |C_{(1,1,1,1)}| = 1.$$

1.6 Class functions and an inner product for characters

Definition. *Let $C(G)$ be the vector space of all functions $f : G \to \mathbb{C}$ which are constant on the conjugacy classes of G. Such f are called class functions.*

Example 37. A basis for $C(G)$ is

$$\{1_C : C \text{ is a conjugacy class of } G\}$$

where 1_C is the function such that $1_C(g) = 1$ if $g \in C$ and 0 if $g \notin C$. Therefore the dimension of $C(G)$ is the number of conjugacy classes of G.

Example 38. Theorem 30 says that the character of any representation is an element in $C(G)$.

Definition. *For any $f, h \in C(G)$, we define the inner product*

$$\langle f, h\rangle_G = \frac{1}{|G|}\sum_{g\in G} f(g)\overline{h(g)}.$$

If the group G is clear from context, we may omit the subscript and write this as $\langle f, h\rangle$.

It can be readily verified that this has the properties required of an inner product, namely that $\langle cf + f', h\rangle = c\langle f, h\rangle + \langle f', h\rangle$, $\langle f, h\rangle = \overline{\langle h, f\rangle}$, and $\langle f, f\rangle > 0$ if $f \neq 0$ for all $f, f', h \in C(G)$ and all $c \in \mathbb{C}$.

The inner product of characters can be written in a useful alternative form. As permitted by Exercise 9, we will express the inner product of two characters as

$$\langle \chi^X, \chi^Y\rangle = \frac{1}{|G|}\sum_{g\in G} \chi^X(g)\chi^Y(g^{-1})$$

when convenient.

Example 39. Let sign be the sign representation, D the defining representation, and L be the left regular representation for S_3. Then

$$\begin{aligned}
\left\langle \chi^{\mathrm{sign}}, \chi^D \right\rangle &= \frac{1}{6} \sum_{\sigma \in S_3} \chi^{\mathrm{sign}}(\sigma) \chi^D(\sigma^{-1}) = 0, \\
\left\langle \chi^{L}, \chi^D \right\rangle &= \frac{1}{6} \sum_{\sigma \in S_3} \chi^{L}(\sigma) \chi^D(\sigma^{-1}) = 3, \\
\left\langle \chi^{D}, \chi^D \right\rangle &= \frac{1}{6} \sum_{\sigma \in S_3} \chi^{D}(\sigma) \chi^D(\sigma^{-1}) = 2.
\end{aligned}$$

Theorem 40 (Frobenius Reciprocity)**.** *Let H be a subgroup of G, X be a representation of H, and Y a representation of G. Then*

$$\left\langle \chi^{X\uparrow_H^G}, \chi^Y \right\rangle_G = \left\langle \chi^X, \chi^{Y\downarrow_H^G} \right\rangle_H .$$

Proof. Using Theorem 27, we have

$$\begin{aligned}
\left\langle \chi^{X\uparrow_H^G}, \chi^Y \right\rangle_G &= \frac{1}{|G|} \frac{1}{|H|} \sum_{g \in G} \sum_{h \in G} \chi^X(h^{-1}gh) \chi^Y(g^{-1}) \\
&= \frac{1}{|G|} \frac{1}{|H|} \sum_{k \in G} \sum_{h \in G} \chi^X(k) \chi^Y(hk^{-1}h^{-1}) \\
&= \frac{1}{|G|} \frac{1}{|H|} \sum_{k \in H} \sum_{h \in G} \chi^X(k) \chi^Y(k^{-1}) \\
&= \left\langle \chi^X, \chi^{Y\downarrow_H^G} \right\rangle_H .
\end{aligned}$$

□

Example 41. Let H and X be the same subgroup of S_3 and representation as found in Examples 13 and 28. Then

$$\left\langle \chi^{X\uparrow_H^{S_3}}, \chi^{\mathrm{sign}} \right\rangle_{S_3} = \left\langle \chi^X, \chi^{\mathrm{sign}\downarrow_H^{S_3}} \right\rangle_H = \frac{1}{2}(2 \cdot 1 - 0 \cdot 1) = 1.$$

Theorem 42. *If X and Y are irreducible representations, then*

$$\left\langle \chi^X, \chi^Y \right\rangle = \begin{cases} 1 & \text{if } X \text{ and } Y \text{ are similar,} \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Suppose X is degree m and Y is degree n. For any $m \times n$ matrix M let

$$\widehat{M} = \sum_{g \in G} X(g) M Y(g^{-1}).$$

Then for any $h \in G$ we see

$$X(h)\widehat{M} = \sum_{g \in G} X(hg) M Y(g^{-1}) = \sum_{k \in G} X(k) M Y(k^{-1}h) = \widehat{M} Y(h). \tag{1}$$

Theorem 21 gives that $\widehat{M}$ is invertible or equal to the zero matrix.

Case 1: $\widehat{M}$ is invertible. Here we have $X(h) = \widehat{M}Y(h)\widehat{M}^{-1}$ and so X and Y are similar. Since by Theorem 25 we have

$$\langle \chi^X, \chi^Y \rangle = \langle \chi^X, \chi^X \rangle,$$

we need to show that $\langle \chi^X, \chi^X \rangle = 1$.

If we define $\widetilde{M}$ by

$$\widetilde{M} = \sum_{g \in G} X(g) M X(g^{-1}),$$

then the reasoning given in (1) shows $X(h)\widetilde{M} = \widetilde{M}X(h)$ for all $h \in G$. Theorem 22 says there is a constant λ with $\widetilde{M} = \lambda I_n$, so

$$n\lambda = \mathrm{trace}(\lambda I_n) = \mathrm{trace}\,\widetilde{M} = \sum_{g \in G} \mathrm{trace}(X(g)MX(g^{-1})) = |G| \,\mathrm{trace}\, M.$$

Thus $\lambda = |G| \,\mathrm{trace}\, M/n$.

Compare the i, ℓ entries of $\widetilde{M} = \lambda I_n = (|G| \,\mathrm{trace}\, M/n)I_n$ to see

$$\sum_{g \in G} \sum_{j,k} X(g)_{i,j} M_{j,k} X(g^{-1})_{k,\ell} = \begin{cases} \dfrac{|G|}{n}(M_{1,1} + \cdots + M_{n,n}) & \text{if } i = \ell, \\ 0 & \text{if } i \neq \ell, \end{cases} \tag{2}$$

where $M_{j,k}$ denotes the j, k entry of M. The matrix M is arbitrary, so we can compare the coefficients of $M_{i,\ell}$ on both sides of (2) to see

$$\sum_{g \in G} X(g)_{i,i} X(g^{-1})_{\ell,\ell} = \begin{cases} |G|/n & \text{if } i = \ell, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore $\langle \chi^X, \chi^X \rangle$ is equal to

$$\begin{aligned} \frac{1}{|G|} \sum_{g \in G} \mathrm{trace}\, X(g) \,\mathrm{trace}\, X(g^{-1}) &= \frac{1}{|G|} \sum_{i,\ell} \sum_{g \in G} X(g)_{i,i} X(g^{-1})_{\ell,\ell} \\ &= \frac{1}{|G|} \sum_{i=1}^{n} \frac{|G|}{n} \\ &= 1, \end{aligned}$$

as needed.

Case 2: $\widehat{M} = 0$. Here the i, ℓ entry of $\widehat{M}$ is

$$0 = \sum_{g \in G} \sum_{j,k} X(g)_{i,j} M_{j,k} Y(g^{-1})_{k,\ell}$$

and so by comparing the coefficients of $M_{j,k}$ on both sides,

$$\sum_{g \in G} X(g)_{i,j} Y(g^{-1})_{k,\ell} = 0$$

for all i, ℓ. From here we can see that $\langle \chi^X, \chi^Y \rangle$ is equal to

$$\frac{1}{|G|} \sum_{g \in G} \operatorname{trace} X(g) \operatorname{trace} Y(g^{-1}) = \frac{1}{|G|} \sum_{i,\ell} \sum_{g \in G} X(g)_{i,i} Y(g^{-1})_{\ell,\ell} = 0,$$

as needed. $\square$

Example 43. The degree 2 representation X for $\mathbb{Z}_4$ found in Example 2 is similar to itself but since

$$\langle \chi^X, \chi^X \rangle = \frac{1}{4}(2^2 + 0^2 + 0^2 + 2^2) = 2,$$

the representation X must be reducible.

Theorem 44. *Suppose X is a representation of G such that X is similar to*

$$X^{(1)} \oplus \cdots \oplus X^{(1)} \oplus \cdots \oplus X^{(k)} \oplus \cdots \oplus X^{(k)}$$

where $X^{(1)}, \ldots, X^{(k)}$ are irreducible representations and $X^{(i)}$ appears in the direct sum m_i times. Then

1. $\left\langle \chi^X, \chi^{X^{(i)}} \right\rangle = m_i$,
2. $\langle \chi^X, \chi^X \rangle = m_1^2 + \cdots + m_k^2$,
3. *X is irreducible if and only if $\langle \chi^X, \chi^X \rangle = 1$, and*
4. *X and Y are similar if and only if $\chi^X = \chi^Y$.*

Proof. For statement 1.,

$$\begin{aligned}
\left\langle \chi^X, \chi^{X^{(i)}} \right\rangle &= \left\langle m_1 \chi^{X^{(1)}} + \cdots + m_k \chi^{X^{(k)}}, \chi^{X^{(i)}} \right\rangle \\
&= m_1 \left\langle \chi^{X^{(1)}}, \chi^{X^{(i)}} \right\rangle + \cdots + m_k \left\langle \chi^{X^{(k)}}, \chi^{X^{(i)}} \right\rangle \\
&= m_i.
\end{aligned}$$

For statement 2.,

$$\begin{aligned}
\langle \chi^X, \chi^X \rangle &= \left\langle m_1 \chi^{X^{(1)}} + \cdots + m_k \chi^{X^{(k)}}, m_1 \chi^{X^{(1)}} + \cdots + m_k \chi^{X^{(k)}} \right\rangle \\
&= \sum_{i,j} m_i m_j \left\langle \chi^{X^{(i)}}, \chi^{X^{(j)}} \right\rangle \\
&= m_1^2 + \cdots + m_k^2.
\end{aligned}$$

For statement 3., $\langle \chi^X, \chi^X \rangle = 1$ if and only if $m_1^2 + \cdots + m_k^2 = 1$, which can only happen if exactly one $m_i = 1$ and the rest are 0.

Finally, for statement 4., if $\chi^X = \chi^Y$, then

$$\left\langle \chi^X, \chi^{X^{(i)}} \right\rangle = \left\langle \chi^Y, \chi^{X^{(i)}} \right\rangle$$

for all $X^{(i)}$ and so X and Y contain the same number of copies of each irreducible representation and therefore X and Y are similar. $\square$

Example 45. Let X be the representation of S_3 given in Example 3. Then $\langle \chi^X, \chi^X \rangle = 1$ and so X is irreducible.

Example 46. Let X be the representation of S_3 given in Example 3 and let 1 and sign be the trivial and sign representations for S_3. Define a representation Y of S_3 such that

$$Y\big((1\,2)\big) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & -1 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \text{and} \qquad Y\big((2\,3)\big) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

This completely defines the representation Y because the permutations $(1\,2)$ and $(2\,3)$ generate S_3. For instance,

$$Y\big((1\,2\,3)\big) = Y\big((1\,2)\big)Y\big((2\,3)\big) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & -1 & 0 \\ -1 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Since characters are constant on conjugacy classes, $\chi^Y(\sigma) = 0$ if the the cycle type of σ is $(2, 1)$ and $\chi^Y(\sigma) = 1$ if the cycle type of σ is (3). From this it may be calculated that

$$\langle \chi^Y, \chi^Y \rangle = 3.$$

The only possible way $m_1^2 + \cdots + m_k^2$ can equal 3 is if $k = 3$ and $m_1 = m_2 = m_3 = 1$. Indeed, we see that

$$\langle \chi^Y, \chi^1 \rangle = 1, \qquad \langle \chi^Y, \chi^{\mathrm{sign}} \rangle = 1, \qquad \langle \chi^Y, \chi^X \rangle = 1,$$

showing that Y is similar to $1 \oplus \mathrm{sign} \oplus X$. We have successfully identified the irreducible components of Y by calculating inner products of characters.

Example 47. Let L be the left regular representation for G and let X an irreducible representation of G of degree d. Then

$$\langle \chi^L, \chi^X \rangle = \frac{1}{|G|} \left(|G|d + 0 + \cdots + 0 \right) = d,$$

and so X appears d times in L.

Since L is degree $|G|$, this last example shows that, up to similarity, the number of irreducible representations of G is at most $|G|$.

1.7 Character tables

Theorem 48. *The set $\{\chi^X : X \text{ is an irreducible representation of } G\}$ is a basis for $C(G)$.*

Proof. Theorem 44 says that $\langle \chi^X, \chi^Y \rangle = 0$ for all irreducible representations X and Y which are not similar. Therefore the set

$$\{\chi^X : X \text{ is an irreducible representation of } G\}$$

is linearly independent. It remains to be shown that this set spans $C(G)$.

Take $f \in (\operatorname{span}\{\chi^X : X \text{ is irreducible}\})^\perp$, meaning that $\langle \chi^X, f \rangle = 0$ for all irreducible representations X. Furthermore, since every representation is a direct sum of irreducible representations, we have that $\langle \chi^X, f \rangle = 0$ for all representations X, irreducible or not. We will show that $f = 0$.

For any representation X we define

$$M_X = \frac{1}{|G|} \sum_{g \in G} \overline{f(g)} X(g).$$

It follows that

$$X(h) M_X X(h^{-1}) = \frac{1}{|G|} \sum_{g \in G} \overline{f(g)} X(hgh^{-1}) = \frac{1}{|G|} \sum_{k \in G} \overline{f(h^{-1}kh)} X(k) = M_X.$$

If X is irreducible, then Theorem 22 says $M_x = \lambda I_n$ for some $\lambda \in \mathbb{C}$. Since $\operatorname{trace} M_X = \langle \chi^X, f \rangle = 0$, this constant λ is 0 and so $M_X = 0$ when X is irreducible.

If X is reducible, say $X = T(A \oplus B)T^{-1}$, then we have

$$M_X = M_{T(A \oplus B)T^{-1}} = T(M_A \oplus M_B)T^{-1}.$$

Since any representation X is similar to a direct sum of irreducible representations, this implies M_X is the zero matrix even if X were not irreducible. In particular, using the left regular representation L in place of X, we have

$$0 = M_L = \frac{1}{|G|} \sum_{g \in G} \overline{f(g)} L(g).$$

The matrices in the left regular representation are linearly independent by Theorem 8, so $\overline{f(g)} = 0$ for all $g \in G$. Thus f is the zero function. □

Definition. *The character table for G is a table with rows indexed by the characters of the irreducible representations of G, columns indexed by the conjugacy classes of G, with the χ^X, C_g entry the value of χ^X on C_g.*

Theorem 48 says that the number of irreducible representations, up so similarity, is the number of conjugacy classes of G. This means that the character table for G is square.

Example 49. The character table for S_3 is

	$C_{(3)}$	$C_{(2,1)}$	$C_{(1,1,1)}$
χ^1	1	1	1
χ^X	−1	0	2
χ^{sign}	1	−1	1

where X is the irreducible representation in Example 3. These must be all of the characters of irreducible representations since the character table is square.

Example 50. The character table for $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ is

	C_0	C_1	C_2	C_3
χ^1	1	1	1	1
$\chi^{(2)}$	1	i	-1	$-i$
$\chi^{(3)}$	1	-1	1	-1
$\chi^{(4)}$	1	$-i$	-1	i

since there are four conjugacy classes for $\mathbb{Z}_4$ and it can be seen that $\chi^{(1)}, \chi^{(2)}, \chi^{(3)}$, and $\chi^{(4)}$ are irreducible degree 1 representations.

More generally, the function defined by $X(j) = e^{2\pi ij/n}$ is an irreducible representation for the cyclic group $Z_n = \{0, 1, \ldots, n-1\}$ for each i, and so the k, j entry of the character table for Z_n is $e^{2\pi ikj/n}$.

Theorem 51. *If X and Y are irreducible representations for G and H, then $X \otimes Y$ is an irreducible representation of $G \times H$. Furthermore, if $X^{(1)}, \ldots, X^{(n)}$ and $Y^{(1)}, \ldots, Y^{(k)}$ are all of the irreducible representations (up to similarity) for G and H, then all irreducible representations of $G \times H$ are $X^{(i)} \otimes Y^{(j)}$.*

Proof. We see

$$\begin{aligned}
\langle \chi^{X\otimes Y}, \chi^{X\otimes Y} \rangle_{G\times H} &= \frac{1}{|G \times H|} \sum_{(g,h)\in G\times H} \chi^X(g)\chi^Y(h)\chi^X(g^{-1})\chi^Y(h^{-1}) \\
&= \left(\frac{1}{|G|} \sum_{g\in G} \chi^X(g)\chi^X(g^{-1}) \right) \left(\frac{1}{|H|} \sum_{h\in H} \chi^Y(h)\chi^Y(h^{-1}) \right) \\
&= \langle \chi^X, \chi^X \rangle_G \langle \chi^Y, \chi^Y \rangle_H .
\end{aligned}$$

This inner product is 1 because both X and Y are irreducible.

If there are n conjugacy classes for G and k for H, then there are nk conjugacy classes for $G \times H$. There are nk irreducible representations of the form $X^{(i)} \otimes Y^{(j)}$, so this must be all of them up to similarity. □

Example 50 gives the character table for any cyclic group. The fundamental theorem of finite abelian groups says that any finite abelian group is the product of cyclic groups. Theorems 26 and 51 show that the character tables for these abelian groups can be found by multiplying the characters of the cyclic groups together.

Example 52. The characters for $\mathbb{Z}_2 = 0, 1$ are χ^1 and $\chi^{(2)}$ where $\chi^{(2)}(1) = -1$ and so the character table for $\mathbb{Z}_2 \times \mathbb{Z}_2 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ is

	$C_{(0,0)}$	$C_{(0,1)}$	$C_{(1,0)}$	$C_{(1,1)}$
$\chi^1\chi^1$	1	1	1	1
$\chi^1\chi^{(2)}$	1	-1	1	-1
$\chi^{(2)}\chi^1$	1	1	-1	-1
$\chi^{(2)}\chi^{(2)}$	1	-1	-1	1

Example 53. We can deduce the character table for S_4 using the theory we have built so far. There are five conjugacy classes and thus we are looking for the characters of five irreducible representations. Two of them are the trivial and sign representations.

Example 23 shows the defining representation D for S_4 has character giving the number of fixed points in a permutation. We see

$$\left\langle \chi^D, \chi^D \right\rangle = \frac{1}{24}(1 \cdot 4^2 + 6 \cdot 2^2 + 3 \cdot 0^2 + 8 \cdot 1^2 + 6 \cdot 0^2) = 2$$

and so the defining character has two irreducible components. One of them is the trivial representation because $\left\langle \chi^D, \chi^1 \right\rangle = 1$. Call the other irreducible representation A. Then A is degree 3 and $\chi^D = \chi^1 + \chi^A$ and so $\chi^A = \chi^D - \chi^1$. In other words, $\chi^A(\sigma)$ is one less the number of fixed points in σ.

