# Representation Theory

Lectures by Stuart Martin
Notes by David Mehrle
dfm33@cam.ac.uk

Cambridge University
Mathematical Tripos Part III
Lent 2016

## Contents



Last updated May 20, 2016.

## Contents by Lecture

# 1 Introduction

**Remark 1.1.** "Welcome to rep theory. This is kind of a big audience, so I'll do my best to reduce it by half by at least Monday."

The topics we're going to cover in this course are as follows:

- Overview of representations and characters of finite groups.
- Representations of symmetric groups: Young symmetrizers, Specht modules, branching rule, Gelfand-Tzetlin bases. (This is a very modern approach to representation theory of $S_n$).
- Young Tableaux, hook-length formula, RSK algorithm (and what Serre said was "the most beautiful proof in all of mathematics.")

All of this started with Young, who was actually a clergyman.

**References**

- B.E. Sagan, The Symmetric Group: representations, combinatorial algorithms and symmetric functions (2nd edn), GTM 203, Springer 2001. The classical approach to representation theory of $S_n$.
- T. Ceccherini-Silberstein, F. Scarabotti and F. Tolli, Representation theory of the symmetric groups: the Okounkov-Vershik approach, character formulas and partition algebras, CUP 2010. This has all of the modern stuff in it, including Gelfand-Tzetlin bases.
- R.P. Stanley, Enumerative Combinatorics, Volume 2 (Chapter 7), CUP 2001.
- James and Kerber CUP '86 (You can tell who wrote which bits, because the stuff James wrote is all correct and everything Kerber wrote is wrong).
- A. Kleschchev, Linear and Projective Representations of Symmetric Groups. Very dense, and has more than what we need.
- W. Fulton, Young tableaux, Cambridge University Press, 1997.

## 1.1 Basic Representation Theory

What is representation theory all about? We have groups on one hand, and symmetries of some object on the other hand. It is the total opposite of geometry. In geometry, we have some object and try to figure out what groups describe it. In representation theory, we are given a group and we want to find the things that are described by the groups.

| GROUPS | | SYMMETRIES OF SOMETHING |
|---|---|---|
| symmetric group $S_n$ | $\longleftrightarrow$ | finite set |
| general linear group $GL_n(\mathbb{C})$ | $\longleftrightarrow$ | vector space $V$, $\dim V = n$ |

For us, $\mathrm{GL}_n(\mathbb{C})$ is the main continuous group, and $S_n$ is the main discrete group we will work with.

**Definition 1.2.** Let $G$ be a group. A (complex, finite dimensional, linear) **representation** of $G$ is a homomorphism $\rho\colon G \to \mathrm{GL}(V)$ where $V$ is some finite-dimensional vector space over $\mathbb{C}$.

Equivalently, a representation is a homomorphism $R\colon G \to \mathrm{GL}_n(\mathbb{C})$, in which case we may think about matrices $R(g)$ instead of endomorphisms $\rho(g)$.

Note that we have $R(g_1g_2) = R(g_1)R(g_2)$ and $R(g^{-1}) = R(g)^{-1}$.

**Example 1.3.** Let $C_n$ be the finite cyclic group of order $n$ generated by $g$: $C_n = \{1, g, \dots, g^{n-1}\}$ such that $g^n = 1$. A representation of $G$ on $V$ defines an invertible endomorphism $\rho(g) \in \mathrm{GL}(V)$ with $\rho(1) = \mathrm{id}_V$ and $\rho(g^k) = \rho(g)^k$. Therefore, all other images of $\rho$ are determined by the single operator $\rho(g)$.

So what are all the representations of $C_n$? The one dimensional representations $R\colon C_n \to \mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^\times$ are completely determined by $R(g) = \zeta$ with $\zeta^n = 1$. Hence, $\zeta$ is an $n$-th root of unity. There are $n$ non-isomorphic 1-dimensional representations of $C_n$.

Actually, this isn't an accident:

**Lemma 1.4.** For any abelian group $G$, the number of one-dimensional representations is $|G|$.

**Example 1.5** (Continued from Example 1.3)**.** What about the $d$-dimensional representations? Choose a basis of $V$, such that $\rho(g)$ corresponds to a matrix $M = R(g)$ which takes Jordan Normal Form.

$$M = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_m \end{bmatrix}$$

where the Jordan blocks $J_k$ are of the form

$$J_k = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}$$

Impose the condition that $M^n = \mathrm{id}_V$. But $M^n$ is also block-diagonal, and the blocks of $M^n$ are just powers of the Jordan blocks.

$$M^n = \begin{bmatrix} J_1^n & & & \\ & J_2^n & & \\ & & \ddots & \\ & & & J_m^n \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

Hence, for any Jordan block $J_k$, we must have $J_k^n = 1$. Now let's compute: let $N$ be the Jordan matrix with $\lambda = 0$, so $J_k = \lambda 1 + N$. Hence,

$$J_k^n = (\lambda 1 + N)^n = \lambda^n \mathrm{id} + \binom{n}{1}\lambda^{n-1}N + \ldots + N^n$$

But $N^p$ for any $p$ is a matrix with zeros and ones only with the ones along a line in position $(i, j)$ with $i = j + p$. So $J^n = \mathrm{id}$ only if $\lambda^n = 1$ and $N = 0$. Thus, each $J_k$ is a $1 \times 1$ block and $M$ must be diagonal with respect to this basis.

We have just proved:

**Proposition 1.6.** If $V$ is a representation of $C_n$, there is a basis of $V$ for which the action of every element of $C_n$ is a diagonal matrix, with the $n$-th roots of 1 on the diagonal. In particular, the $d$-dimensional representations of $C_n$ are classified up to isomorphism by unordered $d$-tuples of $n$-th roots of unity.

**Exercise 1.7.** Do the same thing for representations of the infinite cyclic group $(\mathbb{Z}, +)$. Show that the $d$-dimensional representations of $\mathbb{Z}$ are in bijection with the conjugacy classes of $\mathrm{GL}_d(\mathbb{C})$.

**Exercise 1.8.** If $G$ is a finite abelian group, show that the $d$-dimensional isomorphism classes of representations of $G$ are in bijection with unordered $d$-tuples of 1-dimensional representations.

**Definition 1.9.** Two representations $R_1, R_2$ of $G$ are **equivalent** if for each $g \in G$, $R_1(g) = CR_2(G)C^{-1}$ for some fixed nonsingular matrix $C$.

**Definition 1.10** (Operations on Representations)**.** Let $\rho_1 \colon G \to \mathrm{GL}(V_1)$ and $\rho_2 \colon G \to \mathrm{GL}(V_2)$ be two representations of $G$, with $\dim V_1 = k_1$ and $\dim V_2 = k_2$. Then

(a) the **direct sum** of these representations is a representation $\rho = \rho_1 \oplus \rho_2 \colon G \to \mathrm{GL}(V_1 \oplus V_2)$ of dimension $k_1 + k_2$ such that

$$R(g) = \begin{bmatrix} R_1(g) & 0 \\ 0 & R_1(g) \end{bmatrix}$$

(b) the **tensor product** of these representations is a representation $\rho = \rho_1 \otimes \rho_2 \colon G \to \mathrm{GL}(V \otimes W)$, of dimension $k_1 k_2$.

Last time we defined the tensor product and direct sum of representations. There are many more things we could do here: for any operations on vector spaces, there are similar operations on representations, such as symmetric and exterior powers of representations.

**Definition 1.11.**

(1) A representation $\rho$ is called **decomposable** if it is equivalent to a direct sum of two other representations, $\rho \cong \rho_1 \oplus \rho_2$, with $\rho_1, \rho_2$ both nontrivial, that is, $\dim \rho_2, \rho_2 \geqslant 1$.

Otherwise, $\rho$ is **indecomposable**.

(2) A representation $\rho\colon G \to \mathrm{GL}(V)$ is **reducible** if there is a subspace $W \subsetneq V$ with $W \neq \{0\}$ such that all operators $\rho(g)$ preserve $W$ (for all $w \in W$, $\rho(g)(w) \in W$).

Otherwise, $\rho$ is **irreducible** (sometimes called **simple**).

Clearly, irreducible implies indecomposable.

**Theorem 1.12** (Maschke's Theorem)**.** Over $\mathbb{C}$, for all finite groups $G$, a representation $\rho$ is irreducible if and only if $\rho$ is indecomposable.

Moreover, any representation is a direct sum of irreducible representations, that is,

$$\rho \cong \rho_1 \oplus \rho_2 \oplus \ldots \oplus \rho_k$$

with $\rho_i$ irreducible. In this case, we say $\rho$ is **completely reducible** or **semisimple**.

For any representation $\rho$ of $G$, there is a decomposition

$$\rho \cong \rho_1^{\oplus a_1} \oplus \ldots \oplus \rho_k^{\oplus a_k}$$

where all the $\rho_i$ are distinct irreducible representations; the decomposition of $\rho$ into a direct sum of these $k$-many factors is unique, as are the $\rho_i$ that occur and their multiplicities $a_i$. This is called the **isotypical decomposition of** $\rho$.

Finally, there are only finitely many irreducible representations.

**Remark 1.13.** Henceforth, we will call an irreducible representation an **irrep**.

**Remark 1.14.** Questions for rep theory

(1) Classify (construct) the irreps, $\rho_1, \ldots, \rho_\ell$ of $G$.

(2) Decompose the tensor product of two representations $\rho_i$ and $\rho_j$ into irreps, since it's rarely irreducible.

$$\rho_i \otimes \rho_j = (\rho_1 \oplus \ldots \oplus \rho_1) \oplus (\rho_2 \oplus \ldots \oplus \rho_2) \oplus \ldots = \rho_1^{\oplus m_1} \oplus \rho_2^{\oplus m_2} \oplus \ldots$$

How should we calculate the multiplicity $m_k = m_{i,j,k}$ of $\rho_k$ in $\rho_i \otimes \rho_j$?

This second question is still unsolved even in characteristic zero, although when we work with Hopf algebras and quantum groups and the like, there are many more techniques available to throw at it, so more is known in that case.

## 1.2 Symmetric Groups, Young diagrams, and Partitions

**Definition 1.15.** The elements $w$ of the **symmetric group** $S_n$ are bijections $w\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$. The operation is composition of maps, written from right to left $w_1 w_2$.

**Remark 1.16.** There are several ways to write elements of $S_n$:

| | |
|---|---|
| **one-line notation** | $w(1)\, w(2)\, \ldots\, w(n)$ |
| **two-line notation** | $\begin{pmatrix} 1 & 2 & \ldots & n \\ w(1) & w(2) & \ldots & w(n) \end{pmatrix}$ |
| **cycle notation** | $w = (a_1, \ldots, a_n)(b_1, \ldots, b_\ell) \ldots$ |

**Example 1.17.** In $S_5$, consider the permutation $w(1) = 2$, $w(2) = 3$, $w(3) = 1$, $w(4) = 4$, $w(5) = 5$. The two line notation is

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 \end{pmatrix}$$

The cycle notation for $w$ is $(1\ 2\ 3)(4)(5)$.

The product of the two elements $u = 2\ 1\ 3 = (1\ 2) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ and $w = 1\ 3\ 2 = (2\ 3) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$ of $S_3$ is $uw = (1\ 2\ 3)$.

**Example 1.18.** Some representations of $S_n$

(1) The **trivial** representation $w \mapsto (1)$.

(2) The **sign / alternating** representation $w \mapsto \operatorname{sgn}(w) = \begin{cases} +1 & w \text{ even} \\ -1 & w \text{ odd} \end{cases}$

(3) The **defining / standard** representation $R\colon w \mapsto$ permutation matrix of $w$.

That is, $R(w) = (x_{ij})$ where $x_{ij} = \begin{cases} 1 & w(i) = j \\ 0 & \text{otherwise} \end{cases}$.

For example, if $\{e_1, e_2, e_3\}$ is the standard basis of $\mathbb{C}^3$, then $S_3$ acts in the standard representation by permuting coordinates: $R(w)e_i = e_{w(i)}$.

This representation is not irreducible, because the subspace spanned by the sum of basis vectors is invariant.

**Exercise 1.19.** Check that the trivial and alternating reps are all the 1-dimensional representations of $S_n$.

**Exercise 1.20.** Show that the standard representation $R$ of $S_n$ decomposes as $R = R_1 \oplus R_2$, with $R_1$ trivial and $R_2$ an $(n-1)$-dimensional irrep.

**Definition 1.21.** Let $n \in \mathbb{N}$. A **partition** of $n$ is a finite sequence $\lambda = (\lambda_1, \ldots, \lambda_k)$ such that $\lambda_i \in \mathbb{N}$ and $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_k$ and $\lambda_1 + \lambda_2 + \ldots + \lambda_k = n$. Write $\lambda \vdash n$. The set of all partitions of $n$ is $P(n)$.

Conjugacy classes of $S_n$ are parameterized by the partitions of $n$. The conjugacy class associated with the partition $\lambda \vdash n$ consists of all permutations $w \in S_n$ whose cycle decomposition is of the form

$$w = (a_1, \ldots, a_{\lambda_1})(b_1, \ldots, b_{\lambda_2}) \cdots (c_1, \ldots, c_{\lambda_k}).$$

**Remark 1.22.** Irreps of $S_n$ correspond to partitions of $n$. We've seen that conjugacy classes of $S_n$ are defined by cycle type, and cycle types correspond to partitions. Therefore partitions correspond to conjugacy classes, which correspond to irreps.

**Proposition 1.23.** Given $\lambda \in P(n)$, let $c_\lambda \in \mathbb{C}S_n$ be the sum of all the permutations in $S_n$ with cycle type $\lambda$. Then $\{c_\lambda \mid \lambda \in P(n)\}$ is a basis for $Z(\mathbb{C}S_n)$. (General fact: $\dim Z(\mathbb{C}S_n) = \#$ of conjugacy classes.)

*Proof.* First, note that each $c_\mu$ is invariant under conjugation, since the conjugacy classes are by cycle type. Hence, each $c_\mu \in Z(\mathbb{C}S_n)$.

The $c_\mu$ are linearly independent, since there's no way to add things of different cycle types and end up with zero.

To show that the $c_\mu$ are spanning, let $f \in Z(\mathbb{C}S_n)$ and let $\tau \in S_n$. We have that $\tau f = f\tau$ since $f$ lies in the center, so $\tau f \tau^{-1} = f$. So if we write

$$f = \sum_{\sigma \in S_n} a_\sigma \sigma,$$

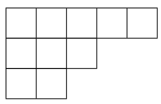
then $a_\sigma = a_{\tau\sigma\tau^{-1}}$. Hence, the $a_\sigma$ are constant on conjugacy classes; if $\sigma, \rho$ have the same cycle type, then $a_\sigma = a_\rho$. So we may write

$$f = \sum_{\lambda \in P(n)} a_\lambda c_\lambda.$$

So the $c_\mu$ are spanning. □

We can conveniently talk about partitions using Young tableaux, but this will first require a long list of definitions.

**Definition 1.24.** A paritition $\lambda \vdash n$ is represented by a what is called alternatively a **Young diagram/frame/Ferrers diagram**, e.g. the partition $(4, 3, 1) \vdash 8$ corresponds to the diagram



**Definition 1.25.** A **box** is described by its coordinates $(x, y)$ where $x$ goes down and $y$ goes across.

**Example 1.26.** For example, in the diagram

| | | | |
|---|---|---|---|
| $a$ | $b$ | $c$ | $d$ |
| $e$ | $f$ | $g$ | |
| $h$ | | | |

$g$ is in box $(2, 3)$.

**Definition 1.27.** The **content** of a box is $c(x, y) = y - x$.

**Definition 1.28.** A **Young tableau of shape** $\lambda$ or $\lambda$-tableau is a bijection between boxes of the Young diagram of shape $\lambda$ and $\{1, \dots, n\}$, e.g.

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

**Definition 1.29.** A Young tableau is called **standard** if the numbers filled into boxes are increasing both along the rows (left-to-right) and along the columns (top-to-bottom).

**Example 1.30.** For example,

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

is standard.

**Definition 1.31.** A box $(x, y)$ is **removable** if there is no box below (in position $(x+1, y)$) or to the right (in position $(x, y+1)$). Precisely, $(x, y)$ is removable if and only if ($x < k$ & $y = \lambda_x > \lambda_{x+1}$) or ($x = k$ & $y = \lambda_k$). Removing such a box produces a Young diagram associated with a partition of $n-1$. Similar for **addable boxes**.

**Example 1.32.** For example,

$$\begin{array}{|c|c|c|c|}\hline a & b & c & d \\ \hline e & f & g \\ \cline{1-3} h \\ \cline{1-1}\end{array}$$

$d, g, h$ are removable but none of the others are. Boxes are addable with coordinates $(1,5)$, $(2,4)$, $(3,2)$ and $(4,1)$, as indicated with $*$'s below.

$$\begin{array}{|c|c|c|c|c|}\hline a & b & c & d & * \\ \hline e & f & g & * \\ \cline{1-4} h & * \\ \cline{1-2} * \\ \cline{1-1}\end{array}$$

**Definition 1.33.** If both $\lambda, \mu \vdash n$, then $\lambda < \mu$ in **lexicographic order** if, for some $i$, $\lambda_j = \mu_j$ for $j < i$ and $\lambda_i < \mu_i$. This defines a total order on $P(n)$.

**Example 1.34.** for example

$$(1^6) < (2, 1^4) < (2^2, 1^2) < \dots < (5, 1) < (6).$$

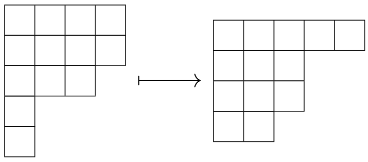
**Definition 1.35.** For $\lambda \vdash n$, $\mathrm{Tab}(\lambda) = \mathrm{SYT}(\lambda) = \{\text{standard } \lambda\text{-tableaux}\}$.

$$\mathrm{SYT}(n) = \bigcup_{\lambda \vdash n} \mathrm{SYT}(\lambda)$$

**Definition 1.36.** Let $\lambda = (\lambda_1, \dots, \lambda_k) \vdash n$. In a Young diagram for $\lambda$, there exists $t := \lambda_1$ columns. The $j$-th column contains exactly $\lambda'_j := |\{i \colon \lambda_i \geqslant j\}|$ boxes. The **conjugate partition** to $\lambda$ is the parittion

$$\lambda' = (\lambda'_1, \dots, \lambda'_t)$$

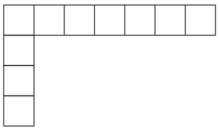
**Example 1.37.** Conjugating the partition $(4,4,3,1,1)$ gives $(5,3,3,2)$, which corresponds to flipping the corresponding Young diagram over the main diagonal:



Notice that conjugating twice gives the original partition back.

**Definition 1.38.** A diagram of $\lambda$ is a **hook** if $\lambda = (n-k, 1^k) = (n-k, 1, 1, \dots, 1)$ for some $0 \leqslant k \leqslant n-1$. $k$ is the **height of the hook**.

**Example 1.39.**



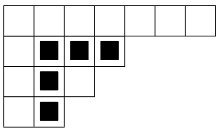
is a hook of height 3.

**Definition 1.40.** Suppose that $\lambda = (\lambda_1, \dots, \lambda_k)$ and $\lambda' = (\lambda'_1, \dots, \lambda'_t)$ are conjugates. The **hook length** of a box of coordinate $(i,j)$ is

$$h(i,j) = h_{ij} := (\lambda_i - j) + (\lambda'_j - i) + i$$

$(\lambda_i - j)$ is the **arm of the hook**: the number of boxes in the same row to the right; $(\lambda'_j - 1)$ is the **leg of the hook**: number of boxes in the same column below, and 1 counts the box in the corner.

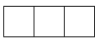
**Example 1.41.** In the partition $\lambda = (7,4,3,2)$



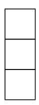
the hook length of the box $(2,2)$ is $h_{22} = 2+2+1 = 5$. The hook is shaded in the diagram.

**Remark 1.42.** The French justify the Young diagrams along the bottom instead of the top. They hate our convention because it violates Descartes's convention whereby the $x$-axis increases to the right and the $y$-axis increases up. The way we draw them is known as the "English convention."

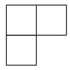
**Example 1.43.** The partitions of 3 correspond to irreps of $S_3$. For example,



is the trivial representation,



is the sign representation, and



is the 2-dimensional component of the regular representation.

**Definition 1.44.** Write $\{V_\lambda \mid \lambda \vdash n\}$ for the irreps of $S_n$ over $\mathbb{C}$. These are called the **Specht modules**.

## 1.3 Things we will prove later

**Remark 1.45.** "No sudden movements or sharp noises; please contain your excitement."

**Proposition 1.46.** $\dim V_\lambda = \#\mathrm{SYT}(\lambda)$.

**Definition 1.47.**

$$f_\lambda := \#\mathrm{SYT}(\lambda)$$

**Proposition 1.48** (Frobenius-Young identity)**.**

$$\sum_{\lambda \vdash n} f_\lambda^2 = n!$$

Combinatorially, Proposition 1.48 states that

$$\#\{(P,Q)\colon P,Q \text{ std tableaux of shape } \lambda, \lambda \vdash n\} = \#S_n = n!$$

and the **Robinson-Schensted-Knuth** (RSK) correspondence gives the bijection.

**Theorem 1.49** (Hook-length formula)**.** $f_\lambda = \dfrac{n!}{\prod\limits_{(i,j)\in\lambda} h(i,j)}$.

**Example 1.50.** If $\lambda = (n,n)$, then

$$f_\lambda = \frac{(2n)!}{(n!)(2\cdot 3\cdots n)(n+1)} = \frac{1}{n+1}\binom{2n}{n}.$$

This is in Richard Stanley's book as part of an infamous exercise, in which he asks you to prove the equivalence of 66 different combinatorial expressions for the Catalan numbers.

**Remark 1.51** (Top Travel Tip)**.** Bring Richard Stanley's book on combinatorics with you when you travel, and you'll never be bored.

Let $G$ be some finite group and let $H \leqslant G$ be a subgroup. There are two operations we can do on representations of $G$ and $H$. Given a representation of $G$, we can restrict to a representation of $H$, and given a representation of $H$, we can induce a representation of $G$. These are linked by Frobenius reciprocity.

**Definition 1.52.** If $V$ is a representation of $G$, then $\mathrm{Res}_H^G V$ is the **restriction** of $V$ to the subgroup $H$. Alternatively written $V{\downarrow_H}$.

**Remark 1.53.** If $V$ is an irreducible $G$-module, then $\operatorname{Res}_H^G V$ need not be irreducible. For example, if $G = S_n$, the irreps are $V_\lambda$ for $\lambda \vdash n$.

To consider $\operatorname{Res}_{S_{n-1}}^{S_n} V_\lambda$, we need to embed $S_{n-1}$ into $S_n$, which we do in the standard way: those permutations of $\{1, 2, \ldots, n\}$ leaving $n$ fixed.

**Definition 1.54.** The **Bratelli diagram** of representations of symmetric groups has

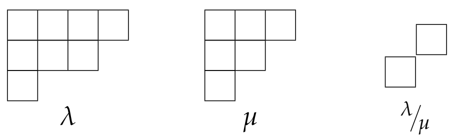
- vertices: irreducible representations of $S_n$
- edges: two vertices $V_\lambda, V_\mu$ are joined by $k$ edges from $\lambda$ to $\mu$ if $V_\mu$ is a representation of $S_{n-1}$ and $V_\lambda$ is a representation of $S_n$, such that $V_\mu$ is an irreducible component of $\operatorname{Res}_{S_{n-1}}^{S_n} V_\lambda$.

**Definition 1.55.** The **Young poset** is the set $\mathbb{Y} = \{\lambda \mid \lambda \vdash n, n \in \mathbb{N}\}$ of all partitions with the poset structure as follows. Let $\mu = (\mu_1, \ldots, \mu_k) \vdash n$ and $\lambda = (\lambda_1, \ldots, \lambda_h) \vdash m$ be partitions in $\mathbb{Y}$. Then we say that

$$\mu \preccurlyeq \lambda \iff m \geqslant n, h \geqslant k \text{ and } \lambda_j \geqslant \mu_j \ \forall j = 1, \ldots, k$$

This simply means that $\mu$ is a subdiagram of $\lambda$.

**Example 1.56.** If $\mu = (3, 2, 1)$ and $\lambda = (4, 3, 1)$, then $\mu \preccurlyeq \lambda$.




We use the notation ${}^{\lambda}/_{\mu}$ to denote the squares that remain after removing $\mu$ from $\lambda$: these are the unshaded ones in the diagram on the right.

**Definition 1.57.** If $\mu, \lambda \in \mathbb{Y}$, then we say that $\lambda$ **covers** (or is covered by) $\mu$ if $\mu \preccurlyeq \lambda$ and $\mu \preccurlyeq \nu \preccurlyeq \lambda, \nu \in \mathbb{Y}$ implies that $\nu = \mu$ or $\nu = \lambda$.

Clearly $\lambda$ covers $\mu$ if and only if $\mu \preccurlyeq \lambda$ and ${}^{\lambda}/_{\mu}$ consists of a single box. We write $\lambda \to \mu$ or $\mu \nearrow \lambda$.

**Definition 1.58.** The **Hasse diagram** of $\mathbb{Y}$ or the **Young (branching) graph** is an oriented graph with vertex set $\mathbb{Y}$ and an arrow $\lambda \to \mu$ if and only if $\lambda$ covers $\mu$.

We will eventually show that the Young graph $\mathbb{Y}$ is the same as the Bratelli diagram for the branching of representations upon restriction from $S_n$ to $S_{n-1}$.

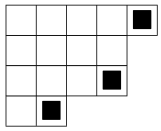
**Lemma 1.59** (Branching Rule)**.**

$$\operatorname{Res}_{S_{n-1}}^{S_n} V_\lambda = \bigoplus_{\substack{\mu \vdash (n-1) \\ \mu \subseteq \lambda}} V_\mu$$

where the direct sum is over all Young diagrams $\mu$ obtained from $\lambda$ by removing a single removable box. Note that if $V_\mu$ occurs at all, it occurs with multiplicity one – the branching is **simple**.

There is a corresponding result for induction, but we won't bother with it right now; see Corollary 5.5.

**Example 1.60.** Consider the Young tableaux $\lambda = (5, 4, 4, 2)$



and the representation $V_\lambda$ is obtained by removing any one of the removable (shaded) boxes.

$$\mathrm{Res}^{S_{15}}_{S_{14}} V_{(5,4,4,2)} = V_{(4,4,4,2)} \oplus V_{(5,4,3,2)} \oplus V_{(5,4,4,1)}$$

To prove Lemma 1.59, we need to think about it combinatorially. To that end, we can associate paths in the Young graph with standard Young tableaux.

**Definition 1.61.** A **path** in the Young graph is a sequence

$$\pi = \left(\lambda^{(n)} \to \lambda^{(n-1)} \to \cdots \to \lambda^{(1)}\right)$$

of partitions $\lambda^{(j)} \vdash j$ such that $\lambda^{(j)}$ covers $\lambda^{(j-1)}$ for $j = 2, 3, \ldots, n$. Notice that a path always ends at the trivial partition $\lambda^{(1)} = (1) \vdash 1$.

Let $n = \ell(\pi)$ be the length of the path $\pi$. Then $\Pi_n(\mathbb{Y})$ is the set of all paths of length $n$ in the Young graph. Further define

$$\Pi(Y) := \bigcup_{n \geqslant 1} \Pi_n(Y).$$

Given $\lambda \vdash n$ and a path $\pi = \left(\lambda^{(n)} \to \lambda^{(n-1)} \to \cdots \to \lambda^{(1)}\right)$, there exists a corresponding standard tableaux $T$ of shape $\lambda$ obtained by placing the integer $k \in \{1, \ldots, n\}$ in the box $\lambda^{(k)}/\lambda^{(k-1)}$.

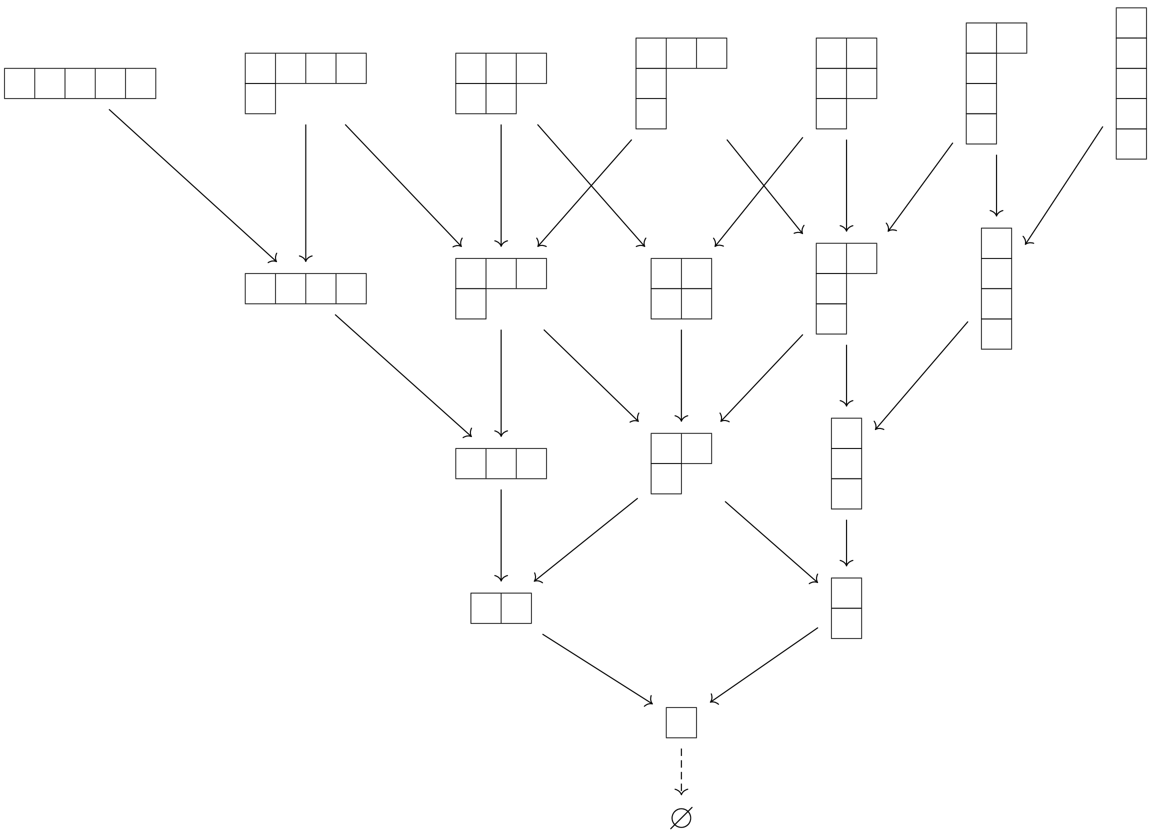



**Figure 1.** Bottom of the Hasse diagram of $\mathbb{Y}$

**Example 1.62.** Consider the path $\pi$

$$\lambda^{(8)} = (4,3,1) \to (4,3) \to (3,3) \to (3,2) \to (2,2) \to (2,1) \to (2) \to (1) = \lambda^{(1)}.$$

We read the path backwards from $\lambda^{(1)}$ to $\lambda^{(8)}$ and add boxes one at a time to reconstruct the standard Young tableaux.

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

So the path $\pi$ corresponds to the standard tableaux

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

**Fact 1.63.** There is a natural bijection between $\Pi_n(\mathbb{Y})$ and $\mathrm{SYT}(n)$, which extends to a bijection $\Pi(\mathbb{Y}) \leftrightarrow \bigcup_{n \geqslant 1} \mathrm{SYT}(n)$.

Once we know that the branching graph for $S_n$ corresponds to the Young graph, we can see deduce interesting results about the representation theory of $S_n$. For instance, we can easily show that $\dim V_\lambda = \#\mathrm{SYT}(\lambda) = f_\lambda$.

Recall the statement of the Branching Rule (Lemma 1.59).

$$\mathrm{Res}^{S_n}_{S_{n-1}}(V_\lambda) = \bigoplus_{\substack{\mu \dashv (n-1) \\ \mu \subseteq \lambda}} V_\mu.$$

Slightly more generally, we could instead restrict to any $m < n$.

$$\mathrm{Res}^{S_n}_{S_m}(V_\lambda) = \mathrm{Res}^{S_{m+1}}_{S_m}\left(\mathrm{Res}^{S_{m+2}}_{S_{m+1}}\left(\cdots \mathrm{Res}^{S_n}_{S_{n-1}}(V_\lambda)\right)\right),$$

and at each step of the consecutive restrictions, the decomposition is simple (multiplicity-free) and it occurs according to the branching graph. Therefore, the multiplicity of $V_\mu$ in the restriction $\mathrm{Res}^{S_n}_{S_m} V_\lambda$ is the number of paths in $\mathbb{Y}$ that start at $\lambda$ and end at $\mu$. This is the number of ways in which you can obtain a diagram of shape $\lambda$ from one of shape $\mu$ from adding successively $(n-m)$ addable boxes to the diagram of shape $\mu$.

Recall Proposition 1.46, which claims that

$$f_\lambda := \dim V_\lambda = \#\mathrm{SYT}(\lambda).$$

We can prove this by counting dimensions in the Branching Rule. We see that

$$\begin{aligned}
\dim V_\lambda &= \sum_{\lambda \to \mu} \dim V_\mu \\
&= \sum_{\lambda \to \mu \to \nu} \dim V_\nu = \cdots \\
&= \sum_{\lambda = \lambda^{(n)} \to \cdots \to \lambda^{(1)} = (1)} \dim V_{(1)} \\
&= \#\text{ of paths from } (\lambda) \text{ to } (1) \text{ in } \mathbb{Y}.
\end{aligned}$$

We'll construct a basis of $V_\lambda$ where each basis vector corresponds to a downward path in the Young graph from $(\lambda)$ to $(1)$. As in the previous lecture, each such path corresponds to a standard Young tableaux of shape $\lambda$.

**Example 1.64.** There are three different paths from $(3,1)$ down to $(1)$ in $\mathbb{Y}$. One such path is the following

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

## 1.4 Back to basics: Young symmetrizers

Let's go back to the problem of constructing these modules $V_\lambda$. The classical method of constructing irreps of $S_n$ is via **Young symmetrizers**. While this approach is fast, it is inferior because it takes a lot of effort to prove things about representations of $S_n$ using Young symmetrizers. In contrast, the Okounkov-Vershik approach that we will take is more effort upfront, but it makes proving things about representations much easier.

**Definition 1.65.** Let $G = \{g_1, \ldots, g_r\}$ be a finite group. The **group algebra** $\mathbb{C}G$ is the $\mathbb{C}$-algebra whose elements are formal linear combinations

$$\alpha_1 g_1 + \cdots + \alpha_r g_r$$

for $\alpha_i \in \mathbb{C}$. Multiplication is given by

$$(\alpha_1 g_1 + \cdots + \alpha_r g_r)(\beta_1 g_1 + \cdots \beta_r g_r) = \sum_{i,j} (\alpha_i \beta_j) g_i g_j = \sum_{k=1}^{r} \Bigg( \sum_{\substack{i,j \\ g_i g_j = g_k}} \alpha_i \beta_j \Bigg) g_k$$

**Example 1.66.** Let $G = S_3$. Then

$$\mathbb{C}G = \mathbb{C}S_3 = \{\alpha 1 + \beta(1\,2) + \gamma(1\,3) + \delta(2\,3) + \varepsilon(1\,2\,3) + \zeta(1\,3\,2)\}$$

is a 6-dimensional $\mathbb{C}$-algebra.

**Definition 1.67.** The **regular representation** $V = \mathbb{C}G$ has a left-action of $G$ by multiplication

$$g(\alpha_1 g_1 + \cdots + \alpha_r g_r) = \alpha_1(g g_1) + \cdots + \alpha_r(g g_r),$$

with $\dim V = |G|$.

Every irreducible representation is contained in the regular representation $G$.

**Theorem 1.68.** *Let $V_i$ be the irreducible representations of $G$, for $i \in I$. Then*

$$\mathbb{C}G = \bigoplus_{i \in I} (\dim V_i) V_i$$

**Corollary 1.69.**

$$\sum_{i \in I} (\dim V_i)^2 = |G|$$

**Example 1.70.**

$$\mathbb{C}S_3 = V_{(3)} \oplus V_{(1,1,1)} \oplus V_{(2,1)} \oplus V_{(2,1)}$$

$V_{(3)}$ corresponds to the trivial representation, $V_{(1,1,1)}$ corresponds to the alternating representation, and $V_{(2,1)}$ is the 2-dimensional component of the standard representation.

So how do we find $V_{(3)}$ in $\mathbb{C}S_3$? Well, it corresponds to a 1-dimensional subspace, so therefore we need a vector to span it. Try $\sum_{w \in S_3} w$. Indeed,

$$V_{(3)} = \left\langle \sum_{w \in S_3} w \right\rangle.$$

Similarly,

$$V_{(1,1,1)} = \left\langle \sum_{w \in S_3} (-1)^w w \right\rangle.$$

We write $(-1)^w = \operatorname{sgn}(w)$.

How do we find $V_{\yng(2,1)} \oplus V_{\yng(2,1)}$ inside $\mathbb{C}S_3$? Consider $c = (1 + (1\,2))(1 - (1\,3)) \in \mathbb{C}S_3$. Take $\mathbb{C}S_3 c \subseteq \mathbb{C}S_3$. The claim is that $V_{\yng(2,1)} \cong \mathbb{C}S_3 c$, which is a special case of the construction called the **Young symmetrizer**.

**Definition 1.71.** Pick any tableau $T$ of shape $\lambda \vdash n$. Define

$$\begin{aligned} P = P_\lambda &= \left\{\sigma \in S_n \mid \sigma \text{ preserves each row of } T\right\} \\ Q = Q_\lambda &= \left\{\sigma \in S_n \mid \sigma \text{ preserves each column of } T\right\} \end{aligned}$$

Now let

$$\begin{aligned} a_\lambda &= \sum_{w \in P} w \in \mathbb{C}S_n \\ b_\lambda &= \sum_{w \in Q} (-1)^w w \end{aligned}$$

Finally, $c_\lambda = a_\lambda b_\lambda$ is the **Young symmetrizer.**

The **Specht module** $V_\lambda$ is the ideal $V_\lambda := \mathbb{C}S_n c_\lambda \subseteq \mathbb{C}S_n$.

**Example 1.72.** If $\lambda = (3, 2)$, and the tableaux is

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

then

$$\begin{aligned} a_\lambda &= (1 + (1\,3) + (1\,4) + (3\,4) + (1\,3\,4) + (1\,4\,3))\,(1 + (2\,5)) \\ b_\lambda &= (1 - (1\,2))(1 - (3\,5)) \end{aligned}$$

**Theorem 1.73.**

(1) Some scalar multiple of $c_\lambda$ is an idempotent: $c_\lambda^2 = n_\lambda c_\lambda$, for some $n_\lambda \in \mathbb{C}$.

(2) $V_\lambda$ is an irrep of $S_n$.

(3) Every irrep of $S_n$ can be obtained in this way for a unique partition $\lambda \vdash n$.

*Proof.* See Fulton & Harris §4.2 / Example sheet. □

**Remark 1.74.** (1) We will ultimately see that each irrep can be defined over $\mathbb{Q}$ instead of over $\mathbb{C}$. In fact, the scalar $n_\lambda$ in Theorem 1.73(1) is $n_\lambda = {}^{n!}/_{\dim V_\lambda}$.

(2) Any tableau gives an irrep, not just standard ones. They will be isomorphic to those of the standard ones, but in with things in a different order.

**Example 1.75.** $\lambda = (n)$, $c_\lambda = a_\lambda = \sum_{w \in S_n} w$

$$V_\lambda = \mathbb{C}S_n c_\lambda = \mathbb{C}S_n \left(\sum_{w \in S_n} w\right) = \text{trivial rep}$$

$\mu = (1^n) = (1, 1, \ldots, 1)$ $c_\mu = b_\mu = \sum_{w \in S_n} (-1)^w w$

$$V_\mu = \mathbb{C}S_n c_\mu = \mathbb{C}S_n \left( \sum_{w \in S_n} (-1)^w w \right) = \text{alternating rep}$$

$\lambda = (2, 1) \vdash 3$

$$c_{(2,1)} = (1 + (2\,1))(1 - (1\,3)) = 1 + (1\,2) - (1\,3) - (1\,3\,2) \in \mathbb{C}S_3.$$

$$V_{(2,1)} = \langle c_{(2,1)}, (1\,3)c_{(2,1)} \rangle$$

**Exercise 1.76.** Prove $V_\lambda \otimes \text{sgn} = V'_\lambda$, where $\lambda'$ is the conjugate partition to $\lambda$.

## 1.5 Coxeter Generators

**Definition 1.77.** $S_n$ is generated by the **adjoint transpositions** $s_i = (i, i+1)$ for $1 \leqslant i \leqslant n-1$ with **coxeter relations**

$$s_i^2 = 1 \qquad s_i s_j s_i = s_j s_i s_j \text{ if } |i - j| = 1 \qquad s_i s_j = s_j s_i \text{ for } |i - j| \geqslant 2$$

**Definition 1.78.** $s_{i_1} \cdots s_{i_k}$ is **reduced** if there is no shorter product. In this case $k$ is the **Coxeter length**.

How do these act on tableaux? If $T$ is a tableau of shape $\lambda$, let $\sigma \in S_n$. We obtain a new tableau $\sigma T$ by replacing $i$ with $\sigma(i)$ for each $i$.

**Example 1.79.** Let $\sigma = (3\,8\,6\,7\,8\,4)(2\,5)$.

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

Notice that $T$ is not standard, but $\sigma T$ is.