Using Exercise 11, another irreducible representation is $\operatorname{sign} \otimes A$. The character of this representation is $\operatorname{sign} \chi^A$.

There is one final irreducible representation, call it B. Example 47 shows that the left regular representation L contains d copies of each degree d irreducible representation, and so if B is degree d, Theorem 44 gives

$$24 = \left\langle \chi^L, \chi^L \right\rangle = 1^2 + 1^2 + 3^2 + 3^2 + d^2.$$

So $d = 2$. Further, since B is the only remaining irreducible representation and since $\operatorname{sign} \otimes B$ is irreducible, it must be that $B = \operatorname{sign} \otimes B$. This can only happen if $\chi^B(\sigma) = 0$ whenever $\operatorname{sign} \sigma = -1$.

At this point our character table is

	$C_{(4)}$	$C_{(3,1)}$	$C_{(2,2)}$	$C_{(2,1,1)}$	$C_{(1,1,1,1)}$
χ^1	1	1	1	1	1
χ^A	-1	0	-1	1	3
χ^B	0	a	b	0	2
$\chi^{\operatorname{sign} \otimes A}$	1	0	-1	-1	3
$\chi^{\operatorname{sign}}$	-1	1	1	-1	1

where $a, b \in \mathbb{C}$. The values of a and b can be found by taking the inner product of χ^B with χ^1 and χ^A:

$$0 = \left\langle \chi^B, \chi^1 \right\rangle = 2 + 8a + 3b,$$
$$0 = \left\langle \chi^B, \chi^A \right\rangle = 6 - 3b.$$

We have $a = -1$ and $b = 2$. Later we will find a much better combinatorial algorithm to find the entries in the character table for S_n.

2 Tableau combinatorics

2.1 Tabloids

Definition. *A tableau of shape λ is a filling of the cells in the Young diagram of the integer partition λ with positive integers. A tableau T of shape $\lambda \vdash n$ is called*

a. *row constant if the integers in each row of T are all the same,*

b. *row nondecreasing if the integers in each row of T are nondecreasing,*

c. *row increasing if the integers in each row of T are (strictly) increasing,*

d. *column strict if T is row nondecreasing and the integers strictly increase up each column, and*

e. *a tabloid if T is row increasing and if each of the integers $1, \ldots, n$ appear exactly once in T.*

Example 54. The 6 possible tableaux of shape $(2, 2)$ filled with $1, 1, 2, 2$ are:

$$\begin{array}{|c|c|}\hline 2 & 2 \\ \hline 1 & 1 \\ \hline \end{array} \qquad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 1 & 2 \\ \hline \end{array} \qquad \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 1 & 2 \\ \hline \end{array} \qquad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 & 1 \\ \hline \end{array} \qquad \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 2 & 1 \\ \hline \end{array} \qquad \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 & 2 \\ \hline \end{array}$$

Example 55. The 10 possible tabloids of shape $(3, 2)$ are:

$$\begin{array}{ccc} 4 & 5 & \\ 1 & 2 & 3 \end{array} \qquad \begin{array}{ccc} 3 & 5 & \\ 1 & 2 & 4 \end{array} \qquad \begin{array}{ccc} 3 & 4 & \\ 1 & 2 & 5 \end{array} \qquad \begin{array}{ccc} 2 & 5 & \\ 1 & 3 & 4 \end{array} \qquad \begin{array}{ccc} 2 & 4 & \\ 1 & 3 & 5 \end{array}$$

$$\begin{array}{ccc} 2 & 3 & \\ 1 & 4 & 5 \end{array} \qquad \begin{array}{ccc} 1 & 5 & \\ 2 & 3 & 4 \end{array} \qquad \begin{array}{ccc} 1 & 3 & \\ 2 & 3 & 5 \end{array} \qquad \begin{array}{ccc} 1 & 4 & \\ 2 & 4 & 5 \end{array} \qquad \begin{array}{ccc} 1 & 2 & \\ 3 & 4 & 5 \end{array}$$

Definition. *The group S_n acts on the set of tabloids of shape $\mu \vdash n$ by using $\sigma \in S_n$ to permute the elements in T and then sorting the integers in each row of T into increasing order. For instance,*

$$(1\,3\,2\,4) \begin{array}{ccc} 3 & 5 & \\ 1 & 2 & 4 \end{array} = \begin{array}{ccc} 2 & 5 & \\ 1 & 3 & 4 \end{array}.$$

Theorem 4 says that every group action corresponds to a representation. This means that for every $\mu \vdash n$ there is a representation X of S_n given by the group action of S_n on tabloids of shape μ. We define ψ^μ to be the character of this representation.

Example 56. Let X be the degree 10 representation of S_5 found by considering the action of S_5 on all the tabloids of shape $(3, 2)$. Listing these tabloids in the order found in Example 55, we have

$$X\big((1\,3\,2\,4)\big) = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Therefore the character $\psi^{(3,2)}\big((1\,3\,2\,4)\big) = 0$. This can also be seen since the permutation $(1\,3\,2\,4)$ does not fix any tabloids of shape $(3,2)$.

Example 57. This is an especially enlightening example in our development. By considering tabloids which are fixed under the action of S_n, we can find that the characters ψ^μ for all $\mu \vdash 4$ on the conjugacy classes of S_4 are:

	$C_{(4)}$	$C_{(3,1)}$	$C_{(2,2)}$	$C_{(2,1,1)}$	$C_{(1,1,1,1)}$
$\psi^{(4)}$	1	1	1	1	1
$\psi^{(3,1)}$	0	1	0	2	4
$\psi^{(2,2)}$	0	0	2	2	6
$\psi^{(2,1,1)}$	0	0	0	2	12
$\psi^{(1,1,1,1)}$	0	0	0	0	24

This is not a character table because all of these representations except for $\psi^{(4)}$ are reducible.

As μ changes from the integer partition (4) to the integer partition $(1,1,1,1)$, the characters ψ^μ change from the trivial representation to the left regular representation. Furthermore, taking inner products with the irreducible characters given for S_4 found in Example 53, we can calculate

$$\begin{aligned} \psi^{(4)} &= \chi^1, \\ \psi^{(3,1)} &= \chi^1 + \chi^A, \\ \psi^{(2,2)} &= \chi^1 + \chi^A + \chi^B, \\ \psi^{(2,1,1)} &= \chi^1 + 2\chi^A + \chi^B + \chi^{\mathrm{sign}\otimes A}, \\ \psi^{(1,1,1,1)} &= \chi^1 + 3\chi^A + 2\chi^B + 3\chi^{\mathrm{sign}\otimes A} + \chi^{\mathrm{sign}}, \end{aligned}$$

showing that as we change from (4) to $(1,1,1,1)$, the number of distinct irreducible representations in ψ^μ increase by 1 at each step.

From here it can be seen that the characters ψ^μ for $\mu \vdash 4$ are linearly independent, and since there are the right number of them, these functions are a basis for

$C(S_4)$. The transition, or change of basis matrix, which turns ψ^μ into the basis given by the irreducible characters is the matrix

$$\begin{array}{c} \\ \chi^1 \\ \chi^A \\ \chi^B \\ \chi^{\mathrm{sign}\otimes A} \\ \chi^{\mathrm{sign}} \end{array} \begin{array}{c} \begin{array}{ccccc} \psi^{(4)} & \psi^{(3,1)} & \psi^{(2^2)} & \psi^{(2,1^2)} & \psi^{(1^4)} \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}.$$

Definition. *Given an integer partition $\mu = (\mu_1, \mu_2, \dots)$, the Young subgroup*

$$S_\mu = S_{\{1,\dots,\mu_1\}} \times S_{\{\mu_1+1,\dots,\mu_1+\mu_2\}} \times \cdots .$$

where for any finite set A we let S_A denote the set of permutations of A.

Example 58. The Young subgroup $S_{(3,2,2)}$ is equal to

$$S_{(3,2,2)} = S_{\{1,2,3\}} \times S_{\{4,5\}} \times S_{\{6,7\}}.$$

Elements in $S_{(3,2,2)}$ are technically ordered triples such as $\big((1\,3), (4\,5), (6)\big)$, but we will simply denote such an element as $(1\,3)(4\,5)(6)$. With this convention it is clear that S_μ is a subgroup of S_n for all $\mu \vdash n$.

Theorem 59. *Let $T_1, \dots, T_k$ be all tabloids of shape $\mu = (\mu_1, \mu_2, \dots) \vdash n$. Let $\pi_i \in S_n$ satisfy $\pi_i T_1 = T_i$. Then $\pi_1, \dots, \pi_k$ is a left transversal for S_μ in S_n.*

Proof. This is the correct number of cosets because

$$\frac{|S_n|}{|S_\mu|} = \frac{n!}{\mu_1!\mu_2!\cdots} = \binom{n}{\mu_1, \mu_2, \dots}$$

and this multinomial coefficient is the number of ways to select μ_1 numbers from $\{1, \dots, n\}$ for the bottom row of T, then select μ_2 numbers from the remaining integers for the second row of T, etc.

Are any of the cosets $\pi_1 S_\mu, \dots, \pi_k S_\mu$ the same? Without loss of generality, take T_1 to be the tabloid with rows containing the integers $1, \dots, \mu_1$, then $\mu_1+1, \dots, \mu_1+\mu_2$, and so on. This way $\sigma T_1 = T_1$ for any $\sigma \in S_\mu$. Therefore, if $\pi_i S_\mu = \pi_j S_\mu$ for some $i \neq j$, then $\pi_i(1) = \pi_j \sigma$ for some $\sigma \in S_\mu$. This would mean that

$$T_i = \pi_i T_1 = \pi_j \sigma T_1 = \pi_j T_1 = T_j,$$

and therefore all cosets are distinct. □

Theorem 60. *The induced representation $1\!\uparrow_{S_\mu}^{S_n}$ has character ψ^μ.*

Proof. Let $\pi_1, \ldots, \pi_k$ be a left transversal for S_μ in S_n and let $T_1, \ldots, T_k$ be the tabloids of shape μ. Then the i, j entry of $1\!\uparrow_{S_\mu}^{S_n}(\sigma)$ is

$$\begin{aligned}\left(1\!\uparrow_{S_\mu}^{S_n}(\sigma)\right)_{i,j} &= \begin{cases}1 & \text{if } \sigma\pi_j \in \pi_i S_\mu\\ 0 & \text{otherwise}\end{cases}\\ &= \begin{cases}1 & \text{if } \sigma T_j = T_i\\ 0 & \text{otherwise,}\end{cases}\end{aligned}$$

which is exactly the representation coming from the group action of S_n in the definition of the character ψ^μ. $\square$

2.2 The irreducible representations of the symmetric group

Definition. *Let $\mathcal{T}^\lambda$ be the vector space of all formal complex linear combinations of tabloids of shape λ.*

Example 61. The vector space $\mathcal{T}^{(2,1)}$ consists of all vectors of the form

$$c_1 \begin{array}{|c|c} \cline{1-1} 3 & \\ \hline 1 & \multicolumn{1}{c|}{2} \\ \hline \end{array} + c_2 \begin{array}{|c|c} \cline{1-1} 2 & \\ \hline 1 & \multicolumn{1}{c|}{3} \\ \hline \end{array} + c_3 \begin{array}{|c|c} \cline{1-1} 1 & \\ \hline 2 & \multicolumn{1}{c|}{3} \\ \hline \end{array}$$

for $c_1, c_2, c_3 \in \mathbb{C}$.

Definition. *Let T be a tableau (or tabloid) filled with $1, \ldots, n$. The column stabilizer C_T is the set of permutations $\sigma \in S_n$ such that i and $\sigma(i)$ are in the same column of T for all $i = 1, \ldots, n$. Define the Young symmetrizer to be the operator on $\mathcal{T}^\lambda$ given by $\kappa_T = \sum_{\sigma \in C_T} (\operatorname{sign} \sigma)\sigma$.*

Example 62. Take $T = \begin{array}{|c|c|} \hline 3 & 4 \\ \hline 1 & 2 \\ \hline \end{array}$ and $T' = \begin{array}{|c|c} \cline{1-1} 3 & \\ \cline{1-1} 2 & \\ \hline 1 & \multicolumn{1}{c|}{4} \\ \hline \end{array}$. Then the column stabilizer C_T is $\{(1), (1\,3), (2\,4), (1\,3)(2\,4)\}$ and so

$$k_T T' = \begin{array}{|c|c} \cline{1-1} 3 & \\ \cline{1-1} 2 & \\ \hline 1 & \multicolumn{1}{c|}{4} \\ \hline \end{array} - \begin{array}{|c|c} \cline{1-1} 1 & \\ \cline{1-1} 2 & \\ \hline 3 & \multicolumn{1}{c|}{4} \\ \hline \end{array} - \begin{array}{|c|c} \cline{1-1} 3 & \\ \cline{1-1} 4 & \\ \hline 1 & \multicolumn{1}{c|}{2} \\ \hline \end{array} + \begin{array}{|c|c} \cline{1-1} 1 & \\ \cline{1-1} 4 & \\ \hline 2 & \multicolumn{1}{c|}{3} \\ \hline \end{array}.$$

The column stabilizer $C_{T'}$ is $\{(1), (1\,2), (1\,3), (2\,3), (1\,2\,3), (1\,3\,2)\}$ and so

$$k_{T'} T = \begin{array}{|c|c|} \hline 3 & 4 \\ \hline 1 & 2 \\ \hline \end{array} - \begin{array}{|c|c|} \hline 3 & 4 \\ \hline 1 & 2 \\ \hline \end{array} - \begin{array}{|c|c|} \hline 1 & 4 \\ \hline 2 & 3 \\ \hline \end{array} - \begin{array}{|c|c|} \hline 2 & 4 \\ \hline 1 & 3 \\ \hline \end{array} + \begin{array}{|c|c|} \hline 1 & 4 \\ \hline 2 & 3 \\ \hline \end{array} + \begin{array}{|c|c|} \hline 2 & 4 \\ \hline 1 & 3 \\ \hline \end{array} = 0.$$

Theorem 63. *Take $\lambda, \mu \vdash n$ with $\lambda \neq \mu$. Then either $k_T T' = 0$ for any tabloids T, T' of shape λ, μ or $k_{T'} T = 0$ for any tabloids T, T' of shape λ, μ.*

Proof. If the first column in the Young diagram of λ is taller than that of μ, then all of the integers in the first column of any tabloid T of shape λ cannot appear in different rows of any tabloid T' of shape μ. Therefore there is a pair of integers i and j in the first column of T and which appear in the same row of T'. Then

$$\kappa_T(i\,j) = \sum_{\sigma \in C_T} (\operatorname{sign} \sigma)\sigma(i\,j) = -\kappa_T,$$

and so we see $\kappa_T T' = -\kappa_T(i\,j)T' = -\kappa_T T'$. This implies $\kappa_T T' = 0$ for all T, T' of shape λ, μ.

If the first column in the Young diagram of μ is taller than that of λ, then a symmetric argument shows $\kappa_{T'} T = 0$.

If the first columns in the Young diagrams of λ and μ are the same, to the second columns and use a similar argument as above to show that either $k_T T' = 0$ or $k_{T'} T = 0$. Continuing in this manner, examining each column in sequence, we find $k_T T' = 0$ or $k_{T'} T = 0$ since eventually two column heights in λ and μ must differ. □

Theorem 64. *If T and T' are tabloids of same shape λ, then $\kappa_T T'$ is equal to* 0*,* $(\kappa_T T)$*, or* $-(\kappa_T T)$*.*

Proof. If there are integers i and j which are in the same column of T and also in the same row of T', then we find $\kappa_T T' = 0$ in the same way as shown in the proof of Theorem 63.

Suppose no such pair of integers i and j exist. If the integer i appears in the same column as j in T, then i and j must be in different rows of T'. If i is in a below j in T' but i is in a row above j in T, then the transposition $(i\,j) \in C_T$ makes i appear below j in $(i\,j)T$.

Examining all such pairs i and j in the same column of T, we can apply the transpositions $(i\,j)$ to T when necessary to show that there exists a $\tau \in C_T$ such that the rows of τT and T', meaning that τT and T' are the same tabloids. Then we have

$$\kappa_T T' = \kappa_T \tau T = \sum_{\sigma \in C_T} (\operatorname{sign} \sigma)\sigma\tau T = (\operatorname{sign} \tau)\kappa_T T,$$

as needed. □

Definition. *The subspace V^λ of $\mathcal{T}^\lambda$ is*

$$V^\lambda = \operatorname{span}\{\kappa_T T : T \text{ is a tabloid of shape } \lambda\}.$$

Example 65. The three possible tabloids when $\lambda = (2, 1)$ are:

$$T_1 = \begin{array}{|c|c}\hline 3 & \\ \hline 1 & \multicolumn{1}{c|}{2} \\ \hline\end{array} \qquad \text{with} \qquad \kappa_{T_1} T_1 = \begin{array}{|c|c}\hline 3 & \\ \hline 1 & \multicolumn{1}{c|}{2} \\ \hline\end{array} - \begin{array}{|c|c}\hline 1 & \\ \hline 2 & \multicolumn{1}{c|}{3} \\ \hline\end{array},$$

$$T_2 = \begin{array}{|c|c}\hline 2 & \\ \hline 1 & \multicolumn{1}{c|}{3} \\ \hline\end{array} \qquad \text{with} \qquad \kappa_{T_2} T_2 = \begin{array}{|c|c}\hline 2 & \\ \hline 1 & \multicolumn{1}{c|}{3} \\ \hline\end{array} - \begin{array}{|c|c}\hline 1 & \\ \hline 2 & \multicolumn{1}{c|}{3} \\ \hline\end{array},$$

$$T_3 = \ytableaushort{1,23} \qquad \text{with} \qquad \kappa_{T_3} T_3 = \ytableaushort{1,23} - \ytableaushort{2,13}.$$

Since $\kappa_{T_3} T_3 = -\kappa_{T_2} T_2$, a basis for $V^{(2,1)}$ is $\{\kappa_{T_1} T_1, \kappa_{T_2} T_2\}$. In other words, $V^{(2,1)}$ is all linear combinations of the form $c_1(\kappa_{T_1} T_1) + c_2(\kappa_{T_2} T_2)$ for $c_1, c_2 \in \mathbb{C}$.