**Definition 1.80.** If $T$ is standard, say $s_i$ is **admissible** for $T$ if $s_i T$ is still standard. So $s_i$ is admissible for $T$ if and only if $i, i+1$ belong to neither the same row nor the same column of $T$.

**Definition 1.81.** For $\pi \in S_n$, an **inversion for** $\pi$ is a pair $(i, j)$ with $i, j \in \{1, \ldots, n\}$ such that $i < j \implies \pi(i) > \pi(j)$.

$\mathcal{I}(\pi) = \{\text{all inversions in } \pi\}$. Let $\ell(\pi) = |\mathcal{I}(\pi)|$.

**Theorem 1.82.** The Coxeter length of $\pi$ equals $\ell(\pi)$.

*Proof.* First, notice that

$$\ell(\pi s_i) = \begin{cases} \ell(\pi) - 1 & \pi(i) > \pi(i+1) \\ \ell(\pi) + 1 & \pi(i+1) > \pi(i) \end{cases}$$

This shows that, if $\pi = s_{i_1} s_{i_2} \cdots s_{i_k}$ is a minimal representation of $\pi$ by Coxeter generators, then

$$\ell(\pi) = \ell(s_{i_1} s_{i_2} \cdots s_{i_k}) \leqslant \ell(s_{i_1} s_{i_2} \cdots s_{i_{k-1}}) + 1 \leqslant \ell(s_{i_1} s_{i_2} \cdots s_{i_{k-2}}) + 2 \leqslant \ldots \leqslant k. \tag{1}$$

So the inversion length is bounded above by the Coxeter length.

It remains to show that the Coxeter length of $\pi$ is bounded below by $\ell(\pi)$. Proof by induction on $\ell(\pi)$. If $\ell(\pi) = 0$, then $\pi$ is the identity, with Coxeter length zero.

Assume now that $\ell(\pi) > 0$. Now let $j \in \{1, 2, \ldots, n\}$ such that $\pi(j) = n$. Define

$$\tau = \pi s_{j_n} s_{j_{n+1}} \cdots s_{n-1}, \tag{2}$$

such that $\tau(n) = n$. Then using (1)

$$\begin{aligned} \ell(\tau) &= \ell(\pi s_{j_n} s_{j_{n+1}} \cdots s_{n-1}) \\ &= \ell(\pi s_{j_n} s_{j_{n+1}} \cdots s_{j_{n-2}}) - 1 \\ &= \ell(\pi s_{j_n} s_{j_{n+1}} \cdots s_{j_{n-3}}) - 2 \\ &\ \ \vdots \\ &= \ell(\pi) - (n - j). \end{aligned}$$

Hence, $\ell(\tau) < \ell(\pi)$, so by induction, $\ell(\tau)$ is the Coxeter length of $\tau$, so $\tau$ can be written as the product of $\ell(\tau)$-many Coxeter generators. (2) now shows us that $\pi$ can be written as the product of $\ell(\tau) + (n - j) = \ell(\pi)$ many Coxeter generators. Hence, the Coxeter length of $\pi$ is bounded above by $\ell(\pi)$. □

**Definition 1.83.** $\lambda = (\lambda_1, \ldots, \lambda_k) \vdash n$.

$T^\lambda =$

| | | | | |
|---|---|---|---|---|
| $1$ | $2$ | | | $\lambda_1$ |
| $\lambda_1 + 1$ | $\lambda_1 + 2$ | | $\lambda_1 + \lambda_2$ | |
| | | | | |
| $\overline{\lambda}_{k-1}$ | | $n$ | | |

where $\overline{\lambda}_k = \lambda_1 + \ldots + \lambda_{k-1} + 1$. $T^\lambda$ is called the **canonical tableau**.

Given $T \in \mathrm{SYT}(\lambda)$, denote by $\sigma_T \in S_n$ the unique permutation such that $\sigma_T T = T^\lambda$.

**Proposition 1.84.** $T \in \mathrm{SYT}(\lambda)$, $\ell = \ell(\sigma_T)$. Then there is a sequence of $\ell$ admissible transpositions which transforms $T$ into $T^\lambda$.

*Proof.* Let $j$ be the number in the rightmost box of the last row of $T$. There are two cases: either $j = n$ or $j \neq n$.

If $j = n$ then, as this box is removable, we can consider standard tableau $\widehat{T}$ of shape $\widehat{\lambda} = (\lambda_1, \ldots, \lambda_k - 1) \vdash (n-1)$ obtained by removing that box. By induction applied to $\widehat{T}$, there exists a sequence of $\widehat{\ell} = \ell(\sigma_{\widehat{T}})$ admissible transpositions which transforms $\widehat{T}$ into $T^{\widehat{\lambda}}$ defined by $\widehat{\lambda}$. The same sequence will transform $T$ into $T^\lambda$ and $\ell = \widehat{\ell}$.

If $j \neq n$, then $s_j$ is an admissible transposition for $T$. Similarly, $s_{j+1}$ is admissible for $s_j T$, and so on, so $s_{n-1}$ is admissible for $s_{n-2}s_{n-3}\cdots s_{j+1}s_j T$. Finally, $s_{n-1}s_{n-2}\cdots s_j T$ has $n$ in the rightmost box on the last row of $T$ and we've reduced to the previous case. □

**Corollary 1.85.** If $T, S \in \mathrm{SYT}(\lambda)$, then $S$ can be obtained from $T$ by applying a sequence of admissible adjacent transpositions.

# 2 The Okounkov-Vershik approach

This approach to the representation theory of $S_n$ is due to Andrei Okounkov and Anatoly Vershik, 1996, 2005.

## 2.1 Main Steps in the Okounkov-Vershik approach

- branching $S_n \to S_{n-1}$ is multiplicity free, see Lemma 1.59
- given an irreducible $S_n$-module $V_\lambda$, branching is simple so the decomposition of $V_\lambda$ into irreducible $S_{n-1}$-modules depends only on the given partition and nothing else. Each module decomposes canonically into irreducible $S_{n-2}$-modules. Iterating we get a canonical decomposition of $V_\lambda$ into irreducible $S_1$-modules. So there exists a canonical basis of $V_\lambda$ determined modulo scalars, called the **Gelfand-Tsetlin basis** (abbreviated GZ-basis).
- Let $Z_n = Z(\mathbb{C}S_n)$ be the center of the group algebra of $S_n$. The **Gelfand-Tsetlin algebra** $GZ_n$ is a (commutative) subalgebra of $\mathbb{C}S_n$ generated by $Z_1 \cup \ldots \cup Z_n$.
- The next step in the Vershik-Okounkov approach is to show that $GZ_n$ consists of all elements of $\mathbb{C}S_n$ that act diagonally in the $GZ$-basis in every irreducible representation. $GZ_n$ is a maximal commutative subalgebra of $\mathbb{C}S_n$ with dimension equal to the sum of dimensions of the distinct irreducible $S_n$-modules. Thus, any vector in the $GZ$-basis (in any irrep) is uniquely determined by the eigenvalues of the elements of the $GZ$-algebra on this vector.
- For $i = 1, \ldots, n$, let $X_i = (1, i) + (2, i) + \ldots + (i-1, i) \in \mathbb{C}S_n$. These are called the **Young-Jucys-Murphy elements** (YJM-elements). We will show that these YJM elements generate the GZ algebra.

- To a GZ-vector $v$ (meaning an element of the GZ-basis for some irrep), we associate a tuple $\alpha(v) = (a_1, a_2, \ldots, a_n)$, where $a_i$ is the eigenvalue of $X_i$ on the vector $v$, and let $\mathrm{Spec}(n) = \{\alpha(v) \mid v \text{ is a GZ-vector}\}$.
- By a previous step, for GZ-vectors $u, v$, we have that $u = v$ if and only if $\alpha(u) = \alpha(v)$. Hence, $|\mathrm{Spec}(n)|$ is equal to the sum of dimensions of the distinct irreducible inequivalent representations of $S_n$.
- The last step in the Vershik-Okounkov approach is to construct a bijection $\mathrm{Spec}(n) \to \mathrm{SYT}(n)$ such that tuples in $\mathrm{Spec}(n)$ whose GZ-vectors belong to the same irrep go to standard Young tableaux of the same shape. This proceeds by induction, using relations

$$s_i^2 = 1, \qquad X_i X_{i+1} = X_{i+1} X_i, \qquad s_i X_i + 1 = X_{i+1} s_i$$

  where $s_i = (i, i+1)$.

**Definition 2.1.** A **part** of a partition $\mu = (\mu_1, \ldots, \mu_k)$ is $\mu_i$ for some $i$.

Let $P_1(n)$ be the set of all pairs $(\mu, i)$ where $\mu \vdash n$ is a partition and $i$ is a part of $\mu$.

A part $\mu_i$ of $\mu$ is **non-trivial** if $\mu_i \geqslant 2$. Let $\#\mu$ be the sum of its nontrivial parts.

Recall by Theorem 1.82 that $\sigma \in S_n$ can be written as a product of $\ell(\sigma)$-many Coexter transpositions and cannot be written as a product of any fewer.

**Remark 2.2** (Conventions)**.**

- All algebras are finite dimensional over $\mathbb{C}$ and unital.
- Subalgebras contain the unit; algebra homomorphisms preserve the unit.
- Given elements or subalgebras $A_1, \ldots, A_n$ of an algebra $A$, denote by $\langle A_1, \ldots, A_n \rangle$ the subalgebra of $A$ generated by $A_1 \cup \ldots \cup A_n$.

**Definition 2.3.** Let $G$ be a group. Let $\{1\} = G_1 \leqslant G_2 \leqslant \ldots \leqslant G_n \leqslant \ldots$ be an (inductive) chain of (finite) subgroups of $G$. Write $G_n^\wedge$ for the set of equivalence classes of finite dimensional complex representations of $G_n$.

$V_\lambda$ is the irreducible $G_n$-module corresponding to $\lambda \in G_n^\wedge$.

**Definition 2.4.** The **branching multigraph / Bratelli diagram** of a chain of groups $\{1\} = G_1 \leqslant G_2 \leqslant \ldots \leqslant G_n \leqslant \ldots$ has

- vertices: the elements of the set $\coprod_{n \geqslant 1} G_n^\wedge$
- edges: two vertices $\lambda, \mu$ are joined by $k$ directed edges from $\lambda$ to $\mu$ whenever $\mu \in G_{n-1}^\wedge$ and $\lambda \in G_n^\wedge$ for some $n$, and the multiplicity of $\mu$ in the restriction of $\lambda$ to $G_{n-1}$ is $k$.

We call $G_n^\wedge$ as the $n$**-th level** of the Bratelli diagram. Write $\lambda \to \mu$ if $(\lambda, \mu)$ is an edge of the diagram.

Assume that the Bratelli diagram is a graph, i.e. all multiplicities of all restrictions are 0 or 1 (**multiplicity free** or **simple branching**).

Take a $G_n$-module $V_\lambda$ with $\lambda \in G_n^\wedge$. By simple branching, the decomposition $V_\lambda = \bigoplus_\mu V_\mu$ where the sum over all $\mu \in G_{n-1}^\wedge$ with $\lambda \to \mu$ is *canonical*. Iterating the decomposition, we obtain a canonical decomposition of $V_\lambda$ into irreducible $G_1$-modules, that is, 1-dimensional subspaces

$$V_\lambda = \bigoplus_T V_T, \tag{3}$$

where the sum is over all possible chains

$$\lambda^{(n)} \to \cdots \to \lambda^{(1)} \tag{4}$$

with $\lambda^{(n)} \in G_i^\wedge$ and $\lambda^{(n)} = \lambda$, (or equivalently, the sum is over all possible tableaux of shape $\lambda$ when $G = S_n$).

**Definition 2.5.** Choosing a nonzero vector $v_T$ in each one-dimensional space $V_T$, we obtain a basis $\{v_T\}$ of $V_\lambda$, called the **Gelfand-Tsetlin basis** (GZ-basis).

By definition of $v_T$, we have $(\mathbb{C}G_i)v_T = V_{\lambda^{(i)}}$ for each $i$, because $v_T \in V_{\lambda^{(i)}}$, which is an irrep, and so the action of $\mathbb{C}G_i$ on $v_T$ can recover the entirety of the irrep $V_{\lambda^{(i)}}$.

Note that the chains (4) are in bijection with directed paths in the Bratelli diagram from $\lambda$ to the unique element $\lambda^{(1)}$ of $G_1^\wedge$.

We have a canonical basis (up to scalars) for $V_\lambda$: the GZ-basis. Can we identify those elements of $\mathbb{C}G_n$ that act diagonally in this basis, for every irrep? In other words, consider the algebra isomorphism

$$\begin{array}{rcl} \phi\colon \mathbb{C}G_n & \xrightarrow{\cong} & \bigoplus_{\lambda \in G_n^\wedge} \operatorname{End} V_\lambda \\ g & \longmapsto & \left(V_\lambda \xrightarrow{g} V_\lambda : \lambda \in G_n^\wedge\right) \end{array} \tag{5}$$

We know that $\phi$ is an isomorphism because, if $\phi(x) = \phi(y)$, then $x$ and $y$ necessarily act the same on each irrep and hence on the regular representation. Therefore, $x = y$. Counting dimensions establishes surjectivity.

**Definition 2.6.** Let $D(V_\lambda)$ be the set of operators on $V_\lambda$ diagonal in the GZ-basis of $V_\lambda$.

What is the image under $\phi^{-1}$ (5) of the subalgebra $\bigoplus_{\lambda \in G_n^\wedge} D(V_\lambda)$ of $\bigoplus_{\lambda \in G_n^\wedge} \operatorname{End}(V_\lambda)$?

**Definition 2.7** (Notation)**.** Let $Z_n = Z(\mathbb{C}G_n)$ be the center of the group algebra.

**Definition 2.8.** We can easily see that $GZ_n = \langle Z_1, \ldots, Z_n \rangle$ is a commutative $\mathbb{C}$-algebra of $\mathbb{C}G_n$. This is the **Gelfand Tsetlin algebra** of the inductive chain of subgroups.

**Theorem 2.9.** $GZ_n$ is the image of $\bigoplus D(V_\lambda)$ under the isomorphism (5). That is, $GZ_n$ consists of elements of $\mathbb{C}G_n$ that act diagonally in the GZ-basis in every irreducible representation of $G_n$. Thus, $GZ_n$ is a maximal commutative subalgebra of $\mathbb{C}G_n$ and its dimension is

$$\dim GZ_n = \sum_{\lambda \in G_n^\wedge} \dim V_\lambda$$

*Proof.* Consider the chain $T$ from (4). For each $i = 1, \dots, n$, we will denote by $p_{\lambda^{(i)}} \in Z_i$ the central idempotent corresponding to the representation defined by $\lambda^{(i)} \in G_i^\wedge$. Define

$$p_T = p_{\lambda^{(1)}} p_{\lambda^{(2)}} \cdots p_{\lambda^{(n)}} \in GZ_n.$$

The image of $p_T$ under (5) is $(f_\mu \colon \mu \in G_n^\wedge)$ where $f_\mu = 0$ if $\mu \neq \lambda$ and $f_\lambda$ is projection onto $V_T$ with respect to (3).

Hence, the image of $GZ_n$ under (5) includes $\bigoplus_{\lambda \in G_n^\wedge} D(V_\lambda)$, which is a commutative maximal subalgebra of $\bigoplus_{\lambda \in G_n^\wedge} \operatorname{End}(V_\lambda)$. Since $GZ_n$ is itself commutative, the result follows. □

**Definition 2.10.** A **GZ-vector** of $G_n$ (modulo scalars) is an element $v_T$ of the GZ-basis for some irrep of $G_n$.

An immediate corollary of Theorem 2.9 says something interesting about these vectors.

**Corollary 2.11** (Corollary of Theorem 2.9)**.**

(i) Let $v \in V_\lambda$, $\lambda \in G_n^\wedge$. If $v$ is an eigenvector (for the action) of every element of $GZ_n$ then (a scalar multiple of) $v$ is a GZ-basis vector of $V_\lambda$, that is, $v$ is of the form $v_T$ for some path $T$.

(ii) Let $u, v$ be two GZ-vectors. If $u, v$ have the same eigenvalues for every element of $GZ_n$, then $u = v$.

**Remark 2.12.** Later we'll find an explicit set of generators for the GZ-algebras of the symmetric groups.

## 2.2 Symmetric groups have simple branching

To prove this, we first need several preliminary theorems from the theory of semisimple algebras. We won't prove them, or even use them often.

**Theorem 2.13** (Artin-Wedderburn Theorem)**.** If $A$ is a semisimple $\mathbb{C}$-algebra, then $A$ decomposes as a direct sum of matrix algebras over $\mathbb{C}$.

**Theorem 2.14** (Double Centralizer)**.** Let $A$ be a finite dimensional central simple algebra (**central** means $Z(A) = \mathbb{C}$), and $B \subseteq A$ a simple subalgebra. Let $C = Z_A(B)$ be the centralizer. Then $C$ is simple, and $Z_A(C) = B$, and moreover $\dim_{\mathbb{C}}(A) = \dim_{\mathbb{C}}(B) \cdot \dim_{\mathbb{C}}(C)$.

**Theorem 2.15.** Let $M$ be a finite dimensional semisimple complex algebra and let $N$ be a semisimple subalgebra. Let $Z(M,N) = Z_M(N)$ be the centralizer of the pair $(M,N)$,

$$Z(M,N) = \{m \in M \mid mn = nm \;\forall\, n \in N\}.$$

Then $Z(M,N)$ is semisimple and the following are equivalent:

(a) the restriction of any finite dimensional complex irreducible representation of $M$ to $N$ is multiplicity free

(b) $Z(M,N)$ is commutative.

We'll apply this theorem in the case where $M = \mathbb{C}S_n$ and $N = \mathbb{C}S_{n-1}$.

*Proof.* Without loss of generality (by Artin-Wedderburn, Theorem 2.13) $M = \bigoplus_{i=1}^{t} M_i$ where each $M_i$ is some matrix algebra. Write elements of $M$ as tuples $(m_1, \ldots, m_k)$ where each $m_i \in M_i$. Let $N_i$ be the image of $N$ under the projection $M \to M_i$. It's a homomorphic image of a semisimple algebra, and so $N_i$ is semisimple.

Now $Z(M,N) = \bigoplus_{i=1}^{k} Z(M_i, N_i)$. By the Double Centralizer Theorem (Theorem 2.14), each $Z(M_i, N_i)$ is simple, and therefore $Z(M,N)$ is semisimple.

Now let's establish the equivalence of (a) and (b). Let

$$V_i = \left\{ (m_1, \ldots, m_k) \in M \;\middle|\; \begin{array}{l} m_j = 0 \text{ for } i \neq j \text{ and with all entries of} \\ m_i \text{ not in the first column equal to zero} \end{array} \right\}.$$

The $V_i$ are all the distinct inequivalent irreducible $M$ modules and the decomposition of $V_i$ into irreducible $N$-modules is identical to the decomposition of $V_i$ into irreducible $N_i$-modules.

Now notice that $Z(M,N)$ is commutative if and only if $Z(M_i, N_i)$ is commutative for all $i$, which is true if and only if all irreducible representations of $Z(M_i, N_i)$ have dimension 1. So it suffices to show that all irreps of $Z(M_i, N_i)$ are have dimension 1 if and only if the restriction of irreps of $M_i$ to $N_i$ is multiplicity free.

First, assume that all irreps of $Z(M_i, N_i)$ have dimension 1. Let $U$ be an irrep of $M_i$ and $V$ an irrep of $N_i$. In particular, $\operatorname{Hom}_{N_i}(V, U)$ is an irrep of $Z(M_i, N_i)$, and so has dimension 1. By Schur's lemma, this is the multiplicity of $V$ in $\operatorname{Res}_{N_i}^{M_i} U$, so the branching is multiplicity-free.

Conversely, assume $Z(M_i, N_i)$ has an irrep of dimension $> 1$. Let $U$ be an irrep of $M_i$ and $V$ an irrep of $N_i$. We have by Schur's Lemma that $\operatorname{End}_{\mathbb{C}}(U) \cong M$, so $\operatorname{End}_{N_i}(U) \cong Z(M_i, N_i)$. Hence, if $\operatorname{Res}_{N_i}^{M_i} U = \bigoplus_j W_j$ is the decomposition of $U$ into simple $N_i$-modules, then

$$Z(M_i, N_i) \cong \operatorname{End}_{N_i}(U) \cong \bigoplus_{j,k} \operatorname{Hom}_{N_i}(W_j, W_k)$$

is a decomposition of $Z(M_i, N_i)$ into irreducible representations. So if there is an irrep of $Z(M_i, N_i)$ with dimension $> 1$, then $W_j \cong W_k$ for some distinct $j, k$. This irrep of $N_i$ then occurs with multiplicity $> 1$ in $\operatorname{Res}_{N_i}^{M_i} U$; branching is not simple. □

## 2.3 Involutive algebras

**Definition 2.16.** Let $F = \mathbb{R}$ or $F = \mathbb{C}$. If $F = \mathbb{C}$, then for $\alpha \in \mathbb{C}$ denote by $\overline{\alpha}$ the complex conjugate. If $F = \mathbb{R}$ and $\alpha \in \mathbb{R}$, then $\overline{\alpha} = \alpha$.