Definition. *A standard tableau of shape $\lambda \vdash n$ is a tabloid such that integers increase within columns (read bottom to top) as well as rows.*

Theorem 66. *A basis for V^λ is $\{\kappa_T T : T$ is a standard tableau of shape $\lambda\}$.*

Proof. Define an ordering among standard tableaux of shape λ such that $T < T'$ whenever the permutation found by reading the rows of T from bottom to top is lexicographically smaller than the permutation found by reading the rows of T'. For instance, we have

$$\ytableaushort{45,123} < \ytableaushort{35,124} < \ytableaushort{34,125} < \ytableaushort{25,134} < \ytableaushort{24,135} .$$

If T is one of these tableaux, then $\sigma T > T$ for all non-identity $\sigma \in C_T$. So, if T' is the smallest possible tableaux, then by looking at the coefficient of T' in the linear combination

$$\sum c_T(\kappa_T T) = 0$$

we can see that $c_{T'} = 0$. Therefore $\kappa_{T'} T'$ cannot appear as a nonzero term in any linear combination of vectors $\kappa_T v_T$ which equal 0.

Continue this line of reasoning with the next smallest tableaux and working our way up, we can see that whenever we have $\sum c_T(\kappa_T T) = 0$, all coefficients c_T must equal 0. Thus we have linear independence.

These vectors also span V^λ by the observation that $k_T T$ is a multiple of $k_{T'} T'$ whenever T and T' have the same columns (as sets), and so we have a basis. $\square$

Example 67. The group action of S_n on tabloids can be extended by linearity to act on $\mathcal{T}^\lambda$. For example, if we let

$$T = \ytableaushort{45,123},$$

then the permutation $(3\,5)$ sends

$$\kappa_T T = \ytableaushort{45,123} - \ytableaushort{15,234} - \ytableaushort{24,135} + \ytableaushort{12,345}$$

to

$$(3\,5)\kappa_T T = \ytableaushort{34,125} - \ytableaushort{13,245} - \ytableaushort{24,135} + \ytableaushort{12,345}.$$

Theorem 68. *If $\mathbf{w} \in V^\lambda$ satisfies $\kappa_T \mathbf{w} = 0$ for all tabloids T of shape λ, then $\mathbf{w}$ is the zero vector.*

Proof. Define an inner product on $\mathcal{T}^\lambda$ by

$$\langle T, T' \rangle = \begin{cases} 1 & \text{if } T = T', \\ 0 & \text{otherwise.} \end{cases}$$

This inner product satisfies $\langle T, T' \rangle = \langle \sigma T, \sigma T' \rangle$ for all $\sigma \in S_n$. Then we have

$$\begin{aligned} \langle \mathbf{w}, \kappa_T T \rangle &= \sum_{\sigma \in C_T} (\operatorname{sign} \sigma) \langle \mathbf{w}, \sigma T \rangle \\ &= \sum_{\sigma \in C_T} (\operatorname{sign} \sigma) \left\langle \sigma^{-1} \mathbf{w}, \sigma^{-1} \sigma T \right\rangle \\ &= \sum_{\sigma^{-1} \in C_T} (\operatorname{sign} \sigma^{-1}) \left\langle \sigma^{-1} \mathbf{w}, T \right\rangle \\ &= \langle \kappa_T \mathbf{w}, T \rangle , \end{aligned}$$

which is equal to 0. Since vectors of the form $k_T T$ span V^λ, the vector $\mathbf{w}$ must be in the orthogonal complement $(V^\lambda)^\perp$. The only vector $\mathbf{w}$ which is in both V^λ and $(V^\lambda)^\perp$ is the zero vector. □

Definition. *For each $\sigma \in S_n$, define a linear map on V^λ by sending $\kappa_T T$ to $\sigma(\kappa_T T)$. The Specht representation $S^\lambda(\sigma)$ is the matrix for this linear map in the basis $\{\kappa_T T :$ T is a standard tableau of shape $\lambda\}$.*

Example 69. If $\lambda = (n)$, then there is only one possible tabloid T of shape λ, the tabloid containing $1, \ldots, n$ in one row. Then $\kappa_T T = T$ and the vector space V^λ is simply the span of T. The Specht representation in this case is the trivial representation.

Example 70. At the other extreme, if $\lambda = (1^n)$, there is only one possible column strict tableau T, the tableau containing $1, \ldots, n$ in one column. The vector space V^λ is simply the span of $\kappa_T T$. The Specht representation in this case is the sign representation.

Example 71. Consider $\lambda = (2, 1)$ and take T_1 and T_2 as in example 65. The permutation $(1\,2)$ sends $\kappa_{T_1} T_1$ to $(\kappa_{T_1} T_1) - (\kappa_{T_2} T_2)$ and sends $\kappa_{T_2} T_2$ to $-(\kappa_{T_2} T_2)$. These give us the columns in the matrix

$$S^{(2,1)}\big((1\,2)\big) = \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}.$$

Similarly, $(1\,2\,3)$ sends $\kappa_{T_1} T_1$ to $-(\kappa_{T_2} T_2)$ and $\kappa_{T_2} T_2$ to $(\kappa_{T_1} T_1) - (\kappa_{T_2} T_2)$, so

$$S^{(2,1)}\big((1\,2\,3)\big) = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}.$$

Indeed, it turns out that $S^{(2,1)}$ is the irreducible representation in Example 3.

Theorem 72. *The Specht representation S^λ is irreducible.*

Proof. Let $W \neq \{0\}$ be a subspace of V^λ invariant under the action of σ for all $\sigma \in S_n$. Since W corresponds to a S^λ invariant subspace of $\mathbb{C}^k$, Theorem 19 says that in order to prove that S^λ is irreducible we can show that $W = V^\lambda$.

The operator κ_T for any tabloid T is a linear combination of permutations σ, and since W is invariant under the action of permutations, W is also invariant under the action of κ_T.

Take $\mathbf{w} \in W$ to be nonzero. The vector $\mathbf{w}$ is a linear combination of terms of the form $\kappa_T T$ where T is a standard tableau, which in turn are linear combinations of tabloids. Therefore, by Theorem 64,

$$\kappa_T \mathbf{w} = c(\kappa_T T) \in W$$

for some $c \in \mathbb{C}$. Since $\mathbf{w}$ is nonzero, Theorem 68 says there must be a T' such that this c is not zero.

Therefore $\sigma(\kappa_{T'} T') \in W$ for all $\sigma \in S_n$. Let T be any other standard tableau of shape λ and take σ to be the permutation such that $\sigma T' = T$. It follows that $\tau \in C_{T'}$ if and only if $\sigma\tau \in C_T$ and therefore

$$\begin{aligned}
\sigma(\kappa_{T'} T') &= \sum_{\tau \in C_{T'}} (\operatorname{sign} \tau) \sigma\tau T' \\
&= (\operatorname{sign} \sigma^{-1}) \sum_{\alpha \in C_T} (\operatorname{sign} \alpha) \alpha T \\
&= (\operatorname{sign} \sigma^{-1}) \kappa_T T'
\end{aligned}$$

where we re-indexed the sum by taking $\alpha = \sigma\tau$. Since $\sigma(\kappa_{T'} T')$ is not zero for any $\sigma \in S_n$, Theorem 64 tells us that have that $\sigma(\kappa_{T'} T') = \pm\kappa_T T$. This means that for every standard tableau T, the vector $\kappa_T T$ is in W. These vectors are a basis for V^λ, and so $W = V^\lambda$. $\square$

Theorem 73. *The representations S^λ and S^μ are not similar when $\lambda \neq \mu$.*

Proof. If $\lambda \neq \mu$, Theorem 63 says that either $k_T T' = 0$ for any tabloids T, T' of shape λ, μ or $k_{T'} T = 0$ for any tabloids T, T' of shape λ, μ. Without loss of generality, assume that $k_T T' = 0$ for any tabloids T, T' of shape λ, μ.

Suppose S^λ and S^μ are similar. This means there is a matrix Θ such that

$$\Theta S^\lambda(\sigma)\mathbf{w} = S^\mu(\sigma)\Theta\mathbf{w}$$

for all $\sigma \in S_n$ and vectors $\mathbf{w}$. Translating this statement in the language of the vector spaces, there is a invertible linear map (a vector space isomorphism) $\vartheta : V^\lambda \to V^\mu$ such that

$$\vartheta(\sigma\mathbf{w}) = \sigma\vartheta(\mathbf{w})$$

for all $\sigma \in S_n$ and $\mathbf{w} \in V^\lambda$. Since κ_T is a linear combination of permutations, the σ in this last expression can be replaced with κ_T for any tabloid T.

If T is a tabloid of shape λ,

$$\vartheta(\kappa_T \mathbf{w}) = \kappa_T \vartheta(\mathbf{w}) = 0$$

since $\kappa_T T' = 0$ for all T, T' of shape λ, μ. This means that $\kappa_T \mathbf{w} = 0$ for all tabloids T of shape λ and all $\mathbf{w} \in V^\lambda$. This contradictions Theorem 68. $\square$

Theorems 72 and 73 combine to tell us that the Specht representations $\{S^\lambda : \lambda \vdash n\}$ are a complete list of irreducible representations—each S^λ is irreducible and there are as many of them as there are conjugacy classes.

At this point the characters of these Specht representations are not particularly easy to calculate. Later we will find a simple combinatorial method to quickly find the value of the character of S^λ on the conjugacy class C_μ.

3 Symmetric functions

3.1 Standard bases for symmetric functions

Definition. *The weight of a tableau T is*

$$w(T) = \prod_{\text{cells } c \text{ in } T} x_{T_c}$$

where T_c is the integer in the cell c in T. For any integer partition λ, we define

a. *the power symmetric function p_λ to equal $\sum w(T)$ where the sum runs over all row constant tableaux T of shape λ,*

b. *the homogeneous symmetric function h_λ to equal $\sum w(T)$ where the sum runs over all row nondecreasing tableaux T of shape λ,*

c. *the elementary symmetric function e_λ to equal $\sum w(T)$ where the sum runs over all row increasing tableaux T of shape λ, and*

d. *the Schur symmetric function s_λ to equal $\sum w(T)$ where the sum runs over all column strict tableaux T of shape λ.*

Example 74. All column strict tableaux of shape $(4, 2, 1)$ with weight $x_1^3 x_2^2 x_3 x_4$ are

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

and so the coefficient of $x_1^3 x_2^2 x_3 x_4$ in $s_{(4,2,1)}$ is 4.

Definition. *A polynomial f in the variables $x_1, \dots, x_N$ is a symmetric polynomial if*

$$f(x_1, \dots, x_N) = f(x_{\sigma(1)}, \dots, x_{\sigma(N)})$$

for all $\sigma \in S_N$. We define $\Lambda_n(x_1, \dots, x_N)$ to be the vector space of symmetric functions with each monomial of total degree n. We will always assume that N is much larger than n.

Example 75. The polynomial

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 + 3x_1x_2 + 3x_1x_3 + 3x_2x_3$$

is a symmetric polynomial. Every monomial in f has total degree 2 and so f is an element in $\Lambda_2(x_1, x_2, x_3)$.

Definition. *Let $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ be an integer partition of n. The monomial symmetric function $m_\lambda = m_\lambda(x_1, \ldots, x_N)$ is the sum of all the monomials with exponents that can be rearranged to give λ.*

Example 76. The monomial symmetric polynomial $m_{(2,1)}(x_1, x_2, x_3)$ is

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

Theorem 77. *The set $\{m_\lambda : \lambda \vdash n\}$ is a basis for Λ_n and therefore the dimension of Λ_n is the number of integer partitions of n.*

Proof. The polynomials m_λ and m_μ have no monomials in common if $\lambda \neq \mu$, so the functions $\{m_\lambda : \lambda \vdash n\}$ are linearly independent.

If $f \in \Lambda_n$ has a monomial of the form $c_\lambda x_1^{\lambda_1} \cdots x_\ell^{\lambda_\ell}$ for some constant c_λ, then it must also contain $c_\lambda m_\lambda$. Thus $\{m_\lambda : \lambda \vdash n\}$ spans Λ_n. $\square$

Theorem 78. *The Schur symmetric function $s_\lambda \in \Lambda_n$ for all $\lambda \vdash n$.*

Proof. Since every element in S_n is a product of adjacent transpositions, it is enough to show that s_λ is unchanged under the action of switching x_i and x_{i+1} for all positive integers i. So we need to show that for every column strict tableau with k occurrences of i and j occurrences of $i+1$, there is a corresponding column strict tableau of the same shape with j occurrences of i and k occurrences of $i+1$.

Take T to be a column strict tableau of shape λ. The appearances of i in relationship to the appearances of $i+1$ in T must look something like the appearances of the 3's and 4's below:

$$\begin{array}{cccccccccccccccc}
3 & 3 & 4 & 4 & 4 & 4 & & & & & & & & & & \\
 & & & 3 & 3 & 3 & 4 & 4 & 4 & & & & & & & \\
 & & & & & & & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 4 & 4
\end{array}$$

Each row in T may have a sequence of i's followed by a sequence of $i+1$'s. Rows are aligned so that an i cannot appear atop another i in the row below.

Suppose a given row of T contains a sequence of k i's followed by j $i+1$'s such that none of these i's or $i+1$'s are immediately above or below cells containing an i or $i+1$. Change this row of T so that it now contains j i's followed by k $i+1$'s. Make this change for every row of T to create the column strict tableau T'. For example, changing the 3's and 4's in the cells displayed above produces:

$$\begin{array}{cccccccccccccccc}
3 & 4 & 4 & 4 & 4 & 4 & & & & & & & & & & \\
 & & & 3 & 3 & 3 & 3 & 4 & 4 & & & & & & & \\
 & & & & & & & 3 & 3 & 3 & 3 & 4 & 4 & 4 & 4 & 4
\end{array}$$

The new column strict tableau has the same shape as before with the number of i's and $i+1$'s switched, as needed. $\square$

Variations on the proof of Theorem 78 can also be used to show that p_λ, h_λ, and e_λ are also symmetric functions of degree n for all $\lambda \vdash n$.

Definition. *The Kostka number $K_{\lambda,\mu}$ is the number of column strict tableau of shape λ and content μ. This is the coefficient of m_λ in s_μ.*

The Kostka matrix is the square matrix indexed by integer partitions of n written in reverse lexicographic order with λ, μ entry equal to $K_{\mu,\lambda}$.

Example 79. The Kostka matrix with rows indexed by λ (the content) and columns indexed by μ (the shape) when $n = 4$ is

$$\begin{array}{c} \\ (4) \\ (3,1) \\ (2^2) \\ (2,1^2) \\ (1^4) \end{array} \begin{array}{c} \begin{array}{ccccc} (4) & (3,1) & (2^2) & (2,1^2) & (1^4) \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 1 & 0 \\ 1 & 3 & 2 & 3 & 1 \end{bmatrix} \end{array}.$$

The astute reader can find a relationship between the Kostka matrix when $n = 4$ and the transition matrix between the bases $\{\psi^\mu : \mu \vdash n\}$ and $\{\chi^\lambda : \lambda \vdash n\}$ for $C(S_4)$ given in Example 57.

The Kostka matrix is the s-to-m transition (change of basis) matrix, which turns a linear combination of Schur functions into a linear combination of monomial symmetric functions by matrix multiplication.

Theorem 80. *The set $\{s_\lambda : \lambda \vdash n\}$ is a basis for Λ_n.*

Proof. If $\lambda < \mu$ in the reverse lexicographic order, then the first part in which λ and μ disagree is larger in λ than in μ. In this case there are no column strict tableau of shape μ and type λ. Further, $K_{\lambda,\lambda} = 1$ for all $\lambda \vdash n$. This tells us that the Kostka matrix is invertible because it is lower triangular with ones along the diagonal. Since $\{m_\lambda : \lambda \vdash n\}$ is a basis for Λ_n, so is $\{s_\lambda : \lambda \vdash n\}$. $\square$

It will turn out that $\{p_\lambda : \lambda \vdash n\}$, $\{h_\lambda : \lambda \vdash n\}$ and $\{e_\lambda : \lambda \vdash n\}$ are also bases is also a basis for Λ_n. We have now defined the five standard bases for Λ_n: the power, homogeneous, elementary, Schur, and monomial symmetric functions.

3.2 Sums involving the power symmetric functions

Example 81. If $\{p_\lambda : \lambda \vdash n\}$ is a basis, then so is $\{p_\lambda / z_\lambda : \lambda \vdash n\}$. This means any symmetric function can be expressed as a linear combination of p_λ's. For example, doing all the gory calculations (on a computer; the computer algebra system Sage has the ability to do symmetric function calculations) one can verify that

$$\begin{aligned}
h_{(4)} &= 1\frac{p_{(4)}}{z_{(4)}} + 1\frac{p_{(3,1)}}{z_{(3,1)}} + 1\frac{p_{(2^2)}}{z_{(2^2)}} + 1\frac{p_{(2,1^2)}}{z_{(2,1^2)}} + 1\frac{p_{(1^4)}}{z_{(1^4)}} \\
h_{(3,1)} &= 1\frac{p_{(3,1)}}{z_{(3,1)}} + 2\frac{p_{(2,1^2)}}{z_{(2,1^2)}} + 4\frac{p_{(1^4)}}{z_{(2,1^2)}} \\
h_{(2,2)} &= 2\frac{p_{(2^2)}}{z_{(2^2)}} + 2\frac{p_{(2,1^2)}}{z_{(2,1^2)}} + 6\frac{p_{(1^4)}}{z_{(1^4)}} \\
h_{(2,1,1)} &= 2\frac{p_{(2,1^2)}}{z_{(2,1^2)}} + 12\frac{p_{(1^4)}}{z_{(1^4)}} \\
h_{(1^4)} &= 24\frac{p_{(1^4)}}{z_{(1^4)}}
\end{aligned}$$

Phrasing these relationships as a matrix multiplication, the h-to-(p/z) transition matrix when $n = 4$ is

$$\begin{array}{c c}
 & \begin{array}{ccccc} (4) & (3,1) & (2^2) & (2,1^2) & (1^4) \end{array} \\
\begin{array}{c} (4) \\ (3,1) \\ (2^2) \\ (2,1^2) \\ (1^4) \end{array} & \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 2 & 4 \\ 0 & 0 & 2 & 2 & 6 \\ 0 & 0 & 0 & 2 & 12 \\ 0 & 0 & 0 & 0 & 24 \end{bmatrix}.
\end{array}$$

This matrix is the same as the matrix which gives the value of the character ψ^μ on the conjugacy class C_λ as seen in Example 57. This is the content of our next theorem.

Theorem 82. *We have* $h_\lambda = \sum_{\mu \vdash n} \psi^\lambda_\mu \frac{p_\mu}{z_\mu}$.

Proof. As seen in Section 2.1, ψ^λ_μ is equal to the number of tabloids T of shape λ such that $\sigma T = T$ for some $\sigma \in C_\mu$. Using this and Theorem 35, we will prove the theorem by proving the identity

$$\begin{aligned}
n! h_\lambda &= \sum_{\mu \vdash n} |C_\mu| (\text{\# of tabloids } T \text{ of shape } \lambda \text{ with } \sigma T = T \text{ for some } \sigma \in C_\mu) p_\mu \\
&= \sum_{\sigma \in S_n \text{ has cycle type } \mu} (\text{\# of tabloids } T \text{ of shape } \lambda \text{ with } \sigma T = T) p_{\mu(\sigma)} \qquad (3)
\end{aligned}$$

where $\mu(\sigma)$ denotes the cycle type of σ.