An $F$-algebra $A$ is **involutive** if it has a conjugate linear anti-automorphism of order 2, that is, a bijective map $x \mapsto x^*$ such that

$$\begin{aligned}(x+y)^* &= x^* + y^* \\ (\alpha x)^* &= \overline{\alpha} x^* \\ (xy)^* &= y^* x^* \\ (x^*)^* &= x\end{aligned}$$

for all $x, y \in A$, $\alpha \in F$. $x^*$ is called the **adjoint** of $x$. An element $x \in A$ is **normal** if it commutes with it's own adjoint: $xx^* = x^*x$. $x$ is **self adjoint** or **hermitian** if $x = x^*$.

**Definition 2.17.** Let $A$ be involutive over $\mathbb{R}$. Then the **$*$-complexification** of $A$ is the algebra whose elements are pairs $(x, y) \in A \times A$, written $x + iy$, and the operations are the obvious ones:

$$\begin{aligned}(x_1 + iy_1) + (x_2 + iy_2) &= (x_1 + x_2) + i(y_1 + y_2) \\ \alpha(x + iy) &= (\alpha x) + i(\alpha y) \\ (x_1 + iy_1)(x_2 + iy_2) &= (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1) \\ (x + iy)^* &= x^* - iy^*\end{aligned}$$

A **real element** of the $*$-complexification is an element of the form $x + i0$ for some $x \in A$.

**Example 2.18.** If $F = \mathbb{R}$ or $\mathbb{C}$, and $G$ is a finite group, $FG$ is involutive under

$$\left(\sum_i \alpha_i g_i\right)^* = \sum_i \overline{\alpha}_i g_i^{-1}. \tag{6}$$

Recall that we defined $Z_n = Z(\mathbb{C}G_n)$ and the **Gelfand-Tsetlin algebra** $GZ_n = \langle Z_1, \ldots, Z_n \rangle$. Recall also that

$$\mathbb{C}G_n = \bigoplus_{\lambda \in G_n^\wedge} \operatorname{End}(V_\lambda).$$

**Remark 2.19.** Here is what we know about the Gelfand-Tsetlin algebra so far, from Theorem 2.9

(0) $GZ_n$ is commutative

(1) $GZ_n$ is the algebra of diagonal matrices with respect to the GZ-basis

(2) $GZ_n$ is a maximal commutative subalgera of $\mathbb{C}G_n$

(3) $v \in V_\lambda$ is in the GZ-basis if and only if $v$ is a common eigenvector of elements of $GZ_n$

(4) Each basis element is uniquely determined by the eigenvalues of elements of $GZ_n$.

If $A \geqslant B$, then $Z(A, B) = \{a \in A \mid ab = ba \; \forall \, b \in B\}$.

**Remark 2.20.** We proved the following in Theorem 2.15. Let $H$ be a subgroup of $G$. Then the following are equivalent:

(1) $\mathrm{Res}_H^G$ is multiplicity-free;

(2) $Z(\mathbb{C}G, \mathbb{C}H)$ is commutative.

Here's some more stuff about involutive algebras.

**Theorem 2.21.** Let $A$ be an involutive $\mathbb{C}$-algebra. Then

(i) An element $x \in A$ is normal if and only if $x = y + iz$ for some self-adjoint $y, z \in A$ that commute.

(ii) $A$ is commutative if and only if every element of $A$ is normal.

(iii) If $A$ is a $*$-complexification of a real involutive algebra, then $A$ is commutative if every real element of $A$ is self-adjoint.

*Proof. Proof of (i).* Assume that $x$ is normal. Then $xx^* = x^*x$. Define $y = \frac{1}{2}(x + x^*)$ and $z = \frac{i}{2}(x^* - x)$. Then we have that

$$y^* = \tfrac{1}{2}(x + x^*)^* = \tfrac{1}{2}(x^* + x) = y$$

$$z^* = (\tfrac{i}{2}(x^* - x))^* = -\tfrac{i}{2}(x^* - x)^* = -\tfrac{i}{2}(x - x^*) = \tfrac{1}{2}(ix^* - ix) = z$$

$$yz = \tfrac{i}{4}(x + x^*)(x^* - x) = \tfrac{i}{4}(xx^* - x^2 + (x^*)^2 + x^*x) = \tfrac{i}{4}((x^*)^2 - x^2)$$

$$zy = \tfrac{i}{4}(x^* - x)(x + x^*) = \tfrac{i}{4}(x^*x + (x^*)^2 - x^2 - xx^*) = \tfrac{i}{4}((x^*)^2 - x^2)$$

$$y + iz = \tfrac{1}{2}(x + x^*) + \tfrac{i^2}{2}(x^* - x) = \tfrac{1}{2}(x + x^* - x^* + x) = x$$

Conversely, if $x = y + iz$ for self-adjoint $y, z \in A$ that commute, then

$$xx^* = (y + iz)(y + iz)^* = (y + iz)(y^* - iz^*) = (y + iz)(y - iz) = y^2 + z^2$$

$$x^*x = (y + iz)^*(y + iz) = (y^* - iz^*)(y + iz) = (y - iz)(y + iz) = y^2 + z^2$$

□

*Proof of (ii).* Let $x, y \in A$. By part $(i)$, write $x = x_1 + ix_2$ for some $x_1, x_2$ that are self-adjoint and commute. Likewise, write $y = y_1 + iy_2$ for some $y_1, y_2$ that are self-adjoint and commute. Note that, since $x_1^* = x_1$ and $y_1 = y_1^*$, and $x_1y_1 \in A$ is normal, then

$$(x_1y_1)^* = x_1y_1 \implies y_1^*x_1^* = x_1y_1 \implies y_1x_1 = x_1y_1,$$

and likewise $x_1$ commutes with $y_2$ and $y_1$ commutes with $x_2$ and $y_2, x_2$ commute. Therefore,

$$\begin{aligned} xy &= (x_1 + ix_2)(y_1 + iy_2) \\ &= x_1 y_1 + ix_1 y_2 + ix_2 y_1 - x_2 y_2 \\ &= y_1 x_1 + iy_2 x_1 + iy_1 x_2 - y_2 x_2 \\ &= (y_1 + iy_2)(x_1 + ix_2) \\ &= yx \end{aligned}$$

□

*Proof of (iii).* Every element of $A$ can be written as $x + iy$ for some real elements $x, y$. It suffices by part (ii) to show that every element of $A$ is normal, and to show that, it suffices to show by part (i) that every element of $A$ is of the form $x + iy$ for $x, y$ self-adjoint and commuting. We know that $x, y$ are self adjoint because they are real elements. We know further that $xy$ is a real element, and therefore $xy = (xy)^* = y^* x^* = yx$. Hence, $x$ and $y$ are self-adjoint and commute, so we are done. □

This concludes the proof of Theorem 2.15. □

**Theorem 2.22.** The centralizer $Z(\mathbb{C}S_n, \mathbb{C}S_{n-1})$ is commutative.

*Proof.* Claim that the involutive subalgebra $Z(\mathbb{C}S_n, \mathbb{C}S_{n-1})$ is the $*$-complexification of $Z(\mathbb{R}S_n, \mathbb{R}S_{n-1})$, and therefore commutative by Theorem 2.21. So it's enough to show that every element of $Z(\mathbb{R}S_n, \mathbb{R}S_{n-1})$ is self-adjoint by Theorem 2.21.

To that end, let $f = \sum_{\pi \in S_n} \alpha_\pi \pi$ for $\alpha_\pi \in \mathbb{R}$ be an element of $Z(\mathbb{R}S_n, \mathbb{R}S_{n-1})$. Fix $\sigma \in S_n$. $S_n$ is **ambivalent** (meaning that every element is conjugate to its inverse) since $\sigma$ and $\sigma^{-1}$ are of the same cycle type.

To produce a permutation $\tau$ in $S_n$ conjugating $\sigma$ to $\sigma^{-1}$: write the permutation $\sigma$ in cycle form and write down $\sigma^{-1}$ in cycle form below such that the lengths correspond to each other. The permutation in $S_n$ taking an element of the top row to the corresponding element in the bottom row cojugates $\sigma$ to $\sigma^{-1}$. For example, in $S_9$, write

$$\begin{aligned} \sigma &= (124)(35)(6879) \\ \sigma^{-1} &= (142)(35)(6978) \end{aligned}$$

and then $\tau = (24)(89)$.

Moreover, we want this $\tau$ to represent an element of $S_{n-1}$. We can always choose a conjugating $\tau$ that fixes any of the numbers that $\sigma$ moves, that is, there is some $\tau$ such that $\tau(n) = n$ and $\tau\sigma\tau^{-1} = \sigma^{-1}$. We can do this by permuting the cycle of $\sigma^{-1}$ containing $n$ until $n$ lines up in both the cycles of $\sigma$ and $\sigma^{-1}$. Continuing the previous example, we can write $(6978) = (7869)$, so

$$\begin{aligned} \sigma &= (124)(35)(6879) \\ \sigma^{-1} &= (142)(35)(7869) \end{aligned}$$

and then $\tau = (24)(67)$ conjugates $\sigma$ to $\sigma^{-1}$.

Therefore, we can choose $\tau \in S_{n-1}$ such that $\tau\sigma\tau^{-1} = \sigma^{-1}$.

Since $\tau \in S_{n-1}$ and $f \in Z(\mathbb{R}S_n, \mathbb{R}S_{n-1})$, we see that $\tau f = f\tau \implies f = \tau f \tau^{-1}$. Hence,

$$f = \tau f \tau^{-1} = \sum_{\pi \in S_n} \alpha_\pi (\tau\pi\tau^{-1}),$$

so $\alpha_\sigma$ are constant on conjugacy classes. Therefore, $\alpha_\sigma = \alpha_{\sigma^{-1}}$. Since $\sigma$ was arbitrary in $S_n$, we see that $f^* = f$ (see (6)). ☐

Denote the centralizer $Z(\mathbb{C}S_n, \mathbb{C}S_{n-1})$ by $Z_{(n-1,1)}$. We'll see a second proof of commutativity in Theorem 2.27.

## 2.4 Young-Jucys-Murphy elements (YJM elements)

Henceforth $G_n = S_n$, so chains in the Bratelli diagram refer to chains in the Bratelli diagram of the symmetric groups.

**Definition 2.23.** For $i = 2, \ldots, n$, define $Y_i$ to be the sum of all $i$-cycles in $S_{n-1}$. By convention, $Y_n = 0$. Define $Y_i'$ as the sum of all $i$-cycles in $S_n$ containing $n$.

For $(\mu, i) \in P_1(n)$, (see Definition 2.1), let $c_{(\mu,i)} \in \mathbb{C}S_n$ be the sum of permutations $\pi$ in $S_n$ such that the type of $\pi$ is $\mu$ and the size of the cycle of $\pi$ containing $n$ is $i$.

**Remark 2.24.** Each of $Y_2, \ldots, Y_{n-1}, Y_2', \ldots, Y_n'$ equals $c_{(\mu,i)}$ for suitable $\mu$ and $i$. In particular, $Y_j = c_{(\mu,1)}$ for $\mu = (j, 1, \ldots, 1)$, and $Y_j' = c_{(\mu,j)}$ for $\mu = (j, 1, \ldots, 1)$.

**Lemma 2.25.**

(i) $\{c_{(\mu,i)} \mid (\mu, i) \in P_1(n)\}$ is a basis of $Z_{(n-1,1)}$. Therefore, we have that $\langle Y_2, \ldots, Y_{n-1}, Y_2', \ldots, Y_n' \rangle \subseteq Z_{(n-1,1)}$.

(ii) $c_{(\mu,i)} \in \langle Y_2, \ldots, Y_k, Y_2', \ldots, Y_k' \rangle$ for $k = \#\mu$.

(iii) $Z_{(n-1,1)} = \langle Y_2, \ldots, Y_{n-1}, Y_2', \ldots, Y_n' \rangle$

(iv) $Z_{n-1} = \langle Y_2, \ldots, Y_{n-1} \rangle$.

*Proof.* (i) The first bit is an exercise, similar to the proof Proposition 1.23 that $\{c_\mu \mid \mu \in P(n)\}$ is a basis of $Z_n$. The second bit follows from Remark 2.24.

(ii) Induction on $\#\mu$. If $\#\mu = 0$, then $c_{(\mu,i)}$ is the identity permutation, which lies in the subalgebra $\langle Y_2, \ldots, Y_k, Y_2', \ldots, Y_k' \rangle$. So now assume true when $\#\mu \leqslant k$. Consider $(\mu, i) \in P_1(n)$ with $\#\mu = k + 1$. Let the nontrivial parts of $\mu$ be $\mu_1, \ldots, \mu_\ell$ in some order.

There are several cases

(a) First, prove it for $i = 1$. Consider the product $Y_{\mu_1} \cdots Y_{\mu_\ell}$ By (i),

$$Y_{\mu_1} \cdots Y_{\mu_\ell} = \alpha_{(\mu,1)} c_{(\mu,1)} + \sum_{(\tau,1)} \alpha_{(\tau,1)} c_{(\tau,1)}.$$

where $\alpha_{(\mu,1)} \neq 0$ and the sum is over all $(\tau, 1)$ with $\#\tau < \#\mu$. Then we are done by induction.

(b) If $i > 1$, without loss of generality assume that $i = \mu_1$. Consider the product $Y'_{\mu_1} Y_{\mu_2} \cdots Y_{\mu_\ell}$. By (i), we see that

$$Y'_{\mu_1} Y_{\mu_2} \cdots Y_{\mu_\ell} = \alpha_{(\mu,i)} c_{(\mu,i)} + \sum_{(\tau,j)} \alpha_{(\tau,j)} c_{(\tau,j)}.$$

where $\alpha_{(\mu,i)} \neq 0$ and the sum is over all $(\tau, j)$ with $\#\tau < \#\mu$. Then we are done by induction.

(iii) Follows from $(i)$ and $(ii)$.

(iv) Similar to $(iii)$.

□

**Definition 2.26.** For $1 \leqslant i \leqslant n$, define the **Young-Jucys-Murphy elements** (YJM elements)

$$X_i = (1, i) + (2, i) + \ldots + (i-1, i) \in \mathbb{C}S_n.$$

By convention, $X_1 = 0$.

This is equal to the sum of all the 2-cycles in $S_i$ minus the sum of all 2-cycles in $S_{i-1}$. Note that $X_i$ is the difference of an element of $Z_i$ and an element of $Z_{i-1}$. Therefore $X_i \notin Z_i$ for all $1 \leqslant i \leqslant n$. $X_i$ and $(i, 1)$ don't commute, for example.

**Theorem 2.27** (Okounkov-Vershnik, 2004)**.**

(i) $Z_{(n-1,1)} = \langle Z_{n-1}, X_n \rangle$

(ii) $GZ_n = \langle X_1, \ldots, X_n \rangle$.

*Proof.* (i) Evidently, $\langle Z_{n-1}, X_n \rangle \subseteq Z_{(n-1,1)}$ because $X_n = Y'_2$ and then apply Lemma 2.25(iii).

Conversely, we already know that $Y_k \in Z_{n-1}$, so it's enough to show that $Y'_2, \ldots, Y'_n \in \langle Z_{n-1}, X_n \rangle$. Since $Y'_2 = X_n$, then $Y'_2 \in \langle Z_{n-1}, X_n \rangle$. This forms the base case for induction.

Now assume that $Y'_2, \ldots, Y'_{k+1} \in \langle Z_{n-1}, X_n \rangle$. We aim to show for an inductive step that $Y'_{k+2} \in \langle Z_{n-1}, X_n \rangle$. We'll sink to computing with elements and just hit this theorem with a club until it dies. Write $Y'_{k+1}$ as

$$Y'_{k+1} = \sum_{i_1,\ldots,i_k} (i_1, \ldots, i_k, n)$$

summed over all distinct $i_1, \ldots, i_k \in \{1, 2, \ldots, n\}$. Consider now $Y'_{k+1} X_n \in \langle Z_{n-1}, X_n \rangle$.

$$Y'_{k+1} X_n = \left( \sum_{i_1,\ldots,i_k} (i_1, \ldots, i_k, n) \right) \left( \sum_{i=1}^{n-1} (i, n) \right) \tag{7}$$

and take a typical element $(i_1, \dots, i_k, n)(i, n)$ of this product. There are two possibilities: either $i \neq i_j$ for any $j = 1, \dots, k$, or $i = i_j$ for some $j$.

- If $i \neq i_j$ for any $j = 1, \dots, k$, then the product is $(i, i_1, \dots, i_k, n)$.
- If $i = i_j$ for some $j$, then the product is $(i_1, \dots, i_j)(i_{j+1}, \dots, n)$.

Hence, (7) becomes

$$\sum_{i,i_1,\dots,i_k} (i, i_1, \dots, i_k, n) + \sum_{i_1,\dots,i_k} \sum_{j=1}^{k} (i_1, \dots, i_j)(i_{j+1}, \dots, i_k, n) \tag{8}$$

The first sum is over all distinct $i, i_1, \dots, i_k \in \{1, \dots, n-1\}$ and the sum on the right is over all distinct $i_1, \dots, i_k \in \{1, 2, \dots, n-1\}$.

Rewrite (8) as

$$Y'_{k+2} + \sum_{(\mu,i)} \alpha_{(\mu,i)} c_{(\mu,i)}$$

where the sum is over all $(\mu, i)$ such that $\#\mu \leqslant k+1$. Now by induction and Lemma 2.25(ii), we have $Y'_{k+2} \in \langle Z_{n-1}, X_n \rangle$.

(ii) Induction on $n$. The cases for $n = 1$ and $n = 2$ are trivial.

Now assume $GZ_{n-1}$ is generated by $\langle X_1, X_2, \dots, X_{n-1} \rangle$. We want to show that $GZ_n = \langle GZ_{n-1}, X_n \rangle$. Clearly, $GZ_n \supseteq \langle GZ_{n-1}, X_n \rangle$, since $X_n$ is the difference of an element of $Z_n$ and an element of $Z_{n-1}$. So we just have to check that $GZ_n \subseteq \langle GZ_{n-1}, X_n \rangle$.

To show this, it's enough to show that $Z_n \subseteq \langle GZ_{n-1}, X_n \rangle$. But this is clear by (i), since $Z_n \subseteq Z_{(n-1,1)} \subseteq \langle Z_{n-1}, X_n \rangle \subseteq \langle GZ_{n-1}, X_n \rangle$.

□

**Remark 2.28.** Theorem 2.27(i) implies that $Z_{(n-1,1)}$ is commutative, because $Z_{(n-1,1)} = \langle Z_{n-1}, X_n \rangle$ and $X_n$ commutes with every element in $Z_{n-1}$. This gives another proof of the fact that $Z_{(n-1,1)}$ is commutative.

**Definition 2.29.** The GZ-basis for $G = S_n$ is called the **Young basis**. By Corollary 2.11(i), the **Young / GZ-vectors** are common eigenvectors for $GZ_n$.

**Definition 2.30.** Let $v$ be a Young vector for $S_n$. $\alpha(v) = (a_1, \dots, a_n) \in \mathbb{C}^n$, where $a_i$ is the eigenvalue of $X_i$ on $v$. Call $\alpha(v)$ the **weight** of $v$. Note that $a_1 = 0$ since $X_1 = 0$.

**Definition 2.31.** Let $\mathrm{Spec}(n) = \{\alpha(v) \mid v \text{ is a Young vector}\}$. This is the **spectrum** of YJM-elements.

By Corollary 2.11(ii),

$$|\mathrm{Spec}(n)| = \dim GZ_n = \sum_{\lambda \in S_n^\wedge} \dim \lambda.$$

By definition, Spec$(n)$ is in natural bijection with chains $T$, as in (4). Given $\alpha \in \text{Spec}(n)$, denote by $v_\alpha$ the Young vector with weight $\alpha$ and $T_\alpha$ the corresponding chain in the Bratelli diagram.

Given a chain $T$ as in (4), we denote the corresponding weight vector $\alpha(v_T)$ by $\alpha(T)$. Hence, we have a one-to-one correspondence $T \mapsto \alpha(T)$; $\alpha \mapsto T_\alpha$ between chains in the Bratelli diagram and Spec$(n)$.

Moreover, there is a natural equivalence relation $\sim$ on Spec$(n)$ defined as follows.

**Definition 2.32.** Let $\alpha, \beta \in \text{Spec}(n)$, $\alpha \sim \beta \iff v_\alpha, v_\beta$ belong to the same irreducible module for $S_n \iff T_\alpha$, $T_\beta$ start at the same vertex.

Clearly, #(Spec$(n)/\sim$) is the number of paths in the Bratelli diagram from level $n$ to level 0, which (although we haven't proved it yet) is equal to $\#\,\text{SYT}(\lambda) = \#\hat{S_n}$. So this gives some circumstantial evidence that looking at Spec$(n)$ is interesting and relevant to representations of $S_n$.