Count the right hand side of (3) by first selecting a permutation $\sigma \in S_n$. Write σ in cycle notation such that maximum integer in each cycle appears first. With the $p_{\mu(\sigma)}$ term, for each cycle of length k in σ, assign a sequence of cells of length k that all contain the same integer c. Acknowledging the role of sums of power symmetric functions in Pólya theory, we will call this c the color of the cycle.

Lastly, select a tabloid T of shape λ such that $\sigma T = T$. The choice of σ, the colors c for each cycle, and the tabloid T accounts for all terms on the right hand side of (3).

Arrange the colored cycles of σ such that smaller colors appear first and, if two cycles have the same color, write the cycle with the smaller maximum element first. Since $\sigma T = T$, the integers in each cycle of σ must appear in a single row of T. Place the cycles of σ into the rows of the Young diagram of λ to indicate which row of T contains the integers in each cycle. Let $\mathcal{U}$ be the set of objects created in this way.

For example, if σ is the permutation

$$(7\ 6)_1\ (10)_1\ (12\ 3\ 9)_1\ (1)_2\ (5\ 2)_2\ (13\ 4\ 11)_2\ (14\ 8)_3$$

where the color of each cycle is denoted by the subscript, then one $U \in \mathcal{U}$ is

$$\begin{array}{|c|c|c|c|c|c|c|}
\cline{1-3}
(13 & 4 & 11)_2 \\
\cline{1-4}
(7 & 6)_1 & (5 & 2)_2 \\
\hline
(10)_1 & (12 & 3 & 9)_1 & (1)_2 & (14 & 8)_3 \\
\hline
\end{array}\ .$$

If we define the weight of $U \in \mathcal{U}$ to be

$$w(U) = \prod_{i=1}^{n} x_{(\text{the color of the cycle containing } i \text{ in } U)}\,,$$

then by construction (3) is equal to $\sum_{U \in \mathcal{U}} w(U)$.

By listing the integers in a given $U \in \mathcal{U}$ (without parentheses) reading left to right and then bottom to top and then recording the color of each cell in a tableau of shape λ, each $U \in \mathcal{U}$ is in a natural 1–1 correspondence with a pair of the form (τ, T') where $\tau \in S_n$ and T' is a tableau of shape λ with nondecreasing rows. For example, the above U corresponds to (τ, T') where

$$\tau = 10\ 12\ 3\ 9\ 1\ 14\ 8\ 7\ 6\ 5\ 2\ 13\ 4\ 11$$

and the row nondecreasing tableau

$$T' = \begin{array}{|c|c|c|c|c|c|c|}
\cline{1-3}
2 & 2 & 2 \\
\cline{1-4}
1 & 1 & 2 & 2 \\
\hline
1 & 1 & 1 & 1 & 2 & 3 & 3 \\
\hline
\end{array}\ .$$

These pairs are counted by $n!h_\mu$, the left hand side of (3), as needed.

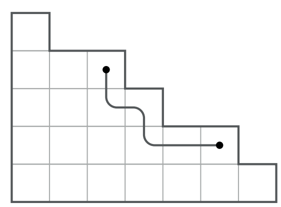
This correspondence between $U \in \mathcal{U}$ and the pairs (τ, T') is a bijection because the placement of the parenthesis in σ can be reconstructed. Indeed, after τ is placed into the Young diagram of shape λ, there is a unique way to insert pairs of parenthesis into each row so that the maximum element in each cycle to appear first and cycles are sorted in increasing order according to maximum element. This unique way is to locate the maximum integer in τ with color c, to place that integer and everything to it's right that also has color c in one cycle, and to iterate. $\square$

3.3 Rim hook tableaux

Definition. *A rim hook of length k is a sequence of k connected cells in the Young diagram of an integer partition which begins in a cell on the northeast boundary and travels along the northeast edge such that its removal leaves the Young diagram of a smaller integer partition. The sign of a rim hook ρ, denoted sign(ρ), is*

$$(-1)^{(\text{the number of rows spanned by } \rho)-1}.$$

Example 83. A rim hook of length 6 with sign $(-1)^{3-1} = +1$ is shown below:



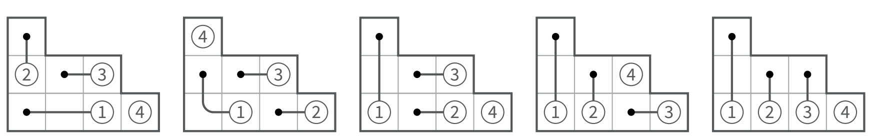
Definition. *Let μ be an integer partition of n and $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ be a composition of n (a composition of n is an integer partition where the parts need not be listed in nonincreasing order).*

A rim hook tableau of shape μ and content λ is a filling of the cells of the Young diagram of μ with rim hooks of lengths $\lambda_1, \ldots, \lambda_\ell$ labeled with $1, \ldots, \ell$ such that the removal of the last i rim hooks leaves the Young diagram of a smaller integer partition for all i.

The sign of a rim hook tableau T is a product of all of the signs of the rim hooks in T and we let

$$\chi^\mu_\lambda = \sum_{T \text{ is a rim hook tableau of shape } \mu \text{ and content } \lambda} \operatorname{sign} T.$$

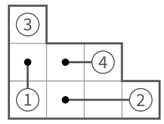
Example 84. All rim hook tableaux of shape $(4, 3, 1)$ and content $(3, 2, 2, 1)$ are



.

The signs are -1, -1, $+1$, -1, and $+1$, and so $\chi^{(4,3,1)}_{(3,2,2,1)} = -1$.

Example 85. We have $\chi^{(4,3,1)}_{(2,3,1,2)} = -1$ because

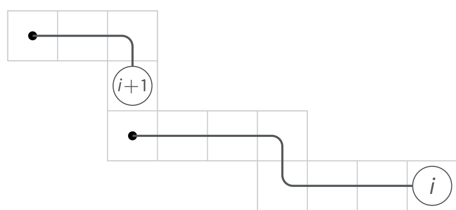


is the only rim hook tableau of shape $(4, 3, 1)$ and content $(2, 3, 1, 2)$.

Theorem 86. *If λ is an integer partition and μ and ν are compositions of n which are rearrangements of one another, then $\chi^\lambda_\mu = \chi^\lambda_\nu$.*

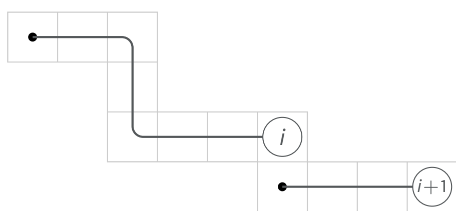
Proof. The result will follow if we can show that the order of two consecutive rim hooks can be interchanged. Then, by interchanging consecutive parts repeatedly, the composition μ can be turned into any rearrangement ν of μ. Therefore we take μ to be the composition $(\mu_1, \dots, \mu_i, \mu_{i+1}, \dots, \mu_\ell)$ and take ν to be the composition μ with μ_i and μ_{i+1} interchanged. We will prove that $\chi^\lambda_\mu = \chi^\lambda_\nu$ by first defining a sign reversing involution I on the set of rim hook tableaux of shape λ and content μ.

Let T be a rim hook tableaux of shape λ and content μ and let C be the cells in T that are occupied by rim hooks i and $i+1$. If C is itself a single rim hook, then let T' be the rim hook tableau found by switching the positions of rim hooks i and $i+1$ in T within C. For example, if rim hooks i and $i+1$ appear in T like this:



,

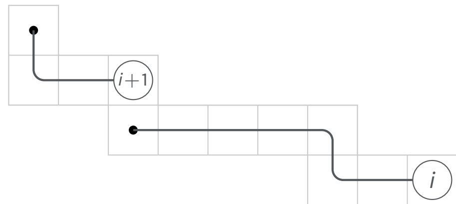
then i and $i+1$ appear in T' like this:



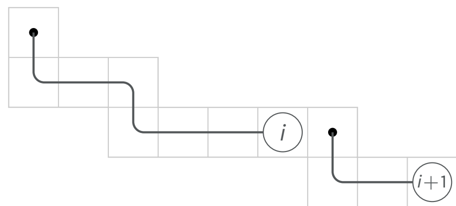
.

Rim hook $i+1$ lies outside (meaning north and/or east) of the rim hook i in T. If the rim hook $i+1$ still lies outside of rim hook i in T', then set $I(T) = T'$. If this does not happen (as depicted above) or if C is not a single rim hook, set $I(T) = T$.

Suppose $I(T) \neq T$ and suppose that rim hook $i+1$ appears before rim hook $i+1$ when reading left to right in T. Since rim hook $i+1$ lies outside of rim hook i in T, the last cell (the most south and east cell) of rim hook $i+1$ must appear below the first cell (the most north and west cell) of rim hook i in T. This situation is depicted here:



Rim hook $i+1$ also appears outside of rim hook i in $I(T)$, which means that the last cell of rim hook i must appear to the left of the first cell in rim hook $i+1$ in $I(T)$. This situation in $I(T)$ is depicted here:



This shows that $\operatorname{sign} T \neq \operatorname{sign} I(T)$ because the gap between the rim hooks i and $i+1$ in T spans two rows while the gap between the rim hooks i and $i+1$ in $I(T)$ does not. Therefore I is a sign reversing involution and

$$\chi^\lambda_\mu = \sum_{\substack{T \text{ is a rim hook tableau of shape } \lambda \\ \text{and content } \mu \text{ such that } I(T) = T}} \operatorname{sign} T.$$

By applying the same involution I to rim hooks i and $i+1$ to rim hook tableau of shape λ but with content ν, we also have

$$\chi^\lambda_\nu = \sum_{\substack{T \text{ is a rim hook tableau of shape } \lambda \\ \text{and content } \nu \text{ such that } I(T) = T}} \operatorname{sign} T.$$

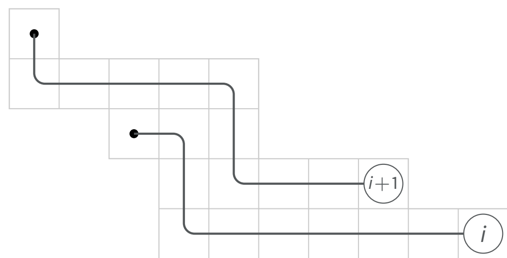
Therefore to complete the proof we will define a sign preserving bijection B from the set of rim hook tableaux T with content μ with $I(T) = T$ to the set of rim hook tableaux T' with content ν with $I(T') = T'$. This bijection will be defined on a case by case basis.

Case 1: Rim hooks i and $i+1$ do not share a border in T. Since either i can be considered to lie outside of $i+1$ or vice versa, T can be considered to have content μ or ν, and so we define $B(T) = T$.

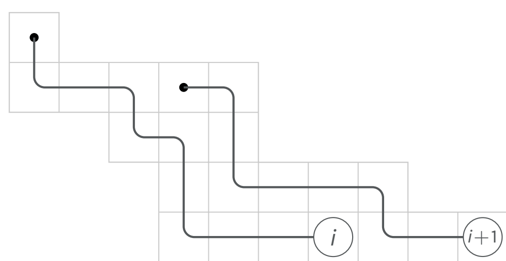
Case 2: The cells C form a single rim hook. Define $B(T)$ to be the tableau found by switching the labels of i and $i+1$ in T' where T' is the rim hook tableau described in the definition of the involution I. Since $I(T) = T$, rim hook $i+1$ lies outside of i in $B(T)$, so $B(T)$ is indeed a rim hook tableau with content ν. It follows that $I(B(T)) = B(T)$ and, in this case, that B is a bijection.

The rim hook $i+1$ lies on the outside of i in T and so the gap between these rim hooks spans two rows in T. The rim hook $i+1$ still lies outside of i in $B(T)$ and so the gap between these rim hooks also spans two rows. Therefore $\operatorname{sign} T = \operatorname{sign} B(T)$ in this case.

Case 3: Not Case 1 *or* 2. Rim hooks i and $i+1$ share a border of length 2 or longer in T with rim hook $i+1$ snaking along the outside border of i, as shown here:



Define $B(T)$ to be the rim hook tableau T created by drawing the rim hook i to snake along the outside cells C instead of rim hook $i+1$ and then switching the i and $i+1$ labels. For example, the image of the above pair of rim hooks is shown here:



Given $B(T)$, we can easily reconstruct T, showing that B is a bijection. The sign of the rim hook tableau changes exactly twice when turning T into $B(T)$: once at the first cell where rim hooks i and $i+1$ touch each other and once at the last cell where i and $i+1$ touch. Therefore the sign of T and $B(T)$ are the same.

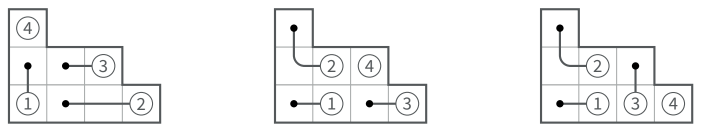
This completes our description of the sign reversing involution I and the bijection B, finishing the proof. $\square$

Example 87. This example illustrates the involution I and the bijection B in the proof of Theorem 86. Suppose we wish to show

$$\chi^{(4,3,1)}_{(3,2,2,1)} = \chi^{(4,3,1)}_{(2,3,1,2)}$$

by starting with a rim hook tableaux with content $(3,2,2,1)$, switching the order of the rim hooks of length 3 and 2, and then switching the order of the rim hooks of length 2 and 1.

In switching the rim hooks of length 3 and 2, the involution I pairs the first and third rim hook tableaux as displayed in Example 84. The bijection B sends the second, fourth, and fifth rim hook tableaux in Example 84 to



Then, to switch the rim hooks of length 2 and 1, the involution I pairs the last two of the above rim hook tableaux. Finally, the bijection B sends the remaining rim hook tableau to the rim hook tableau in Example 85.

Example 88. The matrix $[\chi^{\mu}_{\lambda}]_{\mu,\lambda\vdash n}$ (listing the shapes μ along the rows and the content λ along the columns) is

$$\begin{array}{c} \\ (4) \\ (3,1) \\ (2^2) \\ (2,1^2) \\ (1^4) \end{array} \begin{array}{c} \begin{array}{ccccc} (4) & (3,1) & (2^2) & (2,1^2) & (1^4) \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & 0 & -1 & 1 & 3 \\ 0 & -1 & 2 & 0 & 2 \\ 1 & 0 & -1 & -1 & 3 \\ -1 & 1 & 1 & -1 & 1 \end{bmatrix} \end{array}.$$

For example, the $\mu = (2^2)$ row and $\lambda = (2,1^2)$ column entry is 0 because these two possible rim hook tableaux of shape μ and content λ have opposite signs:

In what may be a surprise, this matrix is the character table for S_4! (See Example 53.)

Example 89. If $\{p_\lambda : \lambda \vdash n\}$ is a basis, then so is $\{p_\lambda/z_\lambda : \lambda \vdash n\}$, and then any symmetric function can be expressed as a linear combination of p_λ's. For example, doing all the gory calculations (on a computer; the computer algebra system Sage has the ability to do symmetric function calculations) one can verify that

$$\begin{aligned}
s_{(4)} &= 1\frac{p_{(4)}}{z_{(4)}} + 1\frac{p_{(3,1)}}{z_{(3,1)}} + 1\frac{p_{(2^2)}}{z_{(2^2)}} + 1\frac{p_{(2,1^2)}}{z_{(2,1^2)}} + 1\frac{p_{(1^4)}}{z_{(1^4)}} \\
s_{(3,1)} &= -1\frac{p_{(4)}}{z_{(4)}} + 0\frac{p_{(3,1)}}{z_{(3,1)}} - 1\frac{p_{(2^2)}}{z_{(2^2)}} + 1\frac{p_{(2,1^2)}}{z_{(2,1^2)}} + 3\frac{p_{(1^4)}}{z_{(1^4)}} \\
s_{(2,2)} &= 0\frac{p_{(4)}}{z_{(4)}} + -1\frac{p_{(3,1)}}{z_{(3,1)}} + 2\frac{p_{(2^2)}}{z_{(2^2)}} + 0\frac{p_{(2,1^2)}}{z_{(2,1^2)}} + 2\frac{p_{(1^4)}}{z_{(1^4)}} \\
s_{(2,1,1)} &= 1\frac{p_{(4)}}{z_{(4)}} + 0\frac{p_{(3,1)}}{z_{(3,1)}} - 1\frac{p_{(2^2)}}{z_{(2^2)}} - 1\frac{p_{(2,1^2)}}{z_{(2,1^2)}} + 3\frac{p_{(1^4)}}{z_{(1^4)}} \\
s_{(1^4)} &= -1\frac{p_{(4)}}{z_{(4)}} + 1\frac{p_{(3,1)}}{z_{(3,1)}} + 1\frac{p_{(2^2)}}{z_{(2^2)}} - 1\frac{p_{(2,1^2)}}{z_{(2,1^2)}} + 1\frac{p_{(1^4)}}{z_{(1^4)}}
\end{aligned}$$

Phrasing these relationships as a matrix multiplication, the s-to-(p/z) transition matrix when $n = 4$ is the same matrix as found in Example 88. So, at least in the case $n = 4$, we see that this same matrix is three things at once:

1. the character table for S_n,
2. the matrix $\left[\chi^\lambda_\mu\right]_{\lambda,\mu\vdash n}$ found by considering rim hook tableaux, and
3. the s-to-(p/z) transition matrix.

The next theorem proves that items 2 and 3 on this list are the same for any n. Later we will also prove that item 1 on this list is the same as items 2 and 3.

Theorem 90. *We have*

$$n!s_\lambda = \sum_{\sigma\in S_n} \chi^\lambda_{\mu(\sigma)} p_{\mu(\sigma)} \tag{4}$$

for all $\lambda \vdash n$.

Proof. Looking at the right hand side of (4), select a colored permutation with colored cycles in the same way as in Theorem 82, accounting for the choice of $\sigma \in S_n$ and the $p_{\mu(\sigma)}$. Arrange the colored cycles in the same way as in Theorem 82 with smaller colors first and then sorted by smallest maximum element. One such σ is

$$(8\ 3\ 1)_1\ (12\ 9)_1\ (7\ 2)_2\ (13\ 4\ 11\ 5\ 10\ 6)_2$$

where the color of each cycle is its subscript.

With the $\chi^\lambda_{\mu(\sigma)}$ term, select a rim hook tableau of shape λ and content ν where ν is the composition giving the lengths of the colored cycles in σ (using Theorem 86). Write down each colored cycle on top of the corresponding rim hook in the rim hook tableau.

For example, if $\lambda = (5, 5, 3)$ and σ is as displayed above, then one object is

$$\begin{array}{|c|c|c|cc}\cline{1-3} (13 & 4 & 11 & & \\ \hline (7 & 2)_2 & 5 & 10 & \multicolumn{1}{c|}{6)_2} \\ \hline (8 & 3 & 1)_1 & (12 & \multicolumn{1}{c|}{9)_1} \\ \hline\end{array}\ .$$

Let $\mathcal{T}$ be the set of all objects created this way. The weight of $T \in \mathcal{T}$ is

$$w(T) = \prod_{i=1}^{n} x_{(\text{the color of the cycle containing } i \text{ in } T)}$$

and the sign is the sign of the underlying rim hook tableau. For example, the weight of the $T \in \mathcal{T}$ shown above is $x_1^5 x_2^8$ and the sign is -1. The right side of (4) is equal to $\sum_{T\in\mathcal{T}} \operatorname{sign}(T)\, w(T)$ by construction.