We want to

- describe the set Spec$(n)$
- describe the relation $\sim$
- calculate the matrix elements in the Young basis
- calculate the characters of irreducible representations of $S_n$

**Remark 2.33.** The book by Curtis-Reiner from 1962 is a good reference for Artin-Wedderburn theory.

Here's the story so far:

- each irrep $V_\lambda$ has a "nice" basis called the GZ-basis $\{v_T\}$, each $v_T$ corresponding to some chain $\lambda^{(n)} \to \lambda^{(n-1)} \to \cdots \to \lambda^{(1)} = (1)$.
- The YJM elements are $X_k = \sum_{i=1}^{k-1}(i,k) \in \mathbb{C}S_n$ for $1 \leqslant k \leqslant n$.
- The GZ-algebra $GZ_n$ is generated by the YJM elements.
- $GZ_n$ is a maximal commutative subalgebra of $\mathbb{C}S_n$ by Theorem 2.27.
- The GZ-basis is the unique basis such that the basis elements are common eigenvectors of the $X_k$. $X_i \cdot v_T = a_i v_T$. Note that $a_i$ depends on $T$ as well as $i$.
- $\alpha(T) = (a_1, \ldots, a_n) \in \mathbb{C}^n$
- We're looking at the spectrum Spec$(n) = \{\alpha(T) \mid T \text{ is a path}\}$.

## 3 Coxeter generators acting on the Young basis

The Young vectors are a simultaneous eigenbasis for the GZ-algebra. The Coxeter generators $s_i = (i, i+1)$ for $1 \leqslant i \leqslant n-1$ commute with each other except for $s_i s_j$ when $|i-j| < 2$. They act "locally" on the Young basis.

**Lemma 3.1.** Let

$$T\colon \lambda^{(n)} \to \lambda^{(n-1)} \to \cdots \to \lambda^{(1)} \tag{9}$$

be a chain with $\lambda^{(k)} \in S_k^\wedge$, and let $1 \leqslant i \leqslant n-1$. Then $s_i \cdot v_T$ is a linear combination of vectors $v_{T'}$, where $T'$ runs over chains of the form

$$\lambda'^{(n)} \to \lambda'^{(n-1)} \to \cdots \to \lambda'^{(1)}$$

with $\lambda'^{(k)} = \lambda^{(k)}$ for $k \neq i$. The coefficients of the linear combination depend only $\lambda^{(i-1)}, \lambda^{(i)}, \lambda^{(i+1)}$ and on the choice of scalar factors for the vectors in the Young basis, i.e. the action of $s_i$ affects only the $i$-th level and depends only on levels $i-1$, $i$, and $i+1$ of the Bratelli diagram.

*Proof.* For $j \geqslant i+1$, since $s_i \in S_j$ and $\mathbb{C}S_j v_T$ is irreducible, then $s_i v_T \in \mathbb{C}S_j s_i v_T = \mathbb{C}S_j v_T \cong V_{\lambda^{(j)}}$ where $V_{\lambda^{(j)}}$ is the irreducible $S_j$-module indexed by $\lambda^{(j)} \in S_j^\wedge$.

For $j \leqslant i-1$, the action of $s_i$ on $V_{\lambda^{(i+1)}}$ is $S_j$-linear because $s_i$ commutes with all the elements of $S_j$. So $s_i v_T$ belongs to the $V_{\lambda^{(j)}}$-isotypical component of $V_{\lambda^{(i+1)}}$. The first bit now follows and the rest is an exercise. □

Now let's compute an explicit action of $s_i$ on $v_T$ in terms of the weights $\alpha(T)$. Check that

$$s_i X_j = X_j s_i \qquad j \neq i, i+1 \tag{10}$$

$$s_i^2 = 1, \quad X_i X_{i+1} = X_{i+1} X_i, \quad s_i X_i + 1 = X_{i+1} s_i \tag{11}$$

**Exercise 3.2.** Prove Lemma 3.1 using (10).

Given $T$ as in (9), let $\alpha(T) = (a_1, \ldots, a_n)$. Let $V$ be the subspace of $V_{\lambda^{(i+1)}}$ generated by $v_T$ and $s_i v_T$. Note that $\dim V \leqslant 2$. Relations (11) imply $V$ is invariant under the actions of $s_i$,$X_i$,$X_{i+1}$.

**Definition 3.3.** $H(2)$ is the algebra generated by $H_1, H_2, s$ with relations

$$s^2 = 1, \quad H_1 H_2 = H_2 H_1, \quad sH_1 + 1 = H_2 s. \tag{12}$$

**Remark 3.4.** $H_2$ is superfluous in the generating set, because $H_2 = sH_1 s + s$. $H(2)$ is the simplest example of the **degenerate affine Hecke algebras**.

**Definition 3.5.** The **degenerate affine Hecke algebra** $H(n)$ is generated by commuting variables $Y_1, \ldots, Y_n$ and Coxeter involutions $s_1, \ldots, s_{n-1}$ with relations

$$s_i Y_j = Y_j s_i \text{ for } j \neq i, i+1 \qquad s_i Y_i + 1 = Y_{i+1} s_i$$

These were introduced by Drinfeld and Cherednik in 1986. If $Y_1 = 0$, then the quotient of $H(n)$ by the corresponding ideal of relations is canonically isomorphic to $\mathbb{C}S_n$.

**Fact 3.6.** Finite-dimensional $C^*$-algebras are semisimple.

**Lemma 3.7.** (i) All irreducible representations of $H(2)$ are at most 2-dimensional.

(ii) for $i = 1, \ldots, n-1$, the image of $H(2)$ in $\mathbb{C}S_n$ obtained by setting $s = s_i = (i, i+1)$, $H_1 = X_i$, $H_2 = X_{i+1}$ is semisimple, i.e. the subalgebra $M$ of $\mathbb{C}S_n$ generated by $s_i$, $X_i$, $X_{i+1}$ is semisimple.

*Proof.* (i) Let $V$ be an irreducible $H(2)$-module. Since $H_1$, $H_2$ commute, they have a common eigenvector $v$. Let $W = \operatorname{Span}(v, sv)$. Then $\dim W \leqslant 2$ and (12) shows that $W$ is a submodule of $V$. Since $V$ is irreducible, $W = V$.

(ii) Let $\operatorname{Mat}(n)$ be the algebra of $n! \times n!$ complex matrices where the rows and columns of these matrices are indexed by permutations in $S_n$. Consider the (left) regular representation of $S_n$. Then in matrix terms, this embeds $S_n$ into $\operatorname{Mat}(n)$.

The matrix in $\operatorname{Mat}(n)$ corresponding to a transposition $(i, j)$ in $S_n$ is real and symmetric. Since $X_i, X_{i+1}$ are sums of transpositions the matrices in $\operatorname{Mat}(n)$ which correspond to them are also real and symmetric. So the subalgebra $M$ is closed under conjugate transpose: $* \colon A \mapsto \overline{A}^T$.

So $\operatorname{Mat}(n)$ is a $C^*$-algebra with involution $*$. As a sub $C^*$-algebra of $\operatorname{Mat}(n)$, it is a finite dimensional $C^*$ algebra, we see that $M$ is semisimple by Fact 3.6. □

**Remark 3.8.** All nontrivial irreps $V$ of $H(2)$ have dimension 2. There exists $v \in V$ such that $H_1 v = av$ and $H_2 v = bv$ for all $a, b \in \mathbb{C}$. If $v$ and $sv$ are linearly independent, then $sH_1 + 1 = H_2 s \implies H_1, H_2$ act in the basis $\langle v, sv \rangle$ via the matrices

$$H_1 \mapsto \begin{bmatrix} a & -1 \\ 0 & b \end{bmatrix} \qquad H_2 \mapsto \begin{bmatrix} b & 1 \\ 0 & a \end{bmatrix} \qquad s \mapsto \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

We're trying to parameterize the Young vectors by elements of $\operatorname{Spec}(n)$ rather than chains $T$. The following theorem gives the action of $s_i$ on the Young basis in terms of weights.

**Theorem 3.9.** Let $T$ is a chain as in (9), and $\alpha(T) = (a_1, \ldots, a_n) \in \operatorname{Spec}(n)$. Take a Young vector $v_\alpha = v_T$. Then

(i) $a_i \neq a_{i+1}$ for all $i$.

(ii) $a_{i+1} = a_i \pm 1 \iff s_i v_\alpha = \pm v_\alpha \iff s_i v_\alpha, v_\alpha$ are linearly dependent.

(iii) for $i = 1, \ldots, n-2$, the following cannot occur: $a_i = a_{i+1} + 1 = a_{i+2}$ and $a_i = a_{i+1} - 1 = a_{i+2}$.

(iv) if $a_{i+1} \neq a_i \pm 1$, then $\alpha' = s_i \alpha = (a_1, \ldots, a_{i-1}, a_{i+1}, a_i, a_{i+2}, \ldots, a_n)$ belongs to $\operatorname{Spec}(n)$ and $\alpha \sim \alpha'$, where $\sim$ is the relation from Definition 2.32. Moreover,

$$v := \left( s_i - \frac{1}{a_{i+1} - a_i} \right) v_\alpha$$

is a scalar multiple of $v_{\alpha'}$. Thus, in the basis $\{v_\alpha, v_{\alpha'}\}$, the actions of $X_i$, $X_{i+1}$and $s_i$ are given by the matrices

$$X_i \mapsto \begin{bmatrix} a_i & 0 \\ 0 & a_{i+1} \end{bmatrix} \qquad X_{i+1} \mapsto \begin{bmatrix} a_{i+1} & 0 \\ 0 & a_i \end{bmatrix} \qquad s_i \mapsto \begin{bmatrix} \frac{1}{a_{i+1}-a_i} & 1 - \frac{1}{(a_{i+1}-a_i)^2} \\ 1 & \frac{-1}{a_{i+1}-a_i} \end{bmatrix}.$$

**Remark 3.10.** “If you’re going to die, please go outside."

*Proof of Theorem 3.9.* Notice that by definition of $\alpha$ and $v_\alpha$, $X_i v_\alpha = a_i v_\alpha$ and $X_{i+1} v_\alpha = a_{i+1} v_\alpha$. So (using the relation $s_i X_{i+1} - 1 = X_i s_i$), $V = \langle v_\alpha, s_i v_\alpha \rangle$ is invariant under the actions of $X_i$, $X_{i+1}$ and $s_i$. Hence, $V$ is invariant under the algebra $M$ generated by $X_i$, $X_{i+1}$ and $s_i$ (see Lemma 3.7).

(i) Suppose first that $v_\alpha$ and $s_i v_\alpha$ are linearly dependent. Then $s_i v_\alpha = \lambda v_\alpha$. Then $s_i^2 = 1 \implies \lambda^2 = 1 \implies \lambda = \pm 1$. So $s_i v_\alpha = \pm v_\alpha$. Then relation (11) ($s_i X_i s_i + s_i = X_{i+1}$) says that $a_i v_\alpha \pm v_\alpha = a_{i+1} v_\alpha$. Therefore, $s_i v_\alpha = \pm v_\alpha \implies a_{i+1} = a_i \pm 1$.

Alternatively, if $v_\alpha$, $s_i v_\alpha$ are linearly independent, let $V$ be the subspace of $V_{\lambda^{(i+1)}}$ they span. Then $V$ is $M$-invariant, and the matrices for $X_i$, $X_{i+1}$ and $s_i$ in the basis $\{v_\alpha, s_i v_\alpha\}$ are

$$X_i \mapsto \begin{bmatrix} a_i & -1 \\ 0 & a_{i+1} \end{bmatrix} \qquad X_{i+1} \mapsto \begin{bmatrix} a_{i+1} & 1 \\ 0 & a_i \end{bmatrix} \qquad s_i \mapsto \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

respectively. The action of $X_i$ on $V_{\lambda^{(i+1)}}$ is diagonalizable, since $V$ is $X_i$ invariant, then the action of $X_i$ on $V$ is also diagonalizable. Hence $a_i \neq a_{i+1}$ by Fact 3.11.

(ii) ($\Longleftarrow$) was done in (i). So now suppose that $a_{i+1} = a_i + 1$ (the proof is similar for $a_{i+1} = a_i - 1$). Assume $v_\alpha$, $s_i v_\alpha$ are linearly independent; if they are dependent, then we are in the situation of (i). Let $V$ be the subspace of $V_{\lambda^{(i+1)}}$ spanned by $s_i v_\alpha$ and $v_\alpha$. $V$ is an $M$-module, and $M$ is semisimple by Lemma 3.7.

Claim that there is only one 1-dimensional $M$-invariant subspace $W$ of $V$. If this is the case, this contradicts the fact that $M$ is semisimple (which means that representations of $M$ are completely reducible, so the compliment of $W$ should be another $M$-invariant 1-dimensional subspace.)

So suppose that $W$ is a 1-dimensional subspace of $V$ invariant under the action of $M$. Then let $W$ be spanned by $bv_\alpha + cs_i v_\alpha$. The fact that $W$ is invariant under the action of $s_i \in M$ implies that $b, c \neq 0$. So without loss of generality, set $b = 1$. Then $s_i(v_\alpha + cs_i v\alpha) = s_i v_\alpha + cv_\alpha$ and $s_i^2 = 1$ implies that $v_\alpha + cs_i v_\alpha = \pm(cv_\alpha + s_i v_\alpha)$. Hence $c = \pm 1$. But $c$ must be $-1$, since you can check that the subspace spanned by $v_\alpha + s_i v_\alpha$ isn’t $M$-invariant. Hence, $W = \mathrm{Span}(v_\alpha - s_i v_\alpha)$.

(iii) Assume $a_i = a_{i+1} - 1 = a_{i+2}$. By $(ii)$, $s_i v_\alpha = v_\alpha$ and $s_{i+1} v_\alpha = -v_\alpha$. Thus, consider the Coxeter relation $s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1}$ and let both sides act on $v_\alpha$. Then we get that $-v_\alpha = v_\alpha$, which is impossible. The other case is similar.

(iv) By $(ii)$, $v_\alpha$, $s_i v_\alpha$ are linearly independent. For $j \neq i, i+1$, we can check that $X_j v = a_j v$. Similarly by (11), $X_i v = a_{i+1} v$ and $X_{i+1} v = a_i v$. Then by Corollary 2.11$(i)$, $\alpha' \in \mathrm{Spec}(n)$ and by Corollary 2.11$(ii)$, $v$ is a scalar multiple of $v_{\alpha'}$. Clearly $\alpha \sim \alpha'$ as $v \in V_{\lambda(n)}$. The matrix representations of $s_i$, $X_i$, and $X_{i+1}$ follow. ☐

**Fact 3.11** (Linear algebra fact #58)**.** Now matrices of the form $\left(\begin{smallmatrix} a & \pm 1 \\ 0 & b \end{smallmatrix}\right)$ are diagonalizable if and only if $a \neq b$, and if so then the eigenvalue $a$ has eigenvector $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ and the eigenvalue $b$ has eigenvector $\left(\begin{smallmatrix} \pm 1/(b-a) \\ 1 \end{smallmatrix}\right)$.

**Definition 3.12.** Let $\alpha = (a_1, \ldots, a_n) \in \mathrm{Spec}(n)$. If $a_i \neq a_{i+1} \pm 1$, then $s_i$ is **admissible** for $\alpha$.

**Fact 3.13.** If $\alpha \in \mathrm{Spec}(n)$ is obtained from $\beta \in \mathrm{Spec}(n)$ by a sequence of admissible transpositions, then $\alpha \sim \beta$.

**Claim 3.14.** $\mathrm{Spec}(n)$ consists of *integral vectors*. That is, each of these $a_i$ are integers. These integers come from the content vectors for the Young tableaux.

Given this, considering the matrix of the action of $s_i$ in Theorem 3.9$(iv)$, if we choose the GZ-basis $\{v_T\}$ appropriately, all irreducible representations of $S_n$ are defined over $\mathbb{Q}$. See Theorem 6.2.

# 4 Content vectors and tableaux

In Vershik-Okounkov theory, the Young tableaux are related to the irreducible representations using **content vectors**.

**Definition 4.1.** Call $\alpha = (a_1, \ldots, a_n) \in \mathbb{Z}^n$ a **content vector** if

(i) $a_1 = 0$.

(ii) for all $i > 1$, $\{a_i - 1, a_i + 1\} \cap \{a_1, \ldots, a_{i-1}\} \neq \varnothing$

(iii) if $a_i = a_j = a$ for some $i < j$, then $\{a-1, a+1\} \subseteq \{a_{i+1}, \ldots, a_{j-1}\}$. That is, between any two occurrences of $a$, there should also be occurrences of $a-1$ and of $a+1$.

Last time we wrote down the definition of a content vector, which was very confusing. This time, we may or may not explain what that actually means. Eventually, we'll show that $\mathrm{Cont}(n) = \mathrm{Spec}(n)$.

**Definition 4.2.** $\mathrm{Cont}(n) \subseteq \mathbb{Z}^n$ is the set of all content vectors of length $n$.

**Example 4.3.** $\mathrm{Cont}(1) = \{0\}$, and $\mathrm{Cont}(2) = \{(0,1), (0,-1)\}$.

**Proposition 4.4.** We can strengthen $(ii)$ and $(iii)$ from Definition 4.1 as follows:

(ii)′ For all $i > 1$, if $a_i > 0$, then $a_j = a_i - 1$ for some $j < i$ and if $a_i < 0$ then $a_j = a_i + 1$ for some $j < i$.

*Proof.* If $a_i > 0$, then by $(i)$ and repeated use of $(ii)$, we construct a sequence $a_i = a_{s_0}, a_{s_1}, \dots, a_{s_k} = 0$ such that $s_0 = i > s_1 > \dots > s_k \geqslant 1$ with $a_{s_h} > 0$ and $|a_{s_h} - a_{s_{h+1}}| = 1$ for all $h = 0, 1, \dots, k-1$. Then, as $h$ varies, $a_{s_h}$ attains all integer values between 0 and $a_i$. In particular, it must attain $a_i - 1$.

The case for $a_i < 0$ is similar. □

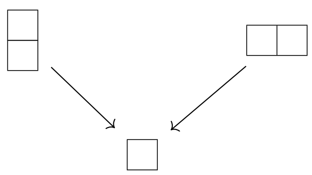
(iii)′ If $i < j$, $a_i = a_j$ and $a_r \neq a_j$ for all $r = i+1, \dots, j-1$, then there exists some unique $s_-, s_+ \in \{i+1, \dots, j-1\}$ such that $a_{s_-} = a_j - 1$ and $a_{s_+} = a_j + 1$.

*Proof.* If $i < j$, $s_-$ and $s_+$ exist by $(iii)$ and uniqueness from the fact that there exists another $s'_-$ such that $a_{s'_-} = a_j - 1$, say with $s_- < s'_-$, then by $(iii)$ there exists $s$ between $s_-$ and $s'_-$ such that $a_s = (a_j - 1) + 1 = a_j$. This is a contradiction. □

**Theorem 4.5.** For all $n \geqslant 1$, $\mathrm{Spec}(n) \subseteq \mathrm{Cont}(n)$.

*Proof.* Proof by induction on $n$. If $n = 1$, this is trivial.

For $n = 2$, The irreducible representations of $S_2$ are the trivial and sign representations. The Bratelli diagram of $S_1 \leqslant S_2$ is



Now $X_2 = (1,2)$ and if $v \in V_{}$, then $X_2 v = v$, while if $w \in V_{}$, $X_2 w = -w$. Hence, $\mathrm{Spec}(2) = \{(0,1), (0,-1)\}$. Now see Example 4.3 to see that this is exactly the content vectors $\mathrm{Cont}(2)$.

Now suppose that $\mathrm{Spec}(n-1) \subseteq \mathrm{Cont}(n-1)$. Let $\alpha \in \mathrm{Spec}(n)$ with $\alpha = (a_1, \dots, a_n)$. As $X_1 = 0$, then clearly $a_1 = 0$ so condition Definition 4.1(i) is satisfied. By the fact that if $\alpha \in \mathrm{Spec}(n)$, then $\alpha' = (a_1, \dots, a_{n-1}) \in \mathrm{Spec}(n-1)$, so we just need to verify that conditions Definition 4.1(ii) and Definition 4.1(iii) for $n$.

Let's show (ii). For the sake of contradiction, assume

$$\{a_n - 1, a_n + 1\} \cap \{a_1, \dots, a_{n-1}\} = \varnothing \tag{13}$$

Now by Theorem 3.9(iv), $(n-1, n)$ is admissible for $\alpha$, which means that $(a_1, \dots, a_{n-2}, a_n, a_{n-1}) \in \mathrm{Spec}(n)$. Hence, $(a_1, \dots, a_{n-2}, a_n) \in \mathrm{Spec}(n-1) \subseteq$

$\mathrm{Cont}(n-1)$. By (13), $\{a_n - 1, a_n + 1\} \cap \{a_1, \ldots, a_{n-2}\} = \varnothing$, contradicting Definition 4.1(ii) for the content vector $(a_1, \ldots, a_{n-2}, a_n) \in \mathrm{Cont}(n-1)$.