We now define a sign reversing and weight preserving involution I on $\mathcal{T}$. Scan the cells of $T \in \mathcal{T}$ from left to right and from bottom to top, looking for the first occurrence of either

Case A: a cell x in the same rim hook as the cell immediately above x, or

Case B: a cell colored c which is immediately below the terminal cell in a c colored rim hook.

The example T displayed above falls into Case A because the cell containing 5 cell and the cell above the 5 are both contained in the same rim hook.

Suppose we are in Case A and let x be the cell which is in the same rim hook as the cell above x. Define $I(T)$ to be the element in $\mathcal{T}$ created by following these instructions:

1. Cut the rim hook containing x at the down step between x and the cell above x. This ends the rim hook above x at the cell above x.

2. Consider the list of integers in the same row as x which are the same color as x. Erase all parentheses currently found in this list and then reinsert parentheses, thereby creating cycles, in the unique manner that forces the maximum element in each cycle to appear first and forces cycles to be sorted in increasing order according to maximum element. (This is the same idea as found in the last paragraph of the proof of Theorem 82.)

For example, if T is the object displayed above, then $I(T)$ is

$(13$	4	$11)_2$		
$(7$	2	$5)_2$	$(10$	$6)_2$
$(8$	3	$1)_1$	$(12$	$9)_1$

.

If T falls into Case A, then $I(T)$ will fall into Case B because the cell x that was in the same rim hook as the cell above x is now a cell colored c which is immediately below the terminal cell in a c colored rim hook.

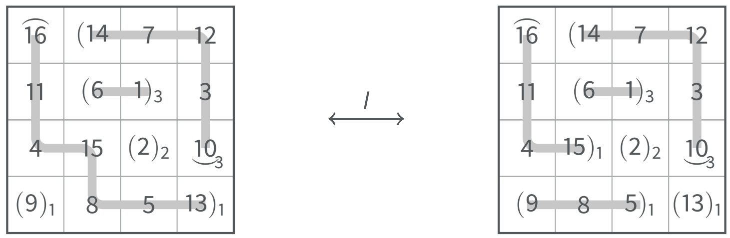
Now suppose T falls into Case B and let x be the cell colored c which is immediately below the terminal cell in a c colored rim hook. Suppose the maximum element in this rim hook below x is m. Define $I(T)$ to be the element in $\mathcal{T}$ created by following these instructions:

1. If x is in a cycle of length 2 or greater, chop x off of the end of its cycle.

2. Extend the rim hook to include x and to include any cycles to the right of x which have a maximum element smaller than m.

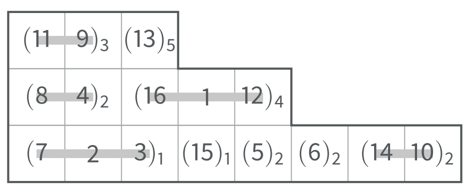
If $T \in \mathcal{T}$ falls into Case B, then $I(T)$ will fall into Case A because the cell x that was the same color as the rim hook below x is now in the same rim hook as the cell above x. Indeed, the function I on a Case B object is the inverse function to I on a Case A object.

If $T \in \mathcal{T}$ does not fall into Case A or Case B, then we define $I(T) = T$. This completes our description of the involution I. As a second example, I pairs these two elements in $\mathcal{T}$:



If T is not a fixed point, then I changes the sign of T because I removes exactly one down step in Case A and introduces exactly one down step into a rim hook in Case B. The involution is weight preserving.

Fixed points under I cannot have any rim hooks which span two or more rows and thus all fixed points have sign $+1$. Fixed points cannot have a cell with color c below another cell with color c and our ordering of the rim hooks guarantees that larger colors appear in rim hooks outside of smaller colors. One example of a fixed point is here:



By listing the integers in a fixed point (without parentheses) and recording the color of each cell in a column strict tableau of shape μ, a fixed point under I is in a natural 1–1 correspondence with a pair of the form (τ, T') where $\tau \in S_n$ and T' is a column strict tableau of shape μ. For example, the above fixed point corresponds to (τ, T') where σ is

$$\sigma = 11\ 9\ 13\ 8\ 4\ 16\ 1\ 12\ 7\ 2\ 3\ 15\ 5\ 6\ 14\ 10$$

and the column strict tableau

$$T' = \begin{array}{|c|c|c|c|c|c|c|c|} \hline 3 & 3 & 5 \\ \cline{1-5} 2 & 2 & 4 & 4 & 4 \\ \hline 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 \\ \hline \end{array} \; .$$

These pairs are counted by $n!s_\mu$, the left hand side of 4, as needed. □

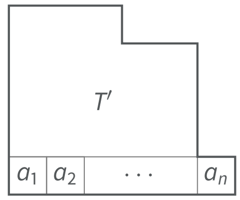
4 The RSK algorithm

4.1 Bumping

Definition. *Let T be a column strict tableau and j an integer. The row insertion of j into T, denoted $T \leftarrow j$, is found by following these rules:*

Insertion rule 0. If T is the empty tableau, then $T \leftarrow j = \boxed{j}$.

If T is not empty, assume T is of the form shown here:



That is, the first row of T is $a_1 \leq \cdots \leq a_n$ and T' is the column strict tableau found by removing the first row of T.

Insertion rule 1. If $a_n \leq j$, then $T \leftarrow j$ is T is the tableau with $\boxed{j}$ appended to the bottom row of T.

Insertion rule 2. If $j < a_n$, then let a_k be the leftmost entry in bottom row of T that is larger than j. Replace a_k with j and insert a_k into T'. In this case we say that j bumps a_k.

Example 91. To row insert a 2 into

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

we replace the 3 in the bottom row with a 2, the 5 in the second row with a 3, and then place a new cell containing a 5 in the third row. Graphically we have

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

Shading the path of the replaced cells in this row insertion process gives

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

This path is called a bumping path. This path must start in the bottom row and move up and weakly to the left and the integers in this path are strictly increasing when read from bottom to top.

Theorem 92. *If T is a column strict tableau and j is a positive integer, then $T \leftarrow j$ is a column strict tableau and the bumping path for the insertion of j in T moves weakly to the left as one proceeds from bottom to top. Moreover, if $sh(T)$ is shape of T, then $sh(T) \subseteq sh(T \leftarrow j)$.*

Proof. The theorem is true if applying insertion rule 0 or insertion rule 1. We therefore assume that insertion rule 2 was applied.

Suppose j bumps a_k in the bottom row of T. The first row of $T \leftarrow j$ is

$$a_1 \cdots a_{k-1}\, j\, a_{k+1} \cdots a_n.$$

By our choice of of a_k, we must have $a_{k-1} \leq j < a_k \leq a_{k+1} \leq \cdots \leq a_n$, meaning that the first row of $T \leftarrow j$ is weakly increasing.

By induction, $T' \leftarrow a_k$ is a column strict tableaux. Thus to show that $T \leftarrow j$ is a column strict tableau, we need only show that $T \leftarrow j$ is strictly increasing in columns in the first two rows. Suppose that $b_1 \leq \cdots \leq b_s$ is the first row of T'. There are two cases.

Case 1. Suppose $s \geq k$. In this case we know $b_k > a_k$, so $T' \leftarrow a_k$ either bumps b_k or it must bump some b_s with $s \leq k$. If b_k is bumped, then the first two cells of the k-th column contains j in the first row and a_k in the second row. Since $j < a_k$, $T \leftarrow j$ satisfies the column strict condition in k-th column as the elements in the first two rows of the remaining columns are the same as in T. Thus $T \leftarrow j$ will be column strict.

If a_k bumps b_s for $s < k$, then we know that in column k the first two elements are j and b_k, but $j < a_k < b_k$ so that we satisfy the column strict condition in column k. In column s, the first two elements are a_s and a_k, but we know that $a_s \leq a_{k-1} \leq j < a_k$ so that we satisfy the column strict condition in column s. In the remaining columns, the first two elements are the same as in T. Thus $T \leftarrow j$ will be column strict.

Case 2. Suppose $s < k$. In this case either a_k is at the end of the first row in T' or it must bump some b_s with $s \leq k$. If a_k is at the end of the first row in T', then either a_k is on top of j if $s = k-1$ (in which case $j < a_k$) or a_k is on top of some a_{s+1} where $s+1 \leq k-1$, in which case $a_{s+1} \leq a_{k-1} \leq j < a_k$. Regardless, $T \leftarrow j$ is column strict in the first two rows of the column that contains a_k. The first two elements in the remaining columns are the same as in T and so $T \leftarrow j$ is column strict.

If a_k bumps b_s for $s < k$, then we know column k has only one element. In column s the first two elements are a_s and a_k, but we know that $a_s \leq a_{k-1} \leq j < a_k$ so that we satisfy the column strict condition in column s. The first two elements are the same as in T in the remaining columns. Thus $T \leftarrow j$ will be column strict.

Lastly, the bumping path moves weakly to the left in the first two rows. By induction, the bumping moves weakly to the left in $T' \leftarrow a_k$. Thus the entire bumping path moves weakly to the left. $\square$

If we are given $T \leftarrow j$ and the location of the cell c_1 in $sh(T \leftarrow j)$ but not $sh(T)$, then we can find both T and j. To reconstruct T and find j, follow these steps to reverse the row insertion process:

1. If the final cell inserted into T is on the bottom row, then this cell contains j and the removal of this cell leaves T.

2. If k is the integer in the final cell inserted into T, then find the rightmost entry in the row below k that is smaller than k, say ℓ. Remove the cell with k, replace ℓ with k, and repeat the process, moving down one row with each step.

Given a word $w_1 w_2 \cdots w_n$, let

$$T \leftarrow w_1 \cdots w_n = (\cdots((T \leftarrow w_1) \leftarrow w_2)\cdots) \leftarrow w_n$$

so that $T \leftarrow w_1 \cdots w_n$ is the result of successively inserting $w_1, w_2, \ldots, w_n$ into T. In such a situation, we denote $T_0 = T$ and $T_i = T \leftarrow w_1 \ldots w_i$ for $i = 1, \cdots, n$. It follows from Theorem 92 that

$$sh(T_0) \subset sh(T_1) \subset sh(T_2) \subset \cdots \subset sh(T_n)$$

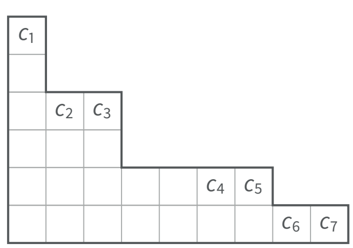
where the notation $\mu \subseteq \lambda$ means that the Young diagram for μ fits inside that of λ. We also let c_i to be the cell in $sh(T_i)$ but not $sh(T_{i-1})$ for $i = 1, \ldots, n$.

Theorem 93. *If $w = w_1 \cdots w_n$ is a word of length $n \geq 2$ with letters in $\{1, 2, \ldots\}$, then*

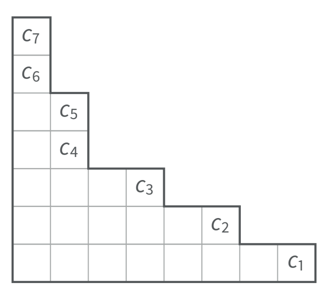
1. *if $w_1 \leq \cdots \leq w_n$, then c_{i+1} is strictly to the right and weakly below c_i for all i, and*

2. *if $w_1 > \cdots > w_n$, then c_{i+1} is strictly above and weakly to the left of c_i for all i.*

These conditions are pictured below:



The case of $w_1 \leq w_2 \leq \cdots \leq w_7$.



The case of $w_1 > w_2 > \cdots > w_7$.

Proof. By induction it is enough to prove either statement by considering only the case $n = 2$.

We prove the first statement by induction on the size of T. If c_1 is in row 1 (which would be the case only if insertion rule 1 were used in $T \leftarrow w_1$), then w_1 is at the end of the row in T_1. Since $w_1 \leq w_2$, the integer w_2 is placed at the end of the first row of T_1 in the insertion $T_1 \leftarrow w_2$. In this case, c_2 is clearly strictly to the right and weakly below c_1.

Suppose the bottom row of T contains $a_1 \leq \cdots \leq a_j$ and w_1 bumps a_k in $T \leftarrow w_1$. By our choice of a_k, we have $a_{k-1} \leq w_1 < a_k$ and w_1 is in the k-th cell of row 1 in T_1. This means that in the insertion $T_1 \leftarrow w_2$, either (i) w_2 is placed in the end of row 1 or (ii) w_2 must bump a_s where $s > k$. In case (i), our result follows since the cell c_1 is the new cell created by $T' \leftarrow a_k$ and we know the that bumping paths move weakly to the left by Theorem 93. In case (ii), the result follows by induction since c_1 and c_2 are the cells created by the insertion $T' \leftarrow a_k a_s$ and $a_k \leq a_s$. This completes the proof of the first statement.

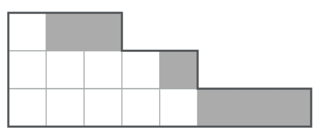
The proof of the second statement is similar and is left to the reader. $\square$

4.2 The Pieri rules

The Pieri rules, found in Theorem 95 and Exercise 29, give a nice description of how to expand the products $h_n s_\mu$ and $e_n s_\mu$ into a sum of Schur symmetric functions.

Definition. *For integer partitions μ and λ, let λ/μ be the cells in the Young diagram for λ but not those in μ. This is type of object is called a skew shape. We say that skew shape λ/μ is a skew row if λ/μ has no two cells in the same column and λ/μ is a skew column if no two cells of λ/μ lie in the same row.*

Example 94. If $\lambda = (8,5,3)$ and $\mu = (5,4,1)$, then the skew shape λ/μ is shaded in the diagram below:



In this example, λ/μ is a skew row.

Theorem 95 (The Pieri rules)**.** *For any partition μ and for all $n \geq 1$,*

$$h_n s_\mu = \sum_{\substack{\lambda/\mu \text{ is a skew row} \\ \text{with } n \text{ cells}}} s_\lambda.$$

Proof. We begin with the expansion of $h_n s_\mu$. Let $\mathcal{H}_{n,\mu}$ denote set of all pairs (S,T) where S is row nondecreasing tableau of shape (n) and T is a column strict tableau of shape μ. Let $SR_{n,\mu}$ denote the set of all column strict tableaux P such that $sh(P)/\mu$ is a skew row of size n. The statement of the theorem is equivalent to

$$\sum_{(S,T)\in\mathcal{H}_{n,\mu}} w(S)w(T) = \sum_{P\in SR_{n,\mu}} w(P)$$

where $w(Q)$ is the usual weight for column strict tableaux.

We claim that the row insertion algorithm allows us to give a weight preserving bijection $\theta : \mathcal{H}_{n,\mu} \to SR_{n,\mu}$. That is, given a pair (S, T), let $a_1 \leq \cdots \leq a_n$ be the elements of S, reading from left to right. We define $\theta(S, T) = P = T \leftarrow a_1 \ldots a_n$. By the first statement in Theorem 93, we know that if $\lambda = sh(P)$, then λ/μ is skew row. Moreover, we know that in the insertion of $a_1 \ldots a_n$ into T, the new cells where created from left to right. This allows us to reverse our steps, showing that θ is one-to-one.

To show that θ is bijection, it remains to be shown that θ is surjective; that is, we must show that if P is any column strict tableau such that $sh(P)/\mu$ is skew row of size n, then there exists a sequence $a_1 \leq \cdots \leq a_n$ and a column strict tableau T of shape μ such that $P = T \leftarrow a_1 \ldots a_n$.

If P did come from an insertion of the form $T \leftarrow a_1 \cdots a_n$ where $a_1 \leq \cdots \leq a_n$, then we know that new cells were created from left to right by Theorem 93. Thus we can reverse the process by first doing the reverse row insertion on the rightmost cell of $sh(P)/\mu$. Thus $P = P_1 \leftarrow a_n$ for some column strict tableau P_1 such that $sh(P_1)/\mu$ is skew row. Next we reverse the row insertion for P_1 using the rightmost cell of $sh(P_1)/\mu$. Thus $P_1 = P_2 \leftarrow a_{n-1}$ so that $P = P_2 \leftarrow a_{n-1}a_n$. Continuing on in this way, we will obtain a sequence $a_1 \cdots a_n$ and a column strict tableau of shape μ such that $P = T \leftarrow a_1 \cdots a_n$.

For example, suppose $\mu = (4, 2, 1)$, $\lambda = (5, 4, 2, 1)$, and P is the column strict tableau shown below

$$\begin{array}{ccccc} 4 & & & & \\ 3 & 4 & & & \\ 2 & 2 & 3 & 3 & \\ 1 & 1 & 2 & 2 & 4 \end{array},$$

where we have shaded the cells in λ/μ. We illustrate the idea of reversing the row insertion process by first undoing the row insertion of the bottom 4, then undoing the row insertion of the right most 3, and so on.

$$\begin{array}{cccc} 4 & & & \\ 3 & 4 & & \\ 2 & 2 & 3 & 3 \\ 1 & 1 & 2 & 2 \end{array} \leftarrow 4, \quad \begin{array}{cccc} 4 & & & \\ 3 & 4 & & \\ 2 & 2 & 3 & \\ 1 & 1 & 2 & 3 \end{array} \leftarrow 24, \quad \begin{array}{cccc} 4 & & & \\ 3 & 4 & & \\ 2 & 2 & & \\ 1 & 1 & 3 & 3 \end{array} \leftarrow 224,$$

$$\begin{array}{cccc} 4 & & & \\ 3 & & & \\ 2 & 4 & & \\ 1 & 2 & 3 & 3 \end{array} \leftarrow 1224, \quad \begin{array}{cccc} 4 & & & \\ 3 & 4 & & \\ 2 & 2 & 3 & 3 \end{array} \leftarrow 11224.$$

The question of why $a_1, \cdots, a_n$ is weakly increasing remains. Appealing to the second statement in Theorem 93, if $a_j > a_{j+1}$ for some j, then the new cell c_{j+1} created by the insertion of a_{j+1} into $T \leftarrow a_1 \cdots a_j$ is strictly above and weakly to left of the new cell c_j created by the insertion of a_j into $T \leftarrow a_1 \cdots a_{j-1}$. But this is not what happens in our process—in our reverse process, c_{i+1} is strictly to the right and weakly below c_i. Hence there can be no such j and so $a_1, \cdots, a_n$ is weakly increasing.

This now proves that θ is a bijection, as needed. □

Theorem 96. *For all $\mu \vdash n$, we have*

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

This gives the entries of the h-to-s transition matrix.

Proof. The result follows immediately from iterating the Pieri rule for the homogeneous symmetric functions.

That is, suppose $\mu = (\mu_1, \ldots, \mu_k)$. Then $h_{\mu_1} = s_{(\mu_1)}$; we can place 1's in the cells of the shape (μ_1). Using the Pieri rule to multiply $h_{\mu_2} s_{(\mu_1)}$, we find all s_λ such that $\lambda/(\mu_1)$ is a skew row. For each such λ, place 2's in the cells of $\lambda/(\mu_1)$. It follows that $h_{\mu_2} s_{(\mu_1)}$ equals the sum over all all s_λ such that the shape λ is the shape of column strict tableau T of weight $x_1^{\mu_1} x_2^{\mu_2}$.

For each such λ, we can use the Pieri rule again to find $h_{\mu_3} s_\lambda$. We mark each new cell added to λ with a 3. Then $h_{\mu_3} h_{\mu_2} s_{(\mu_1)}$ equals the sum over all all s_δ such that the shape δ is the shape of column strict tableau T of weight $x_1^{\mu_1} x_2^{\mu_2} x_3^{\mu_3}$. Continuing in this way proves the theorem. □

Example 97. Below we display the h-to-s transition matrix for all integer partitions $\lambda \vdash 4$:

$$\begin{array}{cc} & \begin{array}{ccccc} h_{(4)} & h_{(3,1)} & h_{(2^2)} & h_{(2,1^2)} & h_{(1^4)} \end{array} \\ \begin{array}{c} s_{(4)} \\ s_{(3,1)} \\ s_{(2,2)} \\ s_{(2,1^2)} \\ s_{(1^4)} \end{array} & \left[\begin{array}{ccccc} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] \end{array}.$$

For example, $h_{(2,1^2)}$ column in the above matrix is determined by these column strict tableau:

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

It is no accident that this exact same matrix appears in Example 57.

4.3 The full RSK algorithm

We now present the RSK algorithm. It is named after Gilbert de Beauregard Robinson, who first described an algorithm equivalent to bumping, Craige Schensted who described the algorithm for permutations, and Donald Knuth, who extended the algorithm from permutations to matrices.

Algorithm. *The input is a nonzero nonnegative integer valued matrix A.*

1. *Begin with P and Q as empty column strict tableaux.*
2. *Let (i, j) be the topmost and then the leftmost nonzero entry in A.*
3. *Change P to $P \leftarrow j$, thereby adding one cell to P. Add a cell containing i to Q in the same position as the cell that was added to P.*
4. *Change A by subtracting* 1 *from the (i, j) entry.*
5. *If A is the zero matrix, stop. Otherwise, go back to step 2.*

The output is the pair (P, Q).

Example 98. Consider applying RSK to $A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & 1 \end{bmatrix}$. Starting at the top left entry of A and moving across each row, the first nonzero entry is at $(1, 1)$. After initializing P and Q to be empty, steps 3 and 4 give

$$P = \begin{array}{|c|}\hline 1 \\ \hline\end{array} \qquad Q = \begin{array}{|c|}\hline 1 \\ \hline\end{array} \qquad A = \begin{bmatrix} 0 & 2 & 0 \\ 3 & 0 & 1 \end{bmatrix}.$$

Iterating until A is the zero matrix, we have

$$P = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline\end{array} \qquad Q = \begin{array}{|c|c|}\hline 1 & 1 \\ \hline\end{array} \qquad A = \begin{bmatrix} 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix},$$

$$P = \begin{array}{|c|c|c|}\hline 1 & 2 & 2 \\ \hline\end{array} \qquad Q = \begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline\end{array} \qquad A = \begin{bmatrix} 0 & 0 & 0 \\ 3 & 0 & 1 \end{bmatrix},$$

$$P = \begin{array}{|c|cc}\cline{1-1} 2 & & \\ \hline 1 & \multicolumn{1}{|c|}{1} & \multicolumn{1}{c|}{2} \\ \hline\end{array} \qquad Q = \begin{array}{|c|cc}\cline{1-1} 2 & & \\ \hline 1 & \multicolumn{1}{|c|}{1} & \multicolumn{1}{c|}{1} \\ \hline\end{array} \qquad A = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 1 \end{bmatrix},$$

$$P = \begin{array}{|c|c|c}\cline{1-2} 2 & 2 & \\ \hline 1 & 1 & \multicolumn{1}{c|}{1} \\ \hline\end{array} \qquad Q = \begin{array}{|c|c|c}\cline{1-2} 2 & 2 & \\ \hline 1 & 1 & \multicolumn{1}{c|}{1} \\ \hline\end{array} \qquad A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix},$$

$$P = \begin{array}{|c|c|cc}\hline 2 & 2 & \multicolumn{2}{c}{} \\ \hline 1 & 1 & \multicolumn{1}{c|}{1} & \multicolumn{1}{c|}{1} \\ \hline \end{array} \qquad Q = \begin{array}{|c|c|cc}\hline 2 & 2 & \multicolumn{2}{c}{} \\ \hline 1 & 1 & \multicolumn{1}{c|}{1} & \multicolumn{1}{c|}{2} \\ \hline \end{array} \qquad A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$P = \begin{array}{|c|c|ccc}\hline 2 & 2 & \multicolumn{3}{c}{} \\ \hline 1 & 1 & \multicolumn{1}{c|}{1} & \multicolumn{1}{c|}{1} & \multicolumn{1}{c|}{3} \\ \hline \end{array} \qquad Q = \begin{array}{|c|c|ccc}\hline 2 & 2 & \multicolumn{3}{c}{} \\ \hline 1 & 1 & \multicolumn{1}{c|}{1} & \multicolumn{1}{c|}{2} & \multicolumn{1}{c|}{2} \\ \hline \end{array} \qquad A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The output of applying RSK to A is the pair (P, Q).

An alternative way of thinking about RSK is to consider the word of pairs

$$w(A) = \begin{matrix} q_1 & q_2 & \dots & q_n \\ p_1 & p_2 & \dots & p_n \end{matrix}$$

obtained by reading the rows of the matrix A from left to right starting at the top row and ending at the bottom row where for each $a_{i,j} > 0$, we write down $a_{i,j}$ pairs of the form $\begin{matrix} i \\ j \end{matrix}$. We call $w(A)$ the bi-word of A. For example, if A is the matrix of our example, then

$$w(A) = \begin{matrix} 1 & 1 & 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 2 & 1 & 1 & 1 & 3. \end{matrix}$$

Then P is found from $\emptyset \leftarrow p_1 \cdots p_n$. The matrix P is sometimes called the insertion tableau. We then use the elements of $q_1 \dots q_n$ to record the growth of P. That is, if c_i is the new cell created by the insertion $(\emptyset \leftarrow p_1 \cdots p_{i-1}) \leftarrow p_i$, then we place q_i in cell c_i of Q. The matrix Q is sometimes called the recording tableau.

Theorem 99. *The RSK algorithm is a bijection between nonnegative integer valued matrices A and pairs of the form (P, Q) where P and Q are column strict tableau of the same shape.*

Proof. Suppose that

$$w(A) = \begin{matrix} q_1 & q_2 & \cdots & q_n \\ p_1 & p_2 & \cdots & p_n \end{matrix}$$

The shapes of P and Q are the same by construction. By Theorem 92, P is a column strict tableau. Moreover, our construction ensures Q is weakly increasing in rows and columns because at any stage, the new cell attached to Q contains a letter i which is greater than or equal to all the previous letters added to Q.

Thus to prove Q is column strict, we need only show that for any letter i, no two i's in the Q can lie in the same column. But this is an immediate consequence of the first statement in Theorem 93. That is, the elements corresponding to the i-th row of A were inserted in weakly increasing order because when we created the word of A, we read the elements form left to right. This insures that if $q_s q_{s+1} \cdots q_t$ is block of i's in $q_1 \dots q_n$, then $p_s \leq \cdots \leq p_t$. But then by the first statement in Theorem 93, new cells which were created by the insertion

$$(\emptyset \leftarrow p_1 \cdots p_{s-1}) \leftarrow p_s \cdots p_t$$

where created from left to right and form a skew row. Thus no two is in Q can be in the same column so that Q is column strict.

It follows that we can read off the the order in which the cells were created in P. That is, for any fixed i, we know by the first statement in Theorem 93, the cells were created from left to right. Thus the last cell c that was created in P corresponds to the right-most cell which contains the largest element in Q. Because we can reverse the row insertion algorithm starting at cell c, it follows that we can reconstruct A from P and Q.

At this point we know that the correspondence $A \to (P, Q)$ is an injection, so it remains to be shown that it is also a surjection. That is, we must show that if P and Q are column strict tableaux of the same shape, then there is a non-negative valued matrix A such that RSK sends A To (P, Q).

The idea is that given (P, Q), we can reverse the bumping process as if came from inserting a bi-word of a matrix A. That is, for each i in Q, we assume that the cells containing i were created from left to right. This allows us to reconstruct a bi-word

$$\begin{array}{cccc} q_1 & q_2 & \cdots & q_n \\ p_1 & p_2 & \cdots & p_n \end{array}$$

by successively reversing the bumping process in P using the cell that contains the largest and then right-most element of Q.

For example, suppose we wish to undo the bumping process for the pair (P, Q) shown below:

$$P = \begin{array}{|c|c|c|c|} \cline{1-2} 2 & 4 & \multicolumn{2}{c}{} \\ \hline 1 & 1 & 3 & 4 \\ \hline \end{array} \qquad Q = \begin{array}{|c|c|c|c|} \cline{1-2} 2 & 2 & \multicolumn{2}{c}{} \\ \hline 1 & 1 & 2 & 2 \\ \hline \end{array}$$

We begin by locating the largest integer in P. If there are ties, select the rightmost. In this example, the rightmost 4 in P is chosen. Then we undo row insertion in P, remove the cell in the same position as Q, and then record this move in the bi-word. This process looks like this:

$$P = \begin{array}{|c|c|c|} \cline{1-2} 2 & 4 & \multicolumn{1}{c}{} \\ \hline 1 & 1 & 3 \\ \hline \end{array} \qquad Q = \begin{array}{|c|c|c|} \cline{1-2} 2 & 2 & \multicolumn{1}{c}{} \\ \hline 1 & 1 & 2 \\ \hline \end{array} \qquad w(A) = \begin{array}{c} 2 \\ 4 \end{array}$$

$$P = \begin{array}{|c|c|} \hline 2 & 4 \\ \hline 1 & 1 \\ \hline \end{array} \qquad Q = \begin{array}{|c|c|} \hline 2 & 2 \\ \hline 1 & 1 \\ \hline \end{array} \qquad w(A) = \begin{array}{cc} 2 & 2 \\ 3 & 4 \end{array}$$

$$P = \begin{array}{|c|c|} \cline{1-1} 2 & \multicolumn{1}{c}{} \\ \hline 1 & 4 \\ \hline \end{array} \qquad Q = \begin{array}{|c|c|} \cline{1-1} 2 & \multicolumn{1}{c}{} \\ \hline 1 & 1 \\ \hline \end{array} \qquad w(A) = \begin{array}{ccc} 2 & 2 & 2 \\ 1 & 3 & 4 \end{array}$$

$$P = \begin{array}{|c|c|} \hline 2 & 4 \\ \hline \end{array} \qquad Q = \begin{array}{|c|c|} \hline 1 & 1 \\ \hline \end{array} \qquad w(A) = \begin{array}{cccc} 2 & 2 & 2 & 2 \\ 1 & 1 & 3 & 4 \end{array}$$

$P = \boxed{2}$ $\qquad Q = \boxed{1}$ $\qquad w(A) = \begin{matrix} 1 & 2 & 2 & 2 & 2 \\ 4 & 1 & 1 & 3 & 4 \end{matrix}$

$P = \varnothing$ $\qquad Q = \varnothing$ $\qquad w(A) = \begin{matrix} 1 & 1 & 2 & 2 & 2 & 2 \\ 2 & 4 & 1 & 1 & 3 & 4 \end{matrix}$

It only remains to be shown that for each i, if $q_s q_{s+1} \cdots q_t$ is block of i's in $q_1 \cdots q_n$, then $p_s \leq \cdots \leq p_t$. Just like in the proof of Theorem 95, we can use the second statement in Theorem 93. That is, $p_j > p_{j+1}$ for some $s \leq j < j+1 \leq t$, then by Theorem 93, the cell c_{j+1} created by inserting p_{j+1} in $\emptyset \leftarrow p_1 \cdots p_j$ would be strictly above cell c_i created by p_j in $\emptyset \leftarrow p_1 \cdots p_{j-1}$. But by construction, c_{j+1} is strictly to the right and weakly below c_j so that there can be no such j.

This shows that RSK is bijection from the set of all non-negative integer valued matrices A onto the collection of pairs of column strict tableaux (P, Q) of the same shape. $\square$

Theorem 100. *For all $\alpha, \beta \vdash n$, let $\mathcal{N}_{\alpha,\beta}$ be the number of nonnegative integer valued matrices A with row sum α and column sum β. Then we have*

$$|S_\alpha|\mathcal{N}_{\alpha,\beta} = \sum_{\sigma \in S_\alpha} \text{(the number of tabloids } T \text{ of shape } \beta \text{ fixed by } \sigma\text{)}.$$

Proof. The left hand side of the identity counts pairs (σ, A) where $\sigma \in S_\alpha$ and A is a nonnegative integer valued matrix with row sum α and column sum β.

Let A' be the matrix A with the integers in A replaced with the sets in a set partition of n by labeling from top to bottom and moving left to right, replacing the number k in A with a set of k consecutive numbers. For example, if $\alpha = (3, 3, 2)$ and $\beta = (3, 2, 2, 1)$, then one matrix A is

$$A = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

For this example matrix A, the matrix A' is

$$A' = \begin{bmatrix} \{1,2,3\} & 0 & 0 & 0 \\ 0 & \{4\} & \{5,6\} & 0 \\ 0 & \{7\} & 0 & \{8\} \end{bmatrix}.$$

Create a tableau T of shape β by

1. filling the bottom row of a tableau with the integers found in the first column of A' read top to bottom and then filling the next row of with the integers in the second column of A', etc, and then

2. permuting the integers in the tableau using the permutation $\sigma \in S_\alpha$.

For example, if the permutation $\sigma = (1\,2\,3)(4\,5)$, then the tableau T corresponding to the matrix A' shown above is

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

Let τ be the permutation in S_α which makes the tableau $T' = \tau T$ a tabloid (so that the rows of T' are increasing). In the running example, $\tau = (1\,3\,2)$. With this construction we have turned the pair (σ, A) into the pair (τ, T') where T' is a tabloid of shape β which is fixed by τ.

This process is a bijection. Indeed, given a pair (τ, T') where $\tau \in S_\alpha$ and T' is a tabloid of shape β which is fixed by τ, we can reconstruct σ and A by first finding $T = \tau^{-1}T'$.

Then, if $\alpha = (\alpha_1, \alpha_2, \dots)$, create a matrix of sets A'' by writing the smallest α_1 numbers into the rows of T in the first column of A'', then writing the next smallest α_2 numbers into the rows of T in the second column of A'', and so on.

For instance, in our running example T is the tableau displayed earlier in this proof and $\alpha = (3, 3, 2)$. This means that the matrix A'' should have the integers $1, 2, 3$ appearing somewhere in the first row, the integers $4, 5, 6$ appearing in the second row, and the integers $7, 8$ in the third row. Our construction gives

$$A'' = \begin{bmatrix} \{2,3,1\} & 0 & 0 & 0 \\ 0 & \{5\} & \{4,6\} & 0 \\ 0 & \{7\} & 0 & \{8\} \end{bmatrix}.$$

because the $2, 1, 3$ integers appear in the first row of T and therefore appear in the first column of A'', the 5 and 7 appear in the second row of T and therefore appear in the second column of T, etc.

Then let σ be the permutation which changes A' into A''. In the running example, this $\sigma = (1\,2\,3)(4\,5)$. Finally, create A from A' by turning each set in A' into its size.

Therefore we have shown that the number of pairs (σ, A) is equal to the number of pairs (τ, T). Since the summation allows for the choice of τ and the choice of T is given by the number of tabloids T of shape β fixed by σ), the pair (τ, T) The is counted by the right hand side of the identity in the statement of the theorem. $\square$

Theorem 101. *For all $\alpha, \beta \vdash n$, we have $\left\langle \psi^\alpha, \psi^\beta \right\rangle = \sum_{\lambda \vdash n} K_{\lambda,\alpha} K_{\lambda,\beta}$.*

Proof. Since $\psi^\alpha = \chi^{1\uparrow_{S_\alpha}^{S_n}}$ by (Theorem 60), we can use Frobenius reciprocity (Theorem 40) to see that

$$\begin{aligned}\left\langle \psi^\alpha, \psi^\beta \right\rangle_{S_n} &= \left\langle \chi^1, \chi^{1\uparrow_{S_\beta}^{S_n} \downarrow_{S_\alpha}^{S_n}} \right\rangle_{S_\alpha} \\ &= \frac{1}{|S_\alpha|} \sum_{\sigma \in S_\alpha} \chi^{1\uparrow_{S_\beta}^{S_n} \downarrow_{S_\alpha}^{S_n}}(\sigma) \\ &= \frac{1}{|S_\alpha|} \sum_{\sigma \in S_\alpha} (\text{the number of tabloids } T \text{ of shape } \beta \text{ fixed by } \sigma).\end{aligned}$$

Theorem 100 says that this last sum is equal to $\mathcal{N}_{\alpha,\beta}$, the number of nonnegative integer valued matrices A with row sum α and column sum β. If we restrict RSK to these matrices A as input, then the output is pairs of the form (P, Q) where P has shape λ and content α and Q has shape λ and content β for some $\lambda \vdash n$. These pairs are counted by

$$\sum_{\lambda \vdash n} K_{\lambda,\alpha} K_{\lambda,\beta},$$

as needed. □

Definition. *Define χ^λ to be the character of the representation found in ψ^λ (this is the representation $1\uparrow_{S_\lambda}^{S_n}$) but not found in ψ^μ for any $\mu < \lambda$ (here we are using the reverse lexicographic order on integer partitions).*

We remark that the notation χ^λ_μ has already been defined to give the number of rim hook tableaux of shape λ and content μ and here we are defining χ^λ (without any subscript) to be the character of a representation. These two ideas will eventually be shown to be the same, but for this section we are thinking of χ^λ as the character of a representation.

Example 102. Let us find χ^λ for all $\lambda = (4), (3, 1)$, and $(2, 2)$. We begin with $\lambda = (4)$, where we see that

$$\chi^{(4)} = \psi^{(4)} = \chi^{1\uparrow_{S_{(n)}}^{S_n}} = \chi^1 = 1$$

and so $\chi^{(4)}$ is the trivial representation.