Now we need to verify Definition 4.1(iii) for $j = n$. Then again for the sake of contradiction, suppose $\alpha$ does not satisfy Definition 4.1(iii) for $j = n$, i.e. assume $a_i = a_n = a$ for some $i < n$. Assume that $i$ is the largest possible index, that is, $a$ does not occur between $a_i$ and $a_n$.

$$a \notin \{a_{i+1}, \ldots, a_{n-1}\} \tag{14}$$

Assume $a - 1 \notin \{a_{i+1}, \ldots, a_{n-1}\}$ $\big($the other case where $a + 1 \notin \{a_{i+1}, \ldots, a_{n-1}\}$ is very similar$\big)$.

Since $(a_1, \ldots, a_{n-1}) \in \mathrm{Cont}(n-1)$, by inductive hypothesis $a + 1$ can only occur in $\{a_{i+1}, \ldots, a_{n-1}\}$ at most once $\big($if twice, then by induction, $a$ also occurs, contradicting maximality of $i$ (14)$\big)$. There are two cases: either $a + 1 \notin \{a_{i+1}, \ldots, a_{n-1}\}$ or $a + 1 \in \{a_{i+1}, \ldots, a_{n-1}\}$.

In the first case, we have $(a_i, \ldots, a_n) = (a, *, \ldots, *, a)$ where $*$ is a number different from $a-1, a, a+1$. We can apply a sequence of $n - i + 1$ admissible transpositions to deduce that $\alpha \sim \alpha' = (\ldots, a, a, \ldots) \in \mathrm{Spec}(n)$. This is a contradiction of Theorem 3.9(i).

In the second case, we have $(a_i, \ldots, a_n) = (a, *, \ldots, *, a+1, *, \ldots, *, a)$ where $*$ is a number different from $a - 1$, $a$, $a + 1$. We can apply a sequence of admissible transpositions to infer that $\alpha \sim \alpha' = (\ldots, a, a+1, a, \ldots) \in \mathrm{Spec}(n)$, contrary to Theorem 3.9(iii). □

**Definition 4.6.** If $\alpha = (a_1, \ldots, a_n) \in \mathrm{Cont}(n)$, and $a_i \neq a_{i+1} \pm 1$, we say that the transposition $s_i$ is **admissible** for $\alpha$. We can define an equivalence relation on $\mathrm{Cont}(n)$: $\alpha \approx \beta$ if $\beta$ can be obtained from $\alpha$ by a sequence of admissible transpositions.

**Remark 4.7.** Given $\alpha \in \mathrm{Cont}(n)$, there can exist $\sigma \in S_n$ such that $\sigma\alpha \notin \mathrm{Cont}(n)$, e.g. $\alpha = (0, 1) \in \mathrm{Cont}(2)$, $\sigma = (1\,2)$, but $\sigma\alpha = (1, 0)$ which is not a content vector, because they always begin with a zero.

**Definition 4.8.** The Young graph $\mathbb{Y}$ has vertices the Young diagrams with two vertices $\mu$ and $\lambda$ connected by a directed edge from $\lambda$ to $\mu$ if and only if $\mu \subseteq \lambda$ and $\lambda \backslash \mu$ is a single box. Write $\lambda \to \mu$ or $\mu \nearrow \lambda$ and say that $\lambda$ **covers** $\mu$.

**Definition 4.9.** The **content** $c(\square)$ of a box $\square$ in a Young diagram is the $y$-coordinate minus the $x$-coordinate. The content of a Tableau of shape $\lambda$ is best given by example.

**Example 4.10.** For $\lambda = (4, 3, 1)$, take a standard tableaux

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

Replace the number in each box by its content

| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| -1 | 0 | 1 | |
| -2 | | | |

Then the content of $T$ is $\alpha = C(T) = (0, 1, -1, 0, 2, 1, 3, -2)$ and is given by $a_i$ is the content of the box with number $i$ in it.

The choice of tableaux of shape $\lambda$ determines the order in which the $c(\square)$ appear in $C(T)$.

**Definition 4.11.** Note that the Young diagram of shape $\lambda = (\lambda_1, \dots, \lambda_k)$ can be divided into **diagonals** numbered $-k+1, -k+2, \dots, 0, 1, 2, \dots, \lambda_1 - 1$. The diagonal numbered $r$ is all those boxes with coordinates $(i, j)$ such that $c(i, j) = j - i = r$.

Recall that $\mathrm{Tab}(\lambda) = \mathrm{SYT}(\lambda)$ is all paths in $\mathbb{Y}$ from $\lambda$ to the unique partition of 1. These correspond bijectively to the standard tableaux of shape $\lambda$. Given a path $T \in \mathrm{Tab}(\lambda)$,

$$\lambda^{(n)} = \lambda \to \lambda^{(n-1)} \to \cdots \to \lambda^{(1)} = (1),$$

recall that we can represent it by taking the Young diagram for $\lambda$ and writing $1, 2, \dots, n$ in the boxes $\lambda^{(1)}, \lambda^{(2)} \backslash \lambda^{(1)}, \dots, \lambda^{(n)} \backslash \lambda^{(n-1)}$, respectively. Let

$$\mathrm{SYT}(n) = \mathrm{Tab}(n) = \bigcup_{\lambda \vdash n} \mathrm{Tab}(\lambda).$$

**Definition 4.12.** Let $T_1 \in \mathrm{Tab}(n)$ and assume $i, i+1$ do not appear in the same row or column of $T_1$. Then switching $i \leftrightarrow i+1$ in $T_1$ preserves the standardness, producing another tableaux $T_2 \in \mathrm{Tab}(n)$. In this case, say $T_2$ can be obtained from $T_1$ by an **admissible** transposition. For $T_1, T_2 \in \mathrm{Tab}(n)$, we write $T_1 \approx T_2$ if $T_2$ can be obtained from $T_1$ by a sequence of (0 or more) admissible transpositions.

**Lemma 4.13.** Let $\Phi\colon \mathrm{Tab}(n) \to \mathrm{Cont}(n)$ be defined as follows: given a tableaux

$$T = \left(\lambda^{(n)} \to \lambda^{(n-1)} \to \cdots \to \lambda^{(1)} = (1)\right) \in \mathrm{Tab}(n),$$

define

$$\Phi(T) = C(T) = \left(c(\lambda^{(1)}), c(\lambda^{(2)} \backslash \lambda^{(1)}), \dots, c(\lambda^{(n)} \backslash \lambda^{(n-1)})\right),$$

to be the content of $T$. Then $\Phi$ is a bijection which takes $\approx$-equivalent standard Young tableaux to $\approx$-equivalent content vectors.

*Proof.* The idea is that the content vector of any SYT satisfies Definition 4.1(i),(ii),(iii), and these conditions uniquely determine the tableaux as a sequence of boxes of the Young diagram.

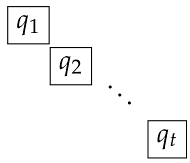
Take $T$ standard and let $C(T) = (a_1, \dots, a_n)$ be it's content. Now $a_1 = 0$: the only way for it to be standard is if 1 is in the upper left spot. So Definition 4.1(i) holds.

If $q \in \{2, \dots, n\}$ is placed in position $(i, j)$ such that $a_q = j - i$, then we have $i > 1$ or $j > 1$. In the first case, consider the number $p$ in box with coordinates $(i-1, j)$ (this is the box above). Then $p < q$ as $T$ is standard, and $a_p = j - i + 1 = a_q + 1$. Similarly, if $j > 1$ consider the number $p'$ in box with coordinates $(i, j-1)$ (this is the box to the left). Then we have $p' < q$ as $T$ is standard, and $a_{p'} = j - 1 - i = a_q - 1$ Hence Definition 4.1(ii) is satisfied.

Now suppose that $a_p = a_q$ with $p < q$. This means that $p, q$ are on the same diagonal. If $(i, j)$ are coordinates of the box containing $q$, then $i, j > 1$. Denote by $q_-$ and $q_+$ the numbers placed in the boxes with coordinates $(i-1, j)$ and $(i, j-1)$, respectively. Note that $q_-, q_+ \in \{p+1, \dots, q-1\}$ because $T$ is standard. By the same argument as above, $a_{q_+} = a_q - 1$ and $a_{q_-} = a_q + 1$. This proves Definition 4.1(iii), thus $C(T) \in \mathrm{Cont}(n)$.

Thus we have shown that $\Phi$ is well-defined.

Now claim that $T \mapsto C(T)$ is injective. Suppose $C(T) = (a_1, \dots, a_n)$, then the diagonal $h$ in $T$ is filled with numbers $q \in \{1, \dots, n\}$ such that $a_q = h$ from northwest to southeast (top-left to down-right).




where $q_1 < \dots < q_t$ and $a_{q_1} = \dots = a_{q_t} = h$ and $a_q \neq h$ if $q \notin \{q_1, \dots, q_t\}$. So if $T_1, T_2 \in \mathrm{Tab}(n)$ have the same content, namely $C(T_1) = C(T_2)$, then they have the same diagonals and must coincide.

Finally, claim that $\Phi \colon T \mapsto C(T)$ is surjective. By induction on $n$. For $n = 1, 2$, the result is clear. So suppose that $\mathrm{Tab}(n-1) \to \mathrm{Cont}(n-1)$ is surjective. Let $\alpha = (a_1, \dots, a_n) \in \mathrm{Cont}(n)$. Then $\alpha' = (a_1, \dots, a_{n-1}) \in \mathrm{Cont}(n-1)$, so by induction hypothesis there is $T' \in \mathrm{Tab}(n-1)$ such that $C(T') = \alpha'$. Now claim that adding the southeast-most (lower-right-most) diagonal box in a diagonal of $T'$ and placing $n$ in this box gives a tableaux $T \in \mathrm{Tab}(n)$ such that $C(T) = \alpha$.

If $a_n \notin \{a_1, \dots, a_{n-1}\}$, then add a box on the first row (if $a_n - 1 \in \{a_1, \dots, a_n\}$) or in the first column (if $a_n + 1 \in \{a_1, \dots, a_{n-1}\}$).

If $a_n \in \{a_1, \dots, a_{n-1}\}$ and $p$ is the largest index $\leqslant n-1$ such that $a_p = a_n$, and $p$ is the largest index $\leqslant n-1$ such that $a_p = a_n$, then if the coordinates of the box containing $p$ are $(i, j)$, place $n$ in the new box with coordinates $(i+1, j+1)$. This box is indeed addable because Proposition 4.4(iii)' guarantees the existence (and uniqueness) of $r, s \in \{p+1, p+2, \dots, n\}$ such that $a_r = a_n + 1$ and $a_s = a_n - 1$:

| | |
|---|---|
| $p$ | $r$ |
| $s$ | $n$ |

(See also the book by the Italians, 3.1.10.) □

**Lemma 4.14.** Suppose $T_1, T_2 \in \mathrm{Tab}(n)$. Then $T_1 \approx T_2$ if and only if the Young diagrams of $T_1$ and $T_2$ have the same shape.

This was already done in Proposition 1.84.

*Proof.* The ($\Longrightarrow$) direction is clear.

For the ($\Longleftarrow$) direction, let $\mu = (\mu_1, \ldots, \mu_r) \vdash n$. Define $R = R_\mu \in \mathrm{Tab}(n)$ as follows: in row 1, write $1, 2, \ldots, \mu_1$ in increasing order, in row 2 write $\mu_1 + 1, \ldots, \mu_1 + \mu_2$, and so on. For example if $\mu = (4, 2, 2, 1) \vdash 9$,

$$R_\mu = \begin{array}{|c|c|c|c|} \hline 1 & 2 & 3 & 4 \\ \hline 5 & 6 \\ \cline{1-2} 7 & 8 \\ \cline{1-2} 9 \\ \cline{1-1} \end{array}$$

In Proposition 1.84, we called this the **canonical tableau**, and denoted it $T^\mu$.

The claim is that for any $T \in \mathrm{Tab}(n)$, $T \approx R_\mu$. Look at the last box of the last row of $T$. Let $i$ be written in this box. Now swap $i \leftrightarrow i+1$ in $T$, which is clearly admissible. Repeat with $i+1 \leftrightarrow i+2$, and then with $i+2 \leftrightarrow i+3$, and so on, ending with $n-1 \leftrightarrow n$. At the end of this sequence of admissible transpositions, we end up with $n$ written in the last box of the last row of $T$. Now repeat for $n-1, n-2, \ldots, 2$. □

**Remark 4.15.** Lemma 4.14 is the same as Corollary 1.85.

Under the bijection $\mathrm{Tab}(n) \longleftrightarrow \mathrm{Cont}(n)$, this is just realizing that the definition of $\alpha \approx \beta$ for $\alpha, \beta \in \mathrm{Cont}(n)$ means that the corresponding tableaux have the same number of boxes in each diagonal.

**Remark 4.16.** Let $s$ be the permutation mapping $R_\mu$ to $T$ from the proof of Lemma 4.14. Then the proof shows that $R_\mu$ can be obtained from $T$ by a sequence of $\ell(s)$-many admissible transpositions. Thus, $T$ can be obtained from $R_\mu$ by a sequence of $\ell(s)$-many admissible transpositions. This says that $\mathrm{Cont}(n)$ is **totally geodesic** subset of $\mathbb{Z}^n$ for the action of $S_n$. This means that along with any two vectors, $\mathrm{Cont}(n)$ contains chains of vectors realizing the minimal path between them.

# 5 Main result and its consequences

**Theorem 5.1.**

(i) $\mathrm{Spec}(n) = \mathrm{Cont}(n)$ and the equivalence relations $\sim$ and $\approx$ coincide.

(ii) $\Phi^{-1}\colon \mathrm{Spec}(n) \to \mathrm{Tab}(n)$ is a bijection and for $\alpha, \beta \in \mathrm{Spec}(n)$, and moreover $\alpha \sim \beta$ if and only if $\Phi^{-1}(\alpha), \Phi^{-1}(\beta)$ have the same Young diagram.

(iii) The branching graph of a chain of symmetric groups is the Young graph $\mathbb{Y}$.

(iv) The spectrum of the Gelfand-Tsetlin algebra $GZ_n$ is the space of paths in $\mathbb{Y}_n$ (= space of standard Young tableaux with $n$ boxes).

*Proof.*

(i)
- By Theorem 4.5, $\mathrm{Spec}(n) \subseteq \mathrm{Cont}(n)$.
- If $\alpha \in \mathrm{Spec}(n), \beta \in \mathrm{Cont}(n), \alpha \approx \beta$, then $\beta \in \mathrm{Spec}(n)$ and $\alpha \sim \beta$. This uses Lemma 4.14 and Theorem 3.9(ii).
- It follows from the previous bullet that given an $\sim$-equivalence class $\mathcal{C}$ of $\mathrm{Spec}(n)$ and an $\approx$-equivalence class $\mathcal{D}$ of $\mathrm{Cont}(n)$, then either $\mathcal{C} \cap \mathcal{D} = \varnothing$, or $\mathcal{D} \subseteq \mathcal{C}$.
- But in fact, these two sets $\mathrm{Spec}(n)/\sim$ and $\mathrm{Cont}(n)/\approx$ have the same cardinality. Recall that $p(n)$ is the number of partitions of $n$.

$$\#(\mathrm{Spec}(n)/\sim) = \#\text{irreps} = \#\text{conjugacy classes of } S_n = p(n)$$

$$\#(\mathrm{Cont}(n)/\approx) = \#(\mathrm{SYT}(n)/\approx) = \#\text{diagrams} = p(n)$$

Therefore $|\,\mathrm{Spec}(n)/\sim| = |\,\mathrm{Cont}(n)/\approx|$.
- This proves Theorem 5.1(i); that is, $\mathrm{Spec}(n) = \mathrm{Cont}(n)$ and the relations $\sim$ and $\approx$ coincide.

(ii) Follows from Lemma 4.14.

(iii) & (iv) We have a natural bijective correspondence between the set of all paths in the branching graph, parameterized by $\mathrm{Spec}(n)$ and the set of all paths in $\mathbb{Y}$, parameterized by $\mathrm{Cont}(n)$. $\left[\text{Combine } \Pi_n(\mathbb{Y}) \leftrightarrow \mathrm{Tab}(n) \text{ with bijection Lemma 4.13 to get a bijection } \Pi_n(\mathbb{Y}) \leftrightarrow \mathrm{Cont}(n).\right]$

Notice also that by Lemma 4.14, if $\alpha, \beta \in \mathrm{Cont}(n)$ correspond to paths

$$\lambda^{(n)} \to \cdots \to \lambda^{(1)} \text{ and } \mu^{(n)} \to \cdots \to \mu^{(1)},$$

respectively, then $\alpha \approx \beta \iff \lambda^{(n)} = \mu^{(n)}$. So we have a bijective correspondence between vertices of these graphs and it's easy to see that this gives a graph isomorphism. □

Following Theorem 5.1, we have a natural correspondence between $S_n^\wedge$ and the $n$-th level of the branching graph $\mathbb{Y}$.

**Definition 5.2.** Given $\lambda \vdash n$, denote by $S^\lambda$ the irreducible representation of $S_n$ spanned by vectors $\{v_\alpha\}$ with $\alpha \in \mathrm{Spec}(n) = \mathrm{Cont}(n)$ corresponding to the standard tableaux of shape $\lambda$. $S^\lambda$ is called the **Specht module**.

Note that $\dim S^\lambda = \#\text{standard } \lambda\text{-tableaux} = f_\lambda$.

Our results give the branching theorems for restriction and induction of Specht modules.

**Corollary 5.3.** Let $0 \leqslant k < n$ and $\lambda \vdash n$ and $\mu \vdash k$. Let $m_{\mu,\lambda} = \left[\mathrm{Res}^{S_n}_{S_k} S^\lambda : S^\mu\right]$ be the multiplicity of $S^\mu$ in $\mathrm{Res}^{S_n}_{S_k} S^\lambda$. Then

$$m_{\mu,\lambda} = \begin{cases} 0 & \mu \not\leq \lambda \\ \#\text{paths in } \mathbb{Y} \text{ from } \lambda \text{ to } \mu & \text{otherwise} \end{cases}$$

In any case, $m_{\mu,\lambda} \leqslant (n-k)!$ and this estimate is sharp.

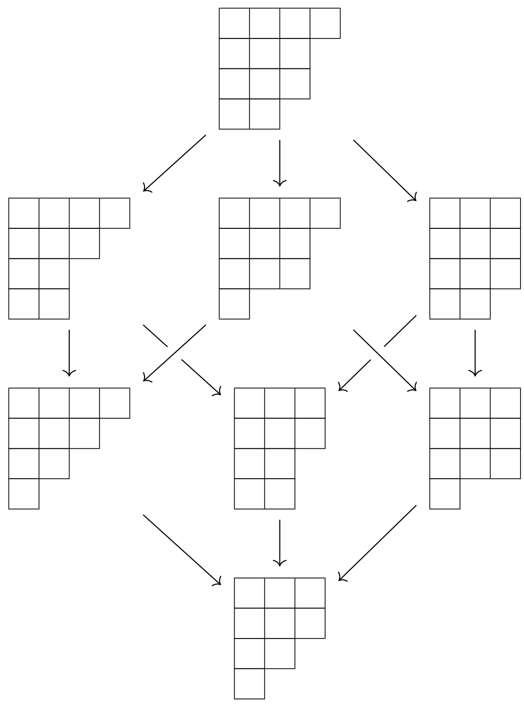
*Proof.*

$$\operatorname{Res}_{S_k}^{S_n} S^\lambda = \operatorname{Res}_{S_k}^{S_{k+1}} \operatorname{Res}_{S_{k+1}}^{S_{k+2}} \cdots \operatorname{Res}_{S_{n-1}}^{S_n} S^\lambda$$

and each step has a decomposition that is multiplicity-free determined by paths in $\mathbb{Y}$. Therefore, $m_{\mu,\lambda}$ is the number of paths in $\mathbb{Y}$ starting at $\lambda$ and ending at $\mu$, or equivalently, the number of ways to obtain the diagram of shape $\lambda$ from the diagram of shape $\mu$ by adding successively $n-k$ addable boxes to the diagram of shape $\mu$ (at each stage you have a diagram of a partition).

So in particular, the multiplicity is at most $(n-k)!$. This bound is sharp when boxes can be added to different rows and columns. □

**Example 5.4.**



There are $6 = (12-9)!$ paths from $(4,3,3,2) \vdash 12$ to $(3,3,2,1) \vdash 9$. So the bound in Corollary 5.3 is sharp.

**Corollary 5.5** (The Branching Rule)**.** For $\lambda \vdash n$

$$\operatorname{Res}_{S_{n-1}}^{S_n} S^\lambda = \bigoplus_{\substack{\mu \vdash (n-1) \\ \lambda \to \mu}} S^\mu$$

summed over all $\mu \vdash (n-1)$ obtained from $\lambda$ by removing one box. Moreover, for $\lambda \vdash (n-1)$,

$$\operatorname{Ind}_{S_{n-1}}^{S_n} S^\mu = \bigoplus_{\substack{\lambda \vdash n \\ \lambda \to \mu}} S^\lambda$$

by Frobenius reciprocity.