The next integer partition is $(3, 1)$. Using Theorem 101, we see

$$\left\langle \psi^{(3,1)}, \chi^{(4)} \right\rangle = \left\langle \psi^{(3,1)}, \psi^{(4)} \right\rangle = \sum_{\lambda \vdash n} K_{\lambda,(3,1)} K_{\lambda,(4)} = K_{(4),(3,1)} = 1.$$

This sum collapsed down to one term since the only column strict tableau of shape λ and content (4) is a row of all 1's (which is when $\lambda = (4)$). Therefore we define $\chi^{(3,1)}$ to be the character which satisfies $\psi^{(3,1)} = \chi^{(3,1)} + \chi^{(4)}$; in other words, we have

$$\chi^{(3,1)} = \psi^{(3,1)} - \chi^{(4)}.$$

Is this character irreducible? We see that

$$\begin{aligned}\left\langle \chi^{(3,1)}, \chi^{(3,1)} \right\rangle &= \left\langle \psi^{(3,1)} - \chi^{(4)}, \psi^{(3,1)} - \chi^{(4)} \right\rangle \\ &= \left\langle \psi^{(3,1)}, \psi^{(3,1)} \right\rangle - 1 \\ &= \sum_{\lambda \vdash n} K^2_{\lambda,(3,1)} - 1 \\ &= K^2_{(3,1),(3,1)} + K^2_{(4),(3,1)} - 1\end{aligned}$$

Since there is only one column strict tableau of shape $(3,1)$ and content $(3,1)$ and there is only one column strict tableau of shape (4) and content $(3,1)$, this inner product is 1, showing that $\chi^{(3,1)}$ is indeed irreducible.

We can continue this process with the integer partition $(2,2)$. We see

$$\left\langle \psi^{(2,2)}, \chi^{(4)} \right\rangle = \left\langle \psi^{(2,2)}, \psi^{(4)} \right\rangle = \sum_{\lambda \vdash n} K_{\lambda,(2,2)} K_{\lambda,(4)} = K_{(4),(2,2)}$$

and we see

$$\begin{aligned}\left\langle \psi^{(2,2)}, \chi^{(3,1)} \right\rangle &= \left\langle \psi^{(2,2)}, \psi^{(3,1)} - \psi^{(4)} \right\rangle \\ &= \sum_{\lambda \vdash n} K_{\lambda,(2,2)} K_{\lambda,(3,1)} - \sum_{\lambda \vdash n} K_{\lambda,(2,2)} K_{\lambda,(4)} \\ &= K_{(3,1),(2,2)} K_{(3,1),(3,1)} + K_{(4),(2,2)} K_{(4),(3,1)} - K_{(4),(2,2)} K_{(4),(4)} \\ &= K_{(3,1),(2,2)}.\end{aligned}$$

Therefore we define $\chi^{(2,2)}$ to be the character which satisfies

$$\psi^{(2,2)} = K_{(4),(2,2)} \chi^{(4)} + K_{(3,1),(2,2)} \chi^{(3,1)} + \chi^{(2,2)};$$

in other words, we have

$$\chi^{(2,2)} = \psi^{(2,2)} - K_{(4),(2,2)} \chi^{(4)} - K_{(3,1),(2,2)} \chi^{(3,1)}.$$

By finding the inner product of $\chi^{(2,2)}$ with itself it can be shown that this character is irreducible.

Theorem 103. *For $\alpha \vdash n$, we have*

a. χ^β *is irreducible for all* $\beta \vdash n$*, and*

b. $\langle \psi^\alpha, \chi^\beta \rangle = K_{\beta,\alpha}$.

Proof. We prove this by induction on the reverse lexicographic order on integer partitions. To begin, we see that $\chi^{(n)}$ is the trivial representation and is therefore irreducible.

Now assume that χ^λ is irreducible and that $\langle \psi^\alpha, \chi^\lambda \rangle = K_{\lambda,\alpha}$ for all $\lambda < \beta$. We have that

$$\chi^\beta = \psi^\beta - \sum_{\lambda < \beta} K_{\lambda,\beta} \chi^\lambda,$$

and so, with the help of Theorem 101, we have

$$\begin{aligned}
\langle \chi^\beta, \chi^\beta \rangle &= \left\langle \psi^\beta - \sum_{\lambda<\beta} K_{\lambda,\beta}\chi^\lambda, \psi^\beta - \sum_{\lambda<\beta} K_{\lambda,\beta}\chi^\lambda \right\rangle \\
&= \sum_{\lambda\vdash n} K^2_{\lambda,\beta} - 2\sum_{\lambda<\beta} K^2_{\lambda,\beta} + \sum_{\lambda<\beta} K^2_{\lambda,\beta} \\
&= K^2_{\beta,\beta},
\end{aligned}$$

which is equal to 1. Therefore χ^β is irreducible.

As for the second statement in the theorem, we see

$$\begin{aligned}
\langle \psi^\alpha, \chi^\beta \rangle &= \left\langle \psi^\alpha, \psi^\beta - \sum_{\lambda<\beta} K_{\lambda,\beta}\chi^\lambda \right\rangle \\
&= \sum_{\lambda\vdash n} K_{\lambda,\alpha}K_{\lambda,\beta} - \sum_{\lambda<\beta} K_{\lambda,\beta}K_{\lambda,\alpha} \\
&= K_{\beta,\alpha}K_{\beta,\beta},
\end{aligned}$$

which is equal to $K_{\beta,\alpha}$, as needed. □

5 The Frobenius map

Definition. *Let $C(S_n)$ be the class functions for the symmetric group S_n and let 1_λ be the element in $C(S_n)$ such that*

$$1_\lambda(\sigma) = \begin{cases} 1 & \text{if } \sigma \in C_\lambda \\ 0 & \text{otherwise.} \end{cases}$$

The Frobenius map F is a linear map which sends symmetric functions of degree n to class functions in $C(S_n)$ defined by

$$F(p_\lambda/z_\lambda) = 1_\lambda$$

where $z_\lambda = n!/|C_\lambda|$. Since $\{p_\lambda/z_\lambda : \lambda \vdash n\}$ and $\{1_\lambda : \lambda \vdash n\}$ are bases for their respective vector spaces, the Frobenius map F is invertible.

Theorem 104. *If ψ^λ_μ is the value of ψ^λ on the conjugacy class C_μ, then*

$$F(h_\lambda) = \psi^\lambda.$$

Proof. By Theorem 82, $F(h_\lambda) = \sum_{\mu\vdash n} \psi^\lambda_\mu F\left(\frac{p_\mu}{z_\mu}\right) = \sum_{\mu\vdash n} \psi^\lambda_\mu 1_\mu = \psi^\lambda$. □

Theorem 105. *Let χ^λ be the character of the irreducible representation corresponding to $\lambda \vdash n$. (This is the character as seen in Theorem 103.) Then*

$$F(s_\lambda) = \chi^\lambda$$

where s_λ is the Schur symmetric function.

Proof. Using Theorem 96, Theorem 104, and the second statement in Theorem 103 in sequence, we see

$$\sum_{\lambda \vdash n} K_{\lambda,\mu} F(s_\lambda) = F(h_\mu) = \psi^\mu = \sum_{\lambda \vdash n} K_{\lambda,\mu} \chi^\lambda.$$

Since both $\{s_\lambda : \lambda \vdash n\}$ and $\{\chi^\lambda : \lambda \vdash n\}$ are bases in their respective vector spaces and since the Frobenius map sends basis vectors to basis vectors, it must be the case that $F(s_\lambda) = \chi^\lambda$. $\square$

Theorem 106. *The value of the irreducible character χ^λ on the conjugacy class C_μ is equal to χ^λ_μ, the signed sum of all rim hook tableaux of shape λ and content μ.*

Proof. The statement in Theorem 90 can be rewritten as

$$s_\lambda = \sum_{\mu \vdash n} \chi^\lambda_\mu \frac{p_\mu}{z_\mu}.$$

The theorem is proved by applying F to the above and using Theorem 105. $\square$

6 Hooks

The amazing Theorem 106 allows us to find the value of the irreducible character χ^λ on any conjugacy class we like by finding rim hook tableaux. In the special case where the conjugacy class $\mu = (1, \dots, 1)$, this implies that the degree of the irreducible representation corresponding to λ is the number of standard tableaux of shape λ.

The number of standard tableaux of shape λ can be a large and unwieldy! For this purpose there is the beautiful hook length formula. It provides a quick and easy way to find the degree of the irreducible representation corresponding to λ.

6.1 The major index for tableaux

Let $\sigma = \sigma_1 \cdots \sigma_n$ be a permutation. If we let $\mathrm{Des}(\sigma)$ be the set of indices i for which $\sigma_i > \sigma_{i+1}$, then the descent and major index statistics can be written as $\mathrm{des}(\sigma) = |\mathrm{Des}(\sigma)|$ and $\mathrm{maj}(\sigma) = \sum_{i \in \mathrm{Des}(\sigma)} i$. We can adapt these definitions for standard tableaux.

Definition. *If T is a standard tableau, we define $Des(T)$ to be the set of indices i such that $i+1$ appears in a row above i. The descent and major index statistic for standard tableau are defined by $des(T) = |Des(T)|$ and $maj(T) = \sum_{i \in Des(T)} i$.*

Theorem 107. *For all $\lambda \vdash n$,*

$$s_\lambda(1, q, q^2, \dots) = \frac{1}{(1-q)(1-q^2)\cdots(1-q^n)} \sum_{T \in ST_\lambda} q^{maj(T)}$$

where ST_λ is the set of standard tableaux of shape λ.

Proof. Let $\mathrm{sum}(T)$ denote the sum of the integers in a tableau T. In Exercise 53 in last quarter's discrete mathematics course, it is seen that $s_\lambda = \sum_{R \in RCS_\lambda} w(R)$ where RCS_λ is the set of reverse column strict tableaux (these are tableaux where the integer labeling weakly decreases in rows and strictly decreases up columns). Therefore $s_\lambda(1, q^1, q^2, \dots)$ is equal to $\sum q^{\mathrm{sum}(R)-n}$ where the sum runs over all $R \in RCS_\lambda$.

This says that we describe the left side of the identity in Theorem 107 by considering reverse column strict tableaux filled with the integers $0, 1, 2, \dots$. For example, one object counted by this sum is here:

3	0	0		
6	4	3		
12	7	7	7	5

The $\sum_{T \in ST_\lambda} q^{\mathrm{maj}(T)}$ term can also provide us with a reverse column strict tableau by reinterpreting the major index statistic in a standard tableau. For example, suppose we let

$T =$

8	10	11		
5	7	9		
1	2	3	4	6

The major index for T is $4 + 6 + 7 + 9$ (the descents are indicated in red). We can create a reverse column strict tableau T' by taking index i that is a descent in T and adding 1 to each cell in T with a position that is less than i. Doing this for the above example gives

$T' =$

1	0	0		
3	2	1		
4	4	4	4	3

This process will not give us all possible reverse column strict tableaux filled with arbitrary integers since by our construction T' must have a 0 in top row and since i must appear in T' whenever $i+1$ appears in T'.

We can now use the $1/((1-q)(1-q^2)\cdots(1-q^n))$ term in 107 to change T' into an arbitrary reverse column strict tableaux. Interpret this term as providing an integer partition μ with exactly n parts (with parts of size 0 allowed). We can use μ to increase the integers in T' by increasing cell i in T with μ_i.

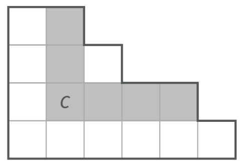
For example, if $\mu = (8, 3, 3, 3, 3, 2, 2, 2, 2, 0, 0)$, then using μ to increase the T' shown above gives

3	0	0		
6	4	3		
12	7	7	7	5

This process is reversible as the tableau T' and the integer partition μ can be reconstructed from this resultant tableaux. This proves the theorem. $\square$

6.2 Hooks

Let λ be an integer partition of n. The hook of a cell c is the "L" shaped subset of cells in the Young diagram of λ consisting of the cell c, all cells to the right of c and in the same row, and all cells above c and in the same column. For example, the hook of a cell c is shaded in the Young diagram below:



.

We define the hook length of the cell c, denoted $h(c)$, to be the number of cells in the hook of c. Below we have filled each cell c in the Young diagram for the integer partition $(6, 5, 3, 2)$ with its hook length $h(c)$:

2	1				
4	3	1			
7	6	4	2	1	
9	8	6	4	3	1

.

Theorem 108. *For any $\lambda = (\lambda_1, \lambda_2, \dots) \vdash n$,*

$$q^{-0\cdot\lambda_1 - 1\cdot\lambda_2 - \cdots} s_\lambda(1, q, q^2, \dots) = \prod_{c\in\lambda} \frac{1}{1 - q^{h(c)}}.$$

where the notation $c \in \lambda$ means that c is a cell in the Young diagram for λ.

Proof. As seen in the proof of Theorem 107, $s_\lambda(1, q, q^2, \dots) = \sum_{T\in CS_\lambda} q^{\mathrm{sum}(T)-n}$ where $\mathrm{sum}(T)$ denotes the sum of the elements in T. Use the factor of q^{-n} in this sum to subtract 1 from each integer in $T \in CS_\lambda$ and use the $q^{-0\cdot\lambda_1-1\cdot\lambda_2-\cdots}$ term to subtract $(i-1)$ from each integer in row i of T. This process changes each $T \in CS_\lambda$ into a tableau filled with nonnegative integers such that

1. the integers weakly increase when reading bottom to top within columns, and
2. the integers weakly increase when reading left to right within rows.

A tableau which satisfies the above conditions is called a reverse plane partition.

Let RPP_λ be the set of reverse plane partitions of shape λ. Given any $T \in RPP_\lambda$, let T_c be the nonnegative integer found in cell c. By defining the weight of $T \in RPP_\lambda$ to be $w(T) = \prod_{\text{cells } c \text{ in } T} q^{T_c}$, we now have

$$q^{-0\cdot\lambda_1 - 1\cdot\lambda_2 - \cdots} s_\lambda(1, q, q^2, \dots) = \sum_{R\in RPP_\lambda} w(R). \tag{6}$$

Let T_λ denote the set of tableaux of shape λ, filled freely with nonnegative integers. Looking at the right side of the equality in the statement of the theorem, write each term in the product $\prod_{c\in\lambda} 1/(1-q^{h(c)})$ as a geometric series and expand, selecting a term of the form $(q^{h(c)})^i$ for each $c \in \lambda$. Record the choices of i made for each cell by placing an i into cell c in a tableau $T \in T_\lambda$. By defining the hook weight of $T \in T_\lambda$ to be $hw(T) = \prod_{\text{cells } c \text{ in } T}(q^{h(c)})^{T_c}$, we now have

$$\prod_{c\in\lambda} \frac{1}{(1-q^{h(c)})} = \sum_{T\in T_\lambda} hw(T). \tag{7}$$

Comparing equations (6) and (7), we see that the theorem can be proved by defining a bijection $\varphi : RPP_\lambda \to T_\lambda$ such that $w(R) = hw(\varphi(R))$ for all $R \in RPP_\lambda$. The bijection we will describe is an algorithm to Abraham Hillman and Richard Grassl.

The input to the algorithm is an element $R \in RPP_\lambda$ with not every entry equal to 0. The image of R is the element $T \in T_\lambda$ found by following these steps:

1. Begin with T the tableau of shape λ with all entries 0.

2. Locate the most north west nonzero element in R, say it lies in cell c.

3. Create a path P which moves down and to the right in R by starting at cell c. The next step in P will move down if the cell below the current cell c contains the same integer as c and the next step in P will move to the right otherwise. Then, after moving to the new cell, change the value of c to be the integer in the current cell. Continue creating P by moving down and to the right in this manner until no more moves are possible.

4. Subtract 1 from each cell in R that lies on the path P.

5. If P begins in column j and ends in row i, then add 1 to the row i, column j entry of T.

6. If R contains a nonzero entry, go back to step 2. If R contains only 0's, then the output of the algorithm is T.

We give an example of this process below, where the appropriate path P is shaded in each step:

$$R = \begin{array}{|c|c|c|c|c|} \cline{1-4} 3 & 3 & 3 & 3 & \multicolumn{1}{c}{} \\ \hline 1 & 1 & 2 & 3 & 3 \\ \hline 0 & 1 & 2 & 2 & 2 \\ \hline \end{array} \qquad T = \begin{array}{|c|c|c|c|c|} \cline{1-4} 0 & 0 & 0 & 0 & \multicolumn{1}{c}{} \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline \end{array},$$

$$R = \begin{array}{|c|c|c|c|c|} \cline{1-4} 2 & 2 & 2 & 2 & \multicolumn{1}{c}{} \\ \hline 1 & 1 & 2 & 2 & 2 \\ \hline 0 & 1 & 2 & 2 & 2 \\ \hline \end{array} \qquad T = \begin{array}{|c|c|c|c|c|} \cline{1-4} 0 & 0 & 0 & 0 & \multicolumn{1}{c}{} \\ \hline 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline \end{array},$$

$$R = \begin{array}{ccccc} 1 & 1 & 1 & 2 & \\ 1 & 1 & 1 & 2 & 2 \\ 0 & 1 & 1 & 1 & 1 \end{array} \qquad T = \begin{array}{ccccc} 0 & 0 & 0 & 0 & \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{array},$$

$$R = \begin{array}{ccccc} 0 & 1 & 1 & 2 & \\ 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \qquad T = \begin{array}{ccccc} 0 & 0 & 0 & 0 & \\ 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \end{array},$$

$$R = \begin{array}{ccccc} 0 & 0 & 0 & 2 & \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \qquad T = \begin{array}{ccccc} 0 & 0 & 0 & 0 & \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \end{array},$$

$$R = \begin{array}{ccccc} 0 & 0 & 0 & 1 & \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \qquad T = \begin{array}{ccccc} 0 & 0 & 0 & 1 & \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \end{array},$$

$$R = \begin{array}{ccccc} 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \qquad T = \begin{array}{ccccc} 0 & 0 & 0 & 1 & \\ 1 & 1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 0 \end{array}.$$

This algorithm moves from left to right down the columns of R, so the last entry added to T lies in the right most nonzero column of T. Furthermore, the weakly increasing rows and columns in R force the last entry added to T to appear in the bottom entry of this column. Therefore, since the the last increased cell in T can be identified, this algorithm is reversible: given a pair R, T, recreate the path P by traveling up and to the left, moving up whenever neighboring entries in different rows are equal and moving left otherwise.

The algorithm is a bijection since it is reversible. The algorithm is also weight preserving since at any step in the algorithm we have $w(R) = hw(T)$ by design. This completes the proof. □ □

Theorems 107 and 108 combine to give the following corollary.

Corollary 109. *For any $\lambda = (\lambda_1, \lambda_2, \ldots) \vdash n$, we have*

$$\sum_{T \in ST_\lambda} q^{maj(T)} = \frac{[n]_q!}{\prod_{c \in \lambda} [h(c)]_q} q^{0\lambda_1 + 1\lambda_2 + \cdots}$$

where ST_λ is the set of standard tableaux of shape λ.

Proof. Comparing the expressions for $s_\lambda(1, q, q^2, \dots)$ in theorems 107 and 108,

$$\frac{1}{(1-q)(1-q^2)\cdots(1-q^n)} \sum_{T \in ST_\lambda} q^{\mathrm{maj}(T)} = \frac{1}{\prod_{c\in\lambda}(1-q^{h(c)})} q^{0\lambda_1 + 1\lambda_2 + \cdots}$$

The corollary follows by solving for $\sum_{T\in ST_\lambda} q^{\mathrm{maj}(T)}$ and simplifying. $\square$ $\square$

Let f^λ denote the number of standard tableaux of shape $\lambda \vdash n$. Taking $q = 1$ in Corollary 109 gives that

$$f^\lambda = \frac{n!}{\prod_{c\in\lambda} h(c)}. \tag{8}$$

for any $\lambda \vdash n$. This identity, which gives a wonderful way to find the number of standard tableaux, is known as the hook length formula. For example, instead of calculating $f^{(3,2)}$ by listing the standard tableaux below

$$\begin{array}{|c|c|c} \hline 4 & 5 & \\ \hline 1 & 2 & \multicolumn{1}{c|}{3} \\ \hline \end{array}, \quad \begin{array}{|c|c|c} \hline 3 & 5 & \\ \hline 1 & 2 & \multicolumn{1}{c|}{4} \\ \hline \end{array}, \quad \begin{array}{|c|c|c} \hline 3 & 4 & \\ \hline 1 & 2 & \multicolumn{1}{c|}{5} \\ \hline \end{array}, \quad \begin{array}{|c|c|c} \hline 2 & 5 & \\ \hline 1 & 3 & \multicolumn{1}{c|}{4} \\ \hline \end{array}, \quad \begin{array}{|c|c|c} \hline 2 & 4 & \\ \hline 1 & 3 & \multicolumn{1}{c|}{5} \\ \hline \end{array},$$

we can instead use the hook length formula to see that $f^{(3,2)} = 5!/(4\cdot 3\cdot 2\cdot 1\cdot 1) = 5$.

Since we proved the hook length formula in a somewhat roundabout way, the intuition giving why the formula is correct may not have been fully formed. To gain insight, consider the following incorrect, but enlightening "proof" of (8):

Proof. (incorrect!) Randomly place the integers $1, \dots, n$ into the Young diagram of shape $\lambda \vdash n$ to create a tableau. Such a naïve filling will be a standard tableau exactly when each cell c is the smallest of the integers in its hook. The probability that this happens for any given $c \in \lambda$ is $1/h(c)$. Therefore the probability of creating a standard tableau is $1/\prod_{c\in\lambda} h(c)$.

There are $n!$ random placements of the integers $1, \dots, n$ into the Young diagram of shape λ, so the number of standard tableaux f^λ is equal to $n!/\prod_{c\in\lambda} h(c)$. $\square$ $\square$

The error in the proof is that the probability of one cell $c_1 \in \lambda$ being the smallest integer in c_1's hook is not independent from the probability of a second cell $c_2 \in \lambda$ being the smallest in c_2's hook. Because the probabilities are not independent, we cannot simply multiply them together in order to find the probability of creating a standard tableau. Strangely, however, since the hook length formula is indeed true, the probability of creating a standard tableau by randomly placing the integers $1, \dots, n$ into a Young diagram is $1/\prod_{c\in\lambda} h(c)$ nevertheless!

7 A summary of the main results

These notes introduce the representation theory of the symmetric group. Chapter 1 introduces matrix representations for a finite group G. The key results:

1. Every group action gives rise to a representation (Theorem 4).

2. Every representation is similar to a direct sum of irreducible representations (Theorem 17).

3. The inner product of the character of a representation A with the character of an irreducible representation B gives the number of times B appears in A (Theorem 44).

Therefore, if we know the characters of all of the irreducible representations of G, then we can take inner products to find out exactly how any given representation breaks up into irreducibles. This is why character tables are useful.

Chapters 2 through 5 specialize the results of Chapter 1 to the symmetric group, with the goal of finding all of the irreducible representations and then to give a nice combinatorial description of their characters. Much depends on understanding the combinatorics of tableaux, with the most relevant type of tableaux being tabloids, column strict tableaux, and rim hook tableaux.

The action of S_n on tabloids of shape μ gives the representation $1\uparrow_{S_\mu}^{S_n}$, the representation with character ψ^μ (Theorem 60). As the integer partition μ varies from (n) to (1^n) in the reverse lexicographic order, this representation contains copies of each of the irreducible representations of S_n, gaining exactly one new irreducible representation at each step (Theorem 103). This is how we defined the irreducible character χ^λ.

Since tableaux are central to the theory, symmetric functions play a significant role since they may be considered multivariate generating functions for tableaux. The connection between the vector space of symmetric functions and the set of functions which are constant on the conjugacy classes of S_n is made explicit by the Frobenius map.

The Frobenius map turns any relationship between the Schur, homogeneous, or p_λ/z_λ symmetric functions into a statement about the irreducible characters χ^λ, the characters ψ^λ, or the indicator function 1_λ for the conjugacy classes of S_n. This map is the bridge which allows us to use symmetric functions (and therefore tableaux) to understand the irreducible characters of S_n.

In particular, the s-to-p/z transition matrix is the character table for S_n. Thus, by Theorem 90, the character table for S_n can be constructed by looking at rim hook tableaux. These characters are the traces of the irreducible representations of the symmetric group, given by the Specht matrices S^λ.

Finally, the hook length formula provides a quick and easy way to find the degree of each irreducible representation.

8 Exercises

1. What if we relax the definition of a matrix representation X to not require $X(g)$ to be invertible? This exercise gives the answer.

a. Let E be an $n \times n$ matrix over $\mathbb{C}$ such that $E \neq I_n$, $E \neq 0$, and $E^2 = E$. Show that there is an $n \times n$ matrix T such that

$$E = T^{-1} \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} T$$

for some k.

b. Let G be a group and X be a function mapping G into the set of (possibly singular) $n \times n$ matrices over $\mathbb{C}$ such that $X(\varepsilon) \neq 0$ and $X(gh) = X(g)X(h)$ for all $g, h \in G$. Show that there is an $n \times n$ matrix T and a matrix representation Y of G such that

$$X(g) = T^{-1} \begin{bmatrix} Y(g) & 0 \\ 0 & 0 \end{bmatrix} T$$

for all $g \in G$.

2. For $\sigma = \sigma_1 \cdots \sigma_n \in S_n$ written in one line notation, $\operatorname{inv} \sigma$ is the number of pairs $i < j$ such that $\sigma_i > \sigma_j$. Define $\operatorname{sign} \sigma = (-1)^{\operatorname{inv} \sigma}$. Show that switching the position of any two integers in σ changes $\operatorname{inv} \sigma$ by an odd number. This fact implies

$$\operatorname{sign}(\sigma\tau) = \operatorname{sign} \sigma \operatorname{sign} \tau$$

where τ is any transposition. Since any permutation can be written as a product of transpositions, this means that the sign representation $X(\sigma) = \operatorname{sign}(\sigma)$ is indeed a matrix representation of S_n.

3. Let $G = \{g_1, \ldots, g_n\}$ be a finite group. The right regular representation R is the representation given by the action of G on itself by right multiplication. That is,

$$R(g)_{i,j} = \begin{cases} 1 & \text{if } g_i g = g_j, \\ 0 & \text{otherwise.} \end{cases}$$

for all $g \in G$. Let L be the left regular representation. Show that $L(g)R(h) = R(h)L(g)$ for all $g, h \in G$.

4. Let X be a representation of H and $H \leq G$. Show that the induced representation $X{\uparrow}_H^G$ is indeed a representation.

5. Show that if X is a representation of K and $K \leq H \leq G$, then $X{\uparrow}_K^G$ and $(X{\uparrow}_K^H){\uparrow}_H^G$ are similar.

6. Give an example of a degree 4 representation for S_3 which does not have the trivial or the sign representation as an irreducible component.

7. This exercise shows that many of our results do not hold over arbitrary fields. Let X be the representation of $\mathbb{Z}_3 = \{0, 1, 2\}$ defined by

$$X(1) = \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}.$$

Show that X is reducible over $\mathbb{C}$ but not over $\mathbb{R}$.

8. This exercise shows that many of our results do not hold for infinite groups.

a. Let G be the infinite group $\mathbb{R}^+$ (the only time we will use an infinite group) with the operation of multiplication. Show that the function X defined by

$$X(x) = \begin{bmatrix} 1 & 0 \\ \ln x & 1 \end{bmatrix}$$

for all $x \in \mathbb{R}^+$ is an irreducible representation.

b. Show that

$$\operatorname{span}\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

is an X-invariant subspace of $\mathbb{C}^2$, showing Theorem 19 does not hold. Where in the proof of Theorem 19 do we use the finiteness of G?

9. Let X be a representation of G. Show that each eigenvalue of $X(g)$ is a root of unity. Use this to show that $\chi^X(g^{-1}) = \chi^{\overline{X}}(g)$ for all $g \in G$.

10. Let C_σ be the conjugacy class for σ in the alternating group A_n.

a. Show that if there is a $\tau \in S_n \setminus A_n$ such τ and σ commute, then

$$C_\sigma = \{\alpha \in S_n : \sigma \text{ and } \alpha \text{ have the same cycle type}\}.$$

b. Show that if there is no $\tau \in S_n \setminus A_n$ such that τ and σ commute, then

$$\{\alpha \in S_n : \sigma \text{ and } \alpha \text{ have the same cycle type}\}$$

is the disjoint union of equal sized conjugacy classes C_σ and $C_{(12)\sigma(12)}$.

11. Prove that if X and Y are representations of G such that X is degree 1 and Y is irreducible, then $X \otimes Y$ is also an irreducible representation.

12. Let X be an irreducible representation of S_n and let A_n be the alternating group. Show that $X\downarrow_{A_n}^{S_n}$ is reducible if and only if $\chi^X(\sigma) = \text{sign}(\sigma)\chi^X(\sigma)$ for all $\sigma \in S_n$.

13. Let X be a nontrivial irreducible representation of G. Show that $\sum_{g \in G} X(g) = 0$.

14. Let N be a normal subgroup of G and let Y be a representation of G/N. Show that the representation X of G defined by $X(g) = Y(gN)$ is irreducible if and only if Y is irreducible.

15. Show that a group G has a nontrivial normal subgroup N (that is, G is not simple) if and only if there is a nontrivial irreducible representation X of G such that $\chi^X(g) = \chi^X(\varepsilon)$ for some $g \neq \varepsilon$.

16. Let $\operatorname{fixed}(\sigma)$ be the number of fixed points (the number of one-cycles) in $\sigma \in S_n$.

a. Use the exponential formula to show that $\displaystyle\sum_{n=0}^{\infty} \sum_{\sigma \in S_n} y^{\operatorname{fixed}(\sigma)} \frac{x^n}{n!} = \frac{e^{yx-x}}{1-x}.$

b. Use part a. to show that $\frac{1}{n!}\sum_{\sigma\in S_n}\mathrm{fixed}(\sigma)^2 = 2$ for $n \geq 2$.

c. Show that the defining representation for S_n breaks up into two irreducible representations. Conclude that the function $\chi : S_n \to \mathbb{C}$ defined by $\chi(\sigma) = \mathrm{fixed}(\sigma) - 1$ is the character of an irreducible representation of S_n for $n \geq 2$.

17. The dihedral group D_4 is generated by $a = (1\,2\,3\,4)$ and $b = (1\,2)(3\,4)$. Find the character table for D_4. (It may be useful to use Exercise 14 with the normal subgroups generated by a and by ab, a^3b.)

18. The quaternion group Q is generated by $c = (1\,2\,3\,4)(5\,6\,7\,8)$ and $d = (1\,8\,3\,6)(2\,7\,4\,5)$. It can be checked that $c^4 = d^4 = (1)$, $c^2 = d^2$, and $cd = dc^3$. Note that Q contains exactly one element of order 2, namely c^2, where the dihedral group D_4 contains at least two elements of order 2, namely the a^2 and b in Exercise 17. Therefore Q and D_4 are not isomorphic.

Find the character table for Q. (It may be useful to use Exercise 14 with the normal subgroup generated by $c^2 = d^2 = (1\,3)(2\,4)(5\,7)(6\,8)$.)

19. Let $\chi^{(1)}, \ldots, \chi^{(\ell)}$ be the characters of the irreducible representations of G, let $C_1, \ldots, C_\ell$ be the conjugacy classes of G, and let $\chi_j^{(i)}$ be the value of $\chi^{(i)}$ on C_j. Then the character table of G is the matrix $M = [\chi_j^{(i)}]_{i,j}$.

a. Show that $M^{-1} = \left[\frac{|C_i|}{|G|}\overline{\chi_i^{(j)}}\right]_{i,j}$ by showing that $MM^{-1} = I$.

b. Show that $\sum_{j=1}^{\ell}\chi_i^{(j)}\overline{\chi_k^{(j)}} = \begin{cases} |G|/|C_i| & \text{if } i = k, \\ 0 & \text{otherwise.}\end{cases}$

20. Prove that g and h are in the same conjugacy class of a group G if and only if $\chi^X(g) = \chi^X(h)$ for all representations X of G.

21. A group has exactly 6 distinct conjugacy classes $C_1, \ldots, C_6$. The values of two characters $\chi^{(1)}$ and $\chi^{(2)}$ corresponding to irreducible representations on these conjugacy classes are shown below.

	C_1	C_2	C_3	C_4	C_5	C_6
$\chi^{(1)}$	1	$-i$	i	1	-1	-1
$\chi^{(2)}$	2	0	0	-1	-1	2

Complete the character table, find the size of the group, and find the sizes of each of the conjugacy classes.

22. Let D be the defining representation for S_n and let ψ^μ be the character of the representation defined in Section 2.1. Prove that $\psi^{(n-1,1)} = \chi^D$.

23. Let X be any representation of S_n. Using Frobenius reciprocity (Theorem 40) on the inner product of χ^X and $\psi^{(2,1^{n-2})} = \chi^{1\uparrow_{S_2\times S_1\times\cdots\times S_1}^{S_n}}$, show that the value of χ^X on the conjugacy class containing the transpositions has the same parity as the degree of X.

24. Find the actual matrices for

$$S^{(2,2)}\big((1\,2)\big),\quad S^{(2,2)}\big((1\,2\,3)\big),\quad S^{(2,2)}\big((1\,2\,3\,4)\big),\quad \text{and}\quad S^{(2,2)}\big((1\,2)(3\,4)\big)$$

where $S^{(2,2)}$ is the irreducible Specht representation for S_4. The traces of these matrices should match the corresponding values of the character χ^B in the character table for S_4 in Example 53.

25. Let sign_μ be the sign of the permutations in the conjugacy class C_μ for $\mu \vdash n$. By performing a sign reversing and weight preserving involution on the objects $\mathcal{U}$ in the proof of Theorem 82, prove the identity

$$n!e_\lambda = \sum_{\sigma\in S_n} \mathrm{sign}(\sigma)\psi^\lambda(\sigma)p_{\mu(\sigma)}.$$

Alternatively, this identity is $e_\lambda = \sum_{\mu\vdash n} \mathrm{sign}_\mu\, \psi^\lambda_\mu \dfrac{p_\mu}{z_\mu}$, thereby giving us the entries of the e-to-p/z transition matrix.

26. Show that all rim hook tableaux of shape $\mu \vdash nk$ with content $(k, k, \ldots, k)$ have the same sign.

27. Show that $\mathrm{sign}(\sigma)\chi^\lambda_\mu = \chi^{\lambda'}_\mu$ where λ' is the conjugate partition to λ and σ is a permutation with cycle type μ.

28. Determine all λ with $\chi^\lambda_{(n-2,2)} \neq 0$.

29. Show that for any partition μ and for all $n \geq 1$, $e_n s_\mu = \displaystyle\sum_{\substack{\lambda/\mu \text{ is a skew column}\\ \text{with } n \text{ cells}}} s_\lambda.$

30. A tableau T of shape λ is row strict if integers strictly increase in rows, reading from left to right, and weakly increase in columns, reading from bottom to top. Let T be a row strict tableaux and j an integer. We define the dual row insertion of j into T, denoted $T \overset{d}{\leftarrow} j$, to be the tableau found by following three rules:

1. If T is the empty tableau, then $T \overset{d}{\leftarrow} j$ is the column strict tableaux with 1 cell which contains the integer j.

   If T is not empty, we assume that the first row of T contains the integers $a_1 < \cdots < a_n$ and T' is the row strict tableau that is found by removing the first row of T.

2. If $a_n < j$, then $T \overset{d}{\leftarrow} j$ results from T by adding a cell containing j at the end of the bottom row of T.

3. If $a_k \geq j$, then let a_k be the leftmost entry in bottom row of T that is greater than or equal to j. Replace a_k with j and then dual row insert a_k into T'.

Calculate $\emptyset \overset{d}{\leftarrow} 1 \overset{d}{\leftarrow} 1 \overset{d}{\leftarrow} 3 \overset{d}{\leftarrow} 2 \overset{d}{\leftarrow} 3 \overset{d}{\leftarrow} 1 \overset{d}{\leftarrow} 1 \overset{d}{\leftarrow} 3 \overset{d}{\leftarrow} 2 \overset{d}{\leftarrow} 2 \overset{d}{\leftarrow} 1$.

31. By expanding each term in the infinite product as a geometric series and then interpreting each choice of $x_i^m y_j^m z^m$ as an entry in a nonnegative integer valued matrix A, prove that

$$\prod_{i,j \geq 1} \frac{1}{1 - x_i y_j z} = \sum_{\lambda} s_\lambda(x_1, x_2, \ldots) s_\lambda(y_1, y_2, \ldots) z^{|\lambda|}.$$

32. Let f^λ denote the number of column strict tableaux of shape λ which contain the integers $1, \ldots, n$. Simplify $\sum_{\lambda \vdash n} \left(f^\lambda\right)^2$.

33. Let F be the Frobenius map. Find these values:

a. $F(e_\lambda)$,

b. $F^{-1}(\chi^D)$ where D is the defining representation,

c. $F^{-1}(\chi^L)$ where L is the left regular representation.

34. Find the character table for S_5 using rim hook tableaux.

35. What is the result of applying the algorithm in the proof of Theorem 108 to the reverse plane partition shown below?

4	4	5	7	
3	3	3	3	6
1	1	2	3	3
0	1	2	2	2

36. Find the degree of the irreducible representation corresponding to $\lambda = (6, 6, 5, 3, 2, 1, 1)$. (This should be easy if you use the hook length formula!)

37. Find the character table for A_5. Why is A_5 simple? (See Exercise 15).

38. Let S^λ be the Specht representation corresponding to $\lambda \vdash n$. How do each of these representations break up as a direct sum of irreducible representations?

a. $S^{(2,2)} \big\downarrow_{S_{(2,2)}}^{S_4}$.

b. $1\!\uparrow_{S_{(3,2,1)}}^{S_6}$.

c. $\left(S^{(2,1)} \otimes S^{(1,1)}\right)\big\uparrow_{S_{(3,2)}}^{S_5}$