Consider the map $\lambda \mapsto S^\lambda$ sending a partition of $n$ to an irrep of $S_n$. Here's a characterization of this map.

**Corollary 5.6.** For all $n \geqslant 1$, let $\{V_\lambda \mid \lambda \vdash n\}$ be a family of representations of $S_n$ indexed by $\lambda$ such that

(i) $V_{\square} \cong S^{(1)}$ is the trivial and unique representation of $S_1$;

(ii) $V_{\square\square}$ and $V_{\substack{\square\\\square}}$ are the trivial and alternating representations of $S_2$, respectively;

(iii) $\operatorname{Ind}_{S_{n-1}}^{S_n} V_\mu = \bigoplus_{\substack{\lambda \vdash n\\ \lambda \to \mu}} V_\lambda$ for all $\mu \vdash (n-1)$ and $n \geqslant 2$.

Then $V_\lambda$ is irreducible and isomorphic to $S^\lambda$ for all $\lambda \vdash n$.

**Exercise 5.7.** Prove Corollary 5.6. Use the fact that $\lambda \vdash n$ is uniquely determined by $\{\mu \vdash (n-1) \mid \lambda \to \mu\}$.

**Example 5.8.** In Corollary 5.5, take $k = n-2$. Let $\lambda \vdash n$, $\mu \vdash (n-2)$.

If $\mu \not\leq \lambda$, then $[\operatorname{Res}_{S_{n-2}}^{S_n} S^\lambda : S^\mu] = 0$.

If $\mu \leq \lambda$, then $[\operatorname{Res}_{S_{n-2}}^{S_n} S^\lambda : S^\mu] \leqslant 2$. There are two cases.

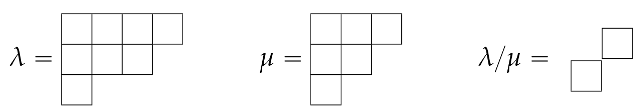
(1) If there is a unique partition $\nu \vdash (n-1)$ such that $\mu \leq \nu \leq \lambda$. So in $\mathbb{Y}$, between $\mu$ and $\lambda$, there is a chain $\lambda \to \nu \to \mu$.

Boxes in $\lambda/\mu$ are on the same row or the same column (if they weren't, there'd be more than one $\nu$). If

$$\lambda = \lambda^{(n)} \to \nu = \lambda^{(n-1)} \to \lambda^{(n-2)} = \mu \to \lambda^{(n-3)} \to \cdots \to \lambda^{(1)}$$

is any path containing $\lambda \to \nu \to \mu$, then it corresponds to a spectral vector $\alpha = (a_1, \ldots, a_n) \in \operatorname{Spec}(n)$ where $a_n = a_{n-1} \pm 1$. ($a_n = a_{n-1} + 1$ if boxes of $\lambda/\mu$ are in the same row, or $a_n = a_{n-1} - 1$ if they're in the same column.) In particular, $s_{n-1} v_\alpha = \pm v_\alpha$ as in Theorem 3.9. Note also that $s_{n-1}$ only affects the $(n-1)$-th level of the diagram and $\nu$ is the only partition between $\mu$ and $\lambda$: see Lemma 3.1.

(2) There are two partitions $\nu, \eta \vdash (n-1)$ such that $\mu \leq \nu, \eta \leq \lambda$. Boxes of $\lambda/\mu$ are on different rows or columns. For example,




and the Bratelli diagram from $\lambda$ to $\mu$ is the square

$$\begin{array}{ccc} & \lambda & \\ \swarrow & & \searrow \\ \nu & & \eta \\ \searrow & & \swarrow \\ & \mu & \end{array}$$

If $\alpha \in \operatorname{Spec}(n)$ corresponds to a path $\lambda \to \nu \to \mu \to \cdots$, then $a_n \neq a_{n-1} \pm 1$, and $\alpha' = (a_1, \dots a_{n-2}, a_n, a_{n-1}) \in \operatorname{Spec}(n)$ corresponds to the path $\lambda \to \eta \to \mu \to \cdots$. The action of $s_{n-1}$ on $v_\alpha$, $v'_\alpha$ is as given in Theorem 3.9(iv) and Lemma 3.1 is also confirmed. See also Theorem 6.2.

# 6 Young's Seminormal and Orthogonal Forms

Let $T^\lambda$ be the "canonical" standard tableau from Proposition 1.84 (alternatively called $R_\lambda$ in the proof of Lemma 4.14). The chain $S_1 \leqslant S_2 \leqslant \ldots \leqslant S_n$ determines a decomposition of every irrep of $S_n$ into 1-dimensional subspaces, and the $GZ$-basis is obtained by choosing a nontrivial vector in each of these subspaces.

If such vectors are normalized with respect to any inner product we say it's an **orthonormal basis**; otherwise it's an **orthogonal basis**. In both cases, the vectors are defined up to a scalar factor (of modulus 1, if normalized).

We saw in Theorem 5.1 that we could parametrize vectors in the $GZ$-basis by standard tableaux: for $T \in \operatorname{Tab}(\lambda)$, let $v_T$ be the corresponding vector in the GZ-basis. We'll discuss the choice of scalar factors in the Young basis $\{v_T\}$ such that all irreps of $S_n$ are defined over $\mathbb{Q}$.

Recall that $\sigma_T$ is the unique permutation that, when applied to the Tableaux $T$, takes you to the standard tableaux $T^\lambda$; see Proposition 1.84. We have $\sigma_T T = T^\lambda$ We will also use Theorem 3.9, so you'd best go take a look at that too.

**Proposition 6.1.** It is always possible to choose the scalar factors of vectors $\{v_T \mid T \in \operatorname{Tab}(n)\}$ in such a way that for the tableaux $T$ in $\operatorname{Tab}(n)$ one has

$$\sigma_T^{-1} v_{T^\lambda} = v_T + \sum_{\substack{R \in \operatorname{Tab}(\lambda) \\ \ell(\sigma_R) < \ell(\sigma_T)}} \alpha_R v_R$$

where $\alpha_R \in \mathbb{C}$ (actually, we'll see in Corollary 6.3 that $\alpha_R \in \mathbb{Q}$), and $\sigma_T$ is as in Proposition 1.84.

*Proof.* Induction on $\ell(\sigma_T)$. At each stage, you choose scalar factors for all $T$ with $\ell(\sigma_T) = \ell$.

If $\ell(\sigma_T) = 1$, then $\sigma_T$ is an admissible transposition for $T^\lambda$ and so by Theorem 3.9 you're done. Can use Theorem 3.9(iv) to choose the scalar factors of $v_T$ (which corresponds to $v_{\alpha'}$ in the statement of that theorem).

Now if $\ell(\sigma_T) > 1$, suppose $\sigma_T = s_{i_1} s_{i_2} \cdots s_{i_{\ell-1}} s_j$ is the standard decomposition of $\sigma_T$ into the product of admissible transpositions (see Corollary 1.85). Then $\sigma_T = \sigma_{T_1} s_j$ where $T_1 = s_j T$ is standard. Note that $\ell(\sigma_{T_1}) = \ell(\sigma_T) - 1$. By induction hypothesis, can assume

$$\sigma_{T_1}^{-1} v_{T^\lambda} = s_j \sigma_T^{-1} v_{T^\lambda} = v_{T_1} + \sum_{\substack{R \in \operatorname{Tab}(\lambda) \\ \ell(\sigma_R) < \ell(\sigma_{T_1})}} \alpha_R^{(1)} v_R \tag{15}$$

Since $T = s_j T_1$, the formula in Theorem 3.9(iv) means we can choose a scalar factor of $v_T$ such that

$$s_j v_{T_1} = v_T + \left(\frac{1}{a_{j+1} - a_j}\right) v_{T_1} \tag{16}$$

where $(a_1, \ldots, a_n)$ is the content of $T_1$. Hence the result follows from (15) and (16), remembering Theorem 3.9 for the computation of $s_j v_R$ for $R \in \mathrm{Tab}(\lambda)$ and $\ell(\sigma_R) < \ell(\sigma_{T_1})$. □

**Theorem 6.2** (Young's seminormal form)**.** Choose vectors of $GZ$-basis of $S_n$ according to Proposition 6.1. Then if $T \in \mathrm{Tab}(\lambda)$, and $C(T) = (a_1, \ldots, a_n)$ is the content of $T$, then the adjacent transpositions $s_j$ act on $v_T$ as follows:

(i) if $a_{j+1} = a_j \pm 1$ then $s_j v_T = \pm v_T$

(ii) if $a_{j+1} \neq a_j \pm 1$ then setting $T' = s_j T$;

$$s_j v_T = \begin{cases} \left(\frac{1}{a_{j+1}-a_j}\right) v_T + v_{T'} & \ell(\sigma_{T'}) > \ell(\sigma_T) \\ \left(\frac{1}{a_{j+1}-a_j}\right) v_T + \left(1 - \frac{1}{(a_{j+1}-a_j)^2}\right) v_{T'} & \ell(\sigma_{T'}) < \ell(\sigma_T) \end{cases}$$

*Proof.* (i) follows from Theorem 3.9(ii)

(ii) also from Theorem 3.9, but you have to check that the action that is described is consistent with the choice made in Proposition 6.1. Need to show that $s_j v_T$ has exactly the required expression. We'll do one case, the other one is similar.

If $\ell(\sigma_{T'}) > \ell(\sigma_T)$, (recall $T' = s_j T$) then we have $\sigma_{T'} = \sigma_T s_j$. We know from Proposition 6.1 that

$$\sigma_T^{-1} v_{T^\lambda} = v_T + \sum_{\substack{R \in \mathrm{Tab}(\lambda) \\ \ell(\sigma_R) < \ell(\sigma_T)}} \alpha_R v_R$$

Then putting these two things together,

$$\sigma_{T'}^{-1} v_{T^\lambda} = v_{T'} + \sum_{\substack{R' \in \mathrm{Tab}(\lambda) \\ \ell(\sigma_{R'}) < \ell(\sigma_{T'})}} \alpha'_{R'} v_{R'}$$

$$s_j(\sigma_T^{-1} v_{T^\lambda}) = s_j v_T + s_j \left( \sum_{\substack{R \in \mathrm{Tab}(\lambda) \\ \ell(\sigma_R) < \ell(\sigma_T)}} \alpha_R v_R \right)$$

Notice that the coefficient of $v_{T'}$ in $s_j v_T$ is 1, which means that the coefficient of $v_T$ in $s_j(\sigma_T^{-1} v_{T^\lambda})$ agrees with the coefficient of $v_T$ in $\sigma_{T'}^{-1} v_{T^\lambda}$. So Theorem 3.9 holds in exactly the required form

The case when $\ell(\sigma_{T'}) < \ell(\sigma_T)$ is analogous, but starting from $\sigma_T = \sigma_{T'} s_j$ and using $\alpha$ as the content of $T'$ when applying Theorem 3.9. □

**Corollary 6.3.** In the bases of Proposition 6.1, Theorem 6.2, the matrix coefficients of the irreducible representations of $S_n$ belong to $\mathbb{Q}$. In particular, the coefficients $\alpha_R$ in Proposition 6.1 are rational.

**Exercise 6.4.** Prove Theorem 6.2 by verifying the given formulae define a representation of $S_n$. That is, verify the Coxeter relations.

**Definition 6.5.** The basis and action described in Theorem 6.2 above are called **Young's Seminormal Form**.

Now normalize the basis $\{v_T \mid T \in \text{Tab}(n)\}$ of $S^\lambda$ by taking

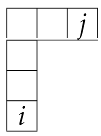
$$w_T = \frac{1}{\|v_T\|_{S^\lambda}} v_T$$

where $\|\cdot\|_{S^\lambda}$ is a norm associated with some arbitrary $S_n$-invariant scalar product on $S^\lambda$ that makes $S^\lambda$ into a unitary representation.

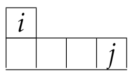
**Definition 6.6.** If $T$ is a standard tableau with $C(T) = (a_1, \ldots, a_n)$. If $i, j \in \{1, \ldots, n\}$, define the **axial distance** from $j$ to $i$ in $T$ to be $a_j - a_i$.

Geometrically, this means that moving from $j$ to $i$ in the tableau, each step to the left or down is counted as $+1$, and each step to the right or up gets a $-1$, then the resulting integer exactly $a_j - a_i$.

**Example 6.7.** In the tableau below, $a_j - a_i = 2 - (-3) = 5$.




Similarly, for this thing,




$a_i - a_j = 2$ which is the number of steps from $i$ to $j$ when counted with signs as in Definition 6.6

**Theorem 6.8.** Consider the orthonormal basis $\{w_T : T \in \text{Tab}(n)\}$. Then

$$s_j w_T = \frac{1}{r} w_T + \sqrt{1 - \frac{1}{r^2}}\, w_{s_j T}$$

where, for $C(T) = (a_1, \ldots, a_n)$, then $r = a_{j+1} - a_j$ is the axial distance from $j+1$ to $j$. In particular, for $a_{j+1} = a_j \pm 1$ then $r = \pm 1$ and $s_j w_T = \pm w_T$.

*Proof.* First, notice that by the choice of the inner product that makes $S^\lambda$ into a unitary representation, each element of $S_n$ is a unitary operator and hence preserves the norm. In particular, $\|s_j v_T\| = \|v_T\|$. Moreover, $v_{T'} \perp v_T$ for all $v_T, v_{T'}$ in the GZ-basis by Maschke's Theorem – these representations are eventually in different irreducible components of $S^\lambda$ upon restriction, and therefore must be orthogonal.

Now let $T' = s_j T$ and suppose that $\ell(\sigma_{T'}) > \ell(\sigma_T)$. Then by Theorem 6.2(ii),

$$\begin{aligned} \|s_j v_T\|^2 &= \|\tfrac{1}{r} v_T + v_{T'}\|^2 \\ &= \tfrac{1}{r^2}\|v_T\|^2 + \|v_{T'}\|^2 \\ \implies \left(1 - \tfrac{1}{r^2}\right)\|v_T\|^2 &= \|v_{T'}\|^2 \end{aligned}$$

using the fact that $s_j v_T = \frac{1}{r} v_T + v_{T'}$. Then in the orthonormal basis

$$\{w_T, w_{T'}\} = \left\{ \frac{1}{\|v_T\|} v_T, \frac{1}{\sqrt{1 - \frac{1}{r^2}}\|v_T\|} v_{T'} \right\}$$

the first line of the formula in Theorem 6.2 becomes

$$s_j w_T = \frac{1}{r} w_T + \sqrt{1 - \frac{1}{r^2}} w_{T'}$$

The case when $\ell(\sigma_{T'}) < \ell(\sigma_T)$ is similar. □

**Definition 6.9.** **Young's Orthogonal Form** for $S^\lambda$ is given by the $w_T$, with

$$s_j w_T = \frac{1}{r} w_T + \sqrt{1 - \frac{1}{r^2}} w_{s_j T}$$

$$s_j w_{s_j T} = -\frac{1}{r} w_{s_j T} + \sqrt{1 - \frac{1}{r^2}} w_T$$

where, if $C(T) = (a_1, \dots, a_n)$, $r = a_{j+1} - a_j$. Thus, with respect to the basis $\{w_T, w_{s_j T}\}$, the Coxeter element $s_j$ is represented by the orthogonal matrix

$$s_j \longmapsto \begin{bmatrix} \frac{1}{r} & \sqrt{1 - \frac{1}{r^2}} \\ \sqrt{1 - \frac{1}{r^2}} & -\frac{1}{r} \end{bmatrix}$$

**Definition 6.10.** The weight $\alpha(T^\lambda)$ of vector $v_{T^\lambda}$ is the maximal weight with respect to the lexicographic order. Call $\alpha(T^\lambda)$ the **highest weight** of $S^\lambda$ and call the vector $v_{T^\lambda}$ the highest weight vector of $S^\lambda$.

**Example 6.11.**

- For $\lambda = (n)$, there is a unique standard tableaux $T = \boxed{1}\boxed{2}\boxed{3}\boxed{\dots}\boxed{n}$ with $C(T) = (0, 1, \dots, n-1)$. Then $s_j w_T = w_T$ for all $1 \leqslant j \leqslant n-1$. We always have that $a_{j+1} = a_j + 1$ so $S^{(n)}$ is trivial.
- $\lambda = (1, 1, \dots, 1)$. Again, there is a unique standard tableaux

$$T = \begin{array}{|c|} \hline 1 \\ \hline 2 \\ \hline 3 \\ \hline \\ \hline n \\ \hline \end{array}$$

  and $C(T) = (0, -1, \dots, -n+1)$. Then $s_j w_T = -w_T$ for all $j$, so $S^\lambda$ is the alternating representation.

- Repeat for $S^{(n-1,1)}$ with the set of standard tableaux

$$T_j = \begin{array}{|c|c|c|c|c|c|c|}\hline 1 & 2 & & j\text{-}1 & j\text{+}1 & & n \\ \hline j \\ \cline{1-1}\end{array}$$

for each $2 \leqslant j \leqslant n$. The content of this vector is

$$C(T) = (0, 1, \dots, j-2, -1, j-1, j, \dots, n-2).$$

# 7 Hook Length Formula

Recall from Definition 1.24 that if $x = (i, j)$ is a box in the Young diagram for $\lambda$, then it defines a hook $\Gamma_x = \{(i, j') \mid j' \geqslant j\} \cup \{(i', j) \mid i' \geqslant i\}$ with hook length $h(x) = h_{ij} = |\Gamma_x|$.

**Theorem 7.1** (Frame, Robinson, Thrall 1954)**.** Let $\lambda = (\lambda_1, \dots, \lambda_k) \vdash n$. Then

$$f_\lambda = \frac{n!}{\prod\limits_{x \in \lambda} h(x)},$$

where $f_\lambda$ is the number of standard $\lambda$-tableaux ($= \dim S^\lambda$).

We'll give a probabalistic "hook-walk" proof due to Greene, Nijenhuis and Wilf from 1979 (which has been described as "cute"). This proof is nice because it actually uses hooks, which previous proofs largely ignored. There are lots of "bijective" proofs of this result based on a technique called the "bijectation" method of Garsic-Milne.

*Proof. (Greene, Nijenhuis, Wilf 1979)* Define

$$F(\lambda) = F(\lambda_1, \dots, \lambda_k) = \begin{cases} \frac{n!}{\prod h_{ij}} & \lambda_1 \geqslant \lambda_2 \geqslant \dots \geqslant \lambda_k \\ 0 & \text{otherwise} \end{cases}$$

For a standard Young tableau $\lambda$, $n$ must appear at a corner (meaning the end of some row and simultaneously the end of a column). Removing this leaves a Young tableau of smaller shape. So Theorem 7.1 follows by induction if we can show that

$$F(\lambda) = \sum_{\alpha=1}^{k} F(\lambda_1, \dots, \lambda_{\alpha-1}, \lambda_\alpha - 1, \lambda_{\alpha+1}, \dots, \lambda_k)$$

where the sum runs over all corners $\alpha$ in the Young tableaux since terms such that $\lambda_{\alpha+1} > \lambda_\alpha - 1$ vanish by definition of $F(\lambda)$. Write

$$F_\alpha := F(\lambda_1, \dots, \lambda_{\alpha-1}, \lambda_\alpha - 1, \lambda_{\alpha+1}, \dots, \lambda_k)$$

for the removal of a corner $\alpha$ from shape $\lambda$.

The idea is to verify that

$$1 = \sum_\alpha \frac{F_\alpha}{F}$$

by using probability, where $F = F(\lambda)$. Here's the procedure

- A box $x = (i, j)$ in the Young diagram for $\lambda$ is chosen at random with probability $1/n$.
- A distinct box $x' = (i', j')$ is chosen at random from among the remaining boxes in the hook $\Gamma_x$ with probability $1/(h(x) - 1)$.
- A new box is chosen at random from the remaining boxes in $\Gamma_{x'}$, and so on, continuing until a corner box $(\alpha, \beta)$ is chosen.
- This completes a single trial. The box $(\alpha, \beta)$ where the process stops is called the **terminal box** of the trial. Note that any corner box can be the terminal box.

Let $p(\alpha, \beta)$ be the probability that a random trial terminates at the box with coordinates $(\alpha, \beta)$.

**Theorem 7.2.** Let $(\alpha, \beta)$ be a corner box. Then $p(\alpha, \beta) = F_\alpha / F$.

*Proof.*

$$\begin{aligned} \frac{F_\alpha}{F} &= \frac{1}{n} \prod_{i=1}^{\alpha-1} \frac{h_{i\beta}}{h_{i\beta} - 1} \prod_{j=1}^{\beta-1} \frac{h_{\alpha j}}{h_{\alpha j} - 1} \\ &= \frac{1}{n} \prod_{i=1}^{\alpha-1} \left(1 + \frac{1}{h_{i\beta} - 1}\right) \prod_{j=1}^{\beta-1} \left(1 + \frac{1}{h_{\alpha j} - 1}\right) \end{aligned} \tag{17}$$

The idea is to interpret each term in (17) as probabilities.

So suppose $\pi \colon (a, b) = (a_1, b_1) \to (a_2, b_2) \to \cdots \to (a_m, b_m) = (\alpha, \beta)$ is a path determined by a trial beginning at $(a, b)$ and ending at $(\alpha, \beta)$. Define the **vertical projection** of $\pi$ as $A = \{a_1, \ldots, a_m\}$ and the **horizontal projection** as $B = \{b_1, \ldots, b_m\}$.

Let $p(A, B \mid a, b)$ be the probability that a random trial which begins at $(a, b)$ has vertical and horizontal projections $A$ and $B$. Claim

**Lemma 7.3.** $p(A, B \mid a, b) = \prod_{\substack{i \in A \\ i \neq a_m}} \frac{1}{h_{i\beta} - 1} \prod_{\substack{j \in B \\ j \neq b_m}} \frac{1}{h_{\alpha j} - 1}$

*Proof Sketch.* Proof by induction on $m$. The base case is easy. For $m > 1$, assume that the statement holds for all $k < m$. Then,

$$\begin{aligned}
p(A, B \mid a, b) &= \frac{1}{h_{ab} - 1}\Big(p(A - a_1, B \mid a_2, b_1) + p(A, B - b_1 \mid a_1, b_2)\Big) \\
&= \frac{1}{h_{ab} - 1}\left(\left(\prod_{\substack{i \in A - a_1 \\ i \neq a_m}} \frac{1}{h_{i\beta} - 1} \prod_{\substack{j \in B \\ j \neq b_m}} \frac{1}{h_{\alpha j} - 1}\right) + \left(\prod_{\substack{i \in A \\ i \neq a_m}} \frac{1}{h_{i\beta} - 1} \prod_{\substack{j \in B - b_1 \\ j \neq b_m}} \frac{1}{h_{\alpha j} - 1}\right)\right) \\
&= \frac{1}{h_{ab} - 1}\left(\prod_{\substack{i \in A - a_1 \\ i \neq a_m}} \frac{1}{h_{i\beta} - 1} \prod_{\substack{j \in B - b_1 \\ j \neq b_m}} \frac{1}{h_{\alpha j} - 1}\right)\left(\frac{1}{h_{\alpha b_1} - 1} + \frac{1}{h_{a_1 \beta} - 1}\right) \\
&= \frac{1}{h_{ab} - 1}\left(\prod_{\substack{i \in A - a_1 \\ i \neq a_m}} \frac{1}{h_{i\beta} - 1} \prod_{\substack{j \in B - b_1 \\ j \neq b_m}} \frac{1}{h_{\alpha j} - 1}\right)\left(\frac{h_{\alpha b_1} - 1 + h_{a_1 \beta} - 1}{(h_{\alpha b_1} - 1)(h_{a_1 \beta} - 1)}\right) \\
&= \frac{1}{h_{ab} - 1}\left(\prod_{\substack{i \in A - a_1 \\ i \neq a_m}} \frac{1}{h_{i\beta} - 1} \prod_{\substack{j \in B - b_1 \\ j \neq b_m}} \frac{1}{h_{\alpha j} - 1}\right)\left(\frac{h_{\alpha \beta} - 1}{(h_{\alpha b_1} - 1)(h_{a_1 \beta} - 1)}\right) \\
&= \prod_{\substack{i \in A \\ i \neq a_m}} \frac{1}{h_{i\beta} - 1} \prod_{\substack{j \in B \\ j \neq b_m}} \frac{1}{h_{\alpha j} - 1} \qquad \square
\end{aligned}$$

Now $p(\alpha, \beta)$ is the sum of the conditional probabilities with respect to the first box chosen. Then, for each such first box, sum over all possible vertical and horizontal projections. Then

$$p(\alpha, \beta) = \frac{1}{n} \sum p(A, B \mid a, b)$$

summed over all $A, B, a, b$ such that $A \subseteq \{1, \ldots, \alpha\}$ and $B \subseteq \{1, \ldots, \beta\}$ and $a = \min A$, $b = \min B$. By Lemma 7.3, this is the same as expanding the products in the right hand side of (17).

This concludes the proof of Theorem 7.2. □

**Corollary 7.4.** $\displaystyle\sum_{\alpha} \frac{F_\alpha}{F} = 1.$

*Proof.* Every trial stops at some terminal box. Therefore, the probabilities $p(\alpha, \beta)$ must all add up to 1. □

This concludes the proof of Theorem 7.1. □

# 8 A bijection that counts

The Robinson-Schenstead-Knuth algorithm was introduced by Robinson with a liberal sprinkling of errors in 1938 and then justified and improved by Schenstead in 1961 and again by Knuth in 1970. It gives a (combinatorial) proof of the identity that

$$\sum_{\lambda \vdash n} f_\lambda^2 = n!.$$

Recall that this says that $|S_n|$ is equal to the number of pairs of standard tableaux of the same shape $\lambda$ as $\lambda$ varies over all partitions of $n$. Denote this bijection by

$$\pi \xleftrightarrow{\text{R-S}} (P, Q)$$

where $\pi \in S_n$ and $P, Q$ are standard tableaux of shape $\lambda$, with $\lambda \vdash n$. The R-S above the arrow is for "Robinson-Schenstead."

So how does this bijection work?

## 8.1 Constructing pairs of tableaux from permutations

Let's first construct a pair of tableaux $(P, Q)$ given a permutation $\pi$. We denote this by $\pi \xmapsto{\text{R-S}} (P, Q)$. Note that the shapes of $P$ and $Q$ need not be the same as the cycle type of $\pi$.

Suppose that $\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ x_1 & x_2 & \dots & x_n \end{pmatrix}$ in 2-line notation. Construct a sequence of tableau pairs

$$(\varnothing, \varnothing) = (P_0, Q_0), (P_1, Q_1), \dots, (P_n, Q_n) = (P, Q) \tag{18}$$

where $x_1, \dots, x_n$ are **inserted** into the $P_j$ and $1, 2, \dots, n$ are **placed** into the $Q_j$ such that the shape of $P_j$ is the shape of $Q_j$ for all $j$.

#### Insertion

Define a **Near Young Tableau** (NYT) as an array with distinct entries whose rows/columns increase (so it is a SYT if elements are in the set $\{1, 2, \dots, n\}$). Given a NYT, $P$,

- Let $x \notin P$
- Let $P_{ij}$ be the entry in row $i$, column $j$ of $P$.
- row insert $x$ into $P$ as follows:
  (a) let $y$ be the least integer such that $P_{1y} > x$
  (b) if no such $y$ exists (means that all elements of the first row are less than $x$), then place $x$ at the end of the first row. Insertion process stops and denote the resulting NYT as $P \leftarrow x$.

(c) if $y$ does exist, replace $P_{1y}$ by $x$. The element $x$ then **bumps** $x' = P_{1y}$ into the second row, i.e. insert $x'$ into the second row of $P$ by the above insertion rule. Either $x'$ is inserted at the end of the second row, or else it bumps an element $x''$ into the third row.

(d) Continue until the element is inserted at the end of a row. Denote the resulting array by $P \leftarrow x$.

**Example 8.1.** Suppose we are going to insert $x = 4$ into

$$P = \begin{array}{|c|c|c|} \hline 1 & 3 & 5 \\ \hline 2 & 7 \\ \cline{1-2} 6 & 9 \\ \cline{1-2} 8 \\ \cline{1-1} \end{array}$$

We first put 4 into the position that 5 occupies in $P$, and so bump 5 into the second row,

$$\begin{array}{|c|c|c|} \hline 1 & 3 & 4 \\ \hline 2 & 7 \\ \cline{1-2} 6 & 9 \\ \cline{1-2} 8 \\ \cline{1-1} \end{array} \leftarrow 5$$

but then 5 bumps 7 into the third row

$$\begin{array}{|c|c|c|} \hline 1 & 3 & 4 \\ \hline 2 & 5 \\ \cline{1-2} 6 & 9 \\ \cline{1-2} 8 \\ \cline{1-1} \end{array} \leftarrow 7$$

and 7 bumps 9 into the third row

$$\begin{array}{|c|c|c|} \hline 1 & 3 & 4 \\ \hline 2 & 5 \\ \cline{1-2} 6 & 7 \\ \cline{1-2} 8 \\ \cline{1-1} \end{array} \leftarrow 9$$

and inserting 9 into the last row just places 9 at the end of the row.

$$\begin{array}{|c|c|c|} \hline 1 & 3 & 4 \\ \hline 2 & 5 \\ \cline{1-2} 6 & 7 \\ \cline{1-2} 8 & 9 \\ \cline{1-2} \end{array}$$

This is $P \leftarrow 4$.

Suppose the result of row insertion of $x$ into $p$ gives $P' = r_x(P)$. Note the insertion rules force $P' = r_x(P)$ to have increasing rows and columns.

### Placement

If $Q$ is a NYT of shape $\mu$ and $(i, j)$ are the coordinates of an addable box for $\mu$, then if $k$ is larger than every element of $Q$, then to place $k$ in $Q$ in the $(i, j)$ box, set $Q_{ij} = k$. The new array must still be a NYT.

**Example 8.2.** If

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

place $k = 9$ in $(i, j) = (2, 3)$ to produce

$$\begin{array}{|c|c|c|} \hline 1 & 2 & 5 \\ \hline 4 & 7 & 9 \\ \hline 6 \\ \cline{1-1} 8 \\ \cline{1-1} \end{array}$$

To construct (18) from the permutation $\pi$,

- start with $(P_0, Q_0) = (\varnothing, \varnothing)$
- assuming $(P_{k-1}, Q_{k-1})$ is constructed, define $P_k = r_{x_k}(P_{k-1})$, $Q_k$ = place $k$ into $Q_{k-1}$ at box $(i, j)$ where the insertion in $P$ terminates.
- The definition of $Q_k$ ensures that the shapes of $P_k$ and $Q_k$ are the same at teach step for all $k$.

**Definition 8.3.** $P = P_n$ is the **insertion tableau** of $\pi$, written $P(\pi)$, and $Q = Q_n$ is the **recording tableau** of $\pi$, written $Q(\pi)$.

**Example 8.4.** Let $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 2 & 7 & 3 & 6 & 1 & 5 \end{pmatrix} \in S_7$

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

**Theorem 8.5** (Robinson-Schenstead Correspondence)**.** The map $\pi \xmapsto{\text{R-S}} (P, Q)$ is a bijection between elements of $S_n$ and pairs of standard tableaux of the same shape $\lambda$, where $\lambda \vdash n$.

**Remark 8.6.** Schenstead is a weird man. He's a physicist, without a faculty position, who makes his money by inventing board games. He invented games called ∗Star, Star, and Y. In 1995, he changed his first name from Craige to Ea, because people kept forgetting the 'e' on the end of his name. Ea is the Babylonian name for the Sumerian god Enki. He then changed it again to Ea Ea because he was afraid of Y2K computer errors or something.

*Proof of Theorem 8.5.* We only need to check that there's an inverse to the procedure, which we $(P, Q) \xmapsto{\text{S}^{-1}\text{-R}^{-1}} \pi$. The idea is to reverse the algorithm step by step.

Given $(P, Q)$, how do we recover $\pi$ uniquely? And we also need to find $\pi$ for any $(P, Q)$. The position occupied by $n$ in $Q$ is the last position to be occupied in the insertion process. Suppose $k$ occupies this position in $P$. It was bumped into this position by some element $j$ in the row above $k$ that is currently the largest of its row less than $k$. Hence can "inverse bump" $k$ into the position occupied by $j$, and now inverse bump $j$ into the row above it by the same procedure. Eventually an element will be placed in the first row, inverse bumping another element $t$ out of the tableau altogether. Thus $t$ was the last element of $P$ to be inserted, i.e. if $\pi(i) = x_i$, then $x_n = t$.

Now locate the position occupied by $n-1$ in $Q_{n-1}$ and repeat the procedure in $P_{n-1}$, obtaining $x_{n-1}$. Continuing in this way, we uniquely construct $\pi$ one element at a time from right to left such that $\pi \mapsto (P, Q)$. ☐

## 8.2 Consequences and Properties of RS Algorithm

**Theorem 8.7.** $\sum_{\lambda \vdash n} f_\lambda^2 = n!$, where $f_\lambda = \#\mathrm{SYT}(\lambda) = \dim S^\lambda$.

**Remark 8.8.** Can define **column insertion** of $x$ into $P$ by replacing "row" by "column" as required in Definition 8.3 and Theorem 8.5.

Given $\pi$, denote by $\pi^{\text{rev}}$ it's **reversal** the permutation $\pi^r(i) = \pi(n+1-i)$. Ea Ea (Schenstead) proved if $P(\pi) = P$ then $P(\pi^{\text{rev}}) = P^T$, (see Sagan 3.2.3).

The recording tableau of $\pi^r$ is characterized by Schützenberger's "operation of evacuation," (see Sagan 3.9).

**Definition 8.9.** Given $m$, a sequence $\mu = (\mu_1, \ldots, \mu_\ell)$ of nonnegative integers is called a **composition** if $\sum_i \mu_i = n$.

**Definition 8.10.** Let $\lambda$ be a partition. A **semi-standard Young tableau** (SSYT) of shape $\lambda$ is an array $T = (T_{ij})$ of positive integers of shape $\lambda$ (so $1 \leqslant i \leqslant \ell(\lambda)$ and $1 \leqslant j \leqslant \lambda_i$) that is weakly increasing in every row and strictly increasing in every column.

$T$ has **weight/type** $\alpha = (\alpha_1, \alpha_2, \ldots)$ if $T$ has $\alpha_i = \alpha_i(T)$ entries equal to $i$. For an SSYT of type $\alpha$, we write $x^T = x^\alpha = x_1^{\alpha_1(T)} x_2^{\alpha_2(T)} \cdots$.

**Example 8.11.**

$$T = \begin{array}{|c|c|c|c|c|c|} \hline 1 & 1 & 1 & 3 & 4 & 4 \\ \hline 2 & 4 & 4 & 5 & 5 \\ \cline{1-5} 5 & 5 & 7 \\ \cline{1-3} 6 & 9 & 9 \\ \cline{1-3} \end{array}$$

is a SSYT of shape $(6, 5, 3, 3)$. It has type $(3, 1, 1, 4, 4, 1, 1, 0, 2)$, and

$$x^\alpha = x^T = x_1^3 x_2 x_3 x_4^4 x_5^4 x_5^4 x_6 x_7 x_9^2.$$

**Definition 8.12.** The **Schur function** $s_\lambda$ is defined by $s_\lambda(x) = \sum_T x^T$ where the sum is over all SSYT $T$ of shape $\lambda$.

Let $K_{\lambda\alpha}$ be the number of SSYT of shape $\lambda$ and type $\alpha$. These are the **Kostka numbers**; clearly

$$s_\lambda = \sum_\alpha K_{\lambda\alpha} x^\alpha$$

summed over all compositions $\alpha$ of $n$, and $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots$.

**Example 8.13.**

(1) $s_{(1)} = x_1 + x_2 + x_3 + \ldots$

(2) $s_{(1^k)} = \sum_{i_1 < i_2 < \ldots < i_k} x_{i_1} x_{i_2} \cdots x_{i_k}$ is the $k$-th elementary symmetric function

(3) $s_{(k)} = \sum_{i_1 \leqslant i_2 \leqslant \ldots \leqslant i_k} x_{i_1} x_{i_2} \cdots x_{i_k}$ is the $k$-th complete homogeneous symmetric function

(4) For a SSYT $T$ of shape $(2, 1)$,

$$T\colon \begin{array}{|c|c|} \hline 1 & 1 \\ \hline 2 \\ \cline{1-1} \end{array}, \quad \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 2 \\ \cline{1-1} \end{array}, \quad \begin{array}{|c|c|} \hline 1 & 1 \\ \hline 3 \\ \cline{1-1} \end{array}, \quad \begin{array}{|c|c|} \hline 1 & 3 \\ \hline 3 \\ \cline{1-1} \end{array}, \quad \ldots, \quad \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array}, \quad \begin{array}{|c|c|} \hline 1 & 3 \\ \hline 2 \\ \cline{1-1} \end{array}, \quad \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 4 \\ \cline{1-1} \end{array}, \quad \begin{array}{|c|c|} \hline 1 & 4 \\ \hline 2 \\ \cline{1-1} \end{array}, \ldots$$

we have that

$$s_{(2,1)} = x_1^2 x_2 + x_1 x_2^2 + x_1^2 x_3 + x_1 x_3^2 + \ldots + 2x_1 x_2 x_3 + 2x_1 x_2 x_4 + \ldots$$

**Lemma 8.14.** The function $s_\lambda$ is symmetric with respect to all possible permutations of the $x_i$.

*Proof.* One way uses representation theory:

$$s_\lambda = \sum_\alpha K_{\lambda\alpha} x^\alpha.$$

summed over all compositions $\alpha$ of $n$. So it would be enough to show that $K_{\lambda\alpha} = K_{\lambda\tilde{\alpha}}$ for all possible rearrangements $\tilde{\alpha}$ of $\alpha$. This uses Young's rule (see Fact 8.16, but that's not actually the Young's rule this refers to).

We give another proof. We show that $s_i \cdot s_\lambda(x) = s_\lambda(x)$ for each adjacent transposition $s_i = (i, i+1)$. Define an involution on SSYT of shape $\lambda$, denoted

$T \to T'$, as follows. We construct this involution such that the number of $i$'s and the number of $(i+1)$'s are exchanged when passing from $T$ to $T'$ (and all the other multiplicities stay the same).

Given $T$, each column contains either an $i, i+1$ pair; exactly one of $i$ or $i+1$; or neither. Call the pairs **fixed** and all other occurrences of $i, i+1$ **free**.

In each row, switch the number of free $i$'s and $(i+1)$'s i.e. if the row consists of $k$ free $i$'s followed by $\ell$ free $(i+1)$'s; then replace them by $\ell$ free $i$'s followed by $k$ free $(i+1)$'s. (See Example 8.15). We call the new SSYT $T'$.

$T'$ is genuinely an SSYT by the definition of free. Since the fixed $i$'s and $(i+1)$'s come in pairs, this map has the desired exchange property. Clearly this is an involution. □

**Example 8.15.**

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

The 2's and 3's in columns 2 through 4 and 7 through 10 are free.

$$T' = \begin{array}{|c|c|c|c|c|c|c|c|c|c|} \hline 1 & 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 \\ \hline 2 & 2 & 2 & 3 & 3 & 3 \\ \cline{1-6} 3 \\ \cline{1-1} \end{array}$$

**Fact 8.16** (Young's Rule, according to Google)**.** A rule for calculating the dose of medicine correct for a child by adding 12 to the child's age, dividing the sum by the child's age, then dividing the adult dose by the figure obtained.

## 8.3 Knuth's generalization of the R-S Algorithm

Instead of starting with a permutation $\pi \in S_n$, begin with some $r \times s$ matrix $A = (a_{ij})$ of nonnegative integers only finitely many nonzero. Stanley calls these **$\mathbb{N}$-matrices of finite support**.

Associate with $A$ a **generalized permutation** (GP)

$$w_A = \begin{pmatrix} i_1 & i_2 & \cdots & i_m \\ j_1 & j_2 & \cdots & j_m \end{pmatrix}$$

such that

(1) $i_1 \leqslant i_1 \leqslant \ldots \leqslant i_m$

(2) $i_r = i_s \ (r \leqslant s) \implies j_r \leqslant j_s$.

(3) for each pair $(i, j)$ there is exactly $a_{ij}$ values of $r$ such that $(i_r, j_r) = (i, j)$.

**Example 8.17.**

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \\ 1 & 1 & 0 \end{pmatrix} \longleftrightarrow w_A = \begin{pmatrix} 1 & 1 & 1 & 2 & 2 & 3 & 3 \\ 1 & 3 & 3 & 2 & 2 & 1 & 2 \end{pmatrix}$$

The Robinson-Schenstead-Knuth (RSK) algorithm maps $A$ (or $w_A$) to pairs $(P, Q)$ of SSYT of the same shape, where $P$ is filled by $j_1, j_2, \ldots$ and $Q$ is filled by $i_1, i_2, \ldots$.

Let $w_A$ be a given GP. We set $(P(0), Q(0)) = (\varnothing, \varnothing)$, and if for some $t < m$, $(P(t), Q(t))$ is defined, then let

(i) $P(t+1) = P(t) \leftarrow j_{t+1}$

(ii) $Q(t+1)$ obtained from $Q(t)$ by placing $i_{t+1}$ (and leaving all other parts of $Q(t)$ unchanged) such that $P(t+1)$, $Q(t+1)$ have the same shape.

The process ends at $(P(m), Q(m))$. Define $(P, Q) = (P(m), Q(m))$. The correspondence $w_A \xmapsto{\text{RSK}} (P, Q)$ is the **RSK algorithm** and $P$ is called the **insertion tableau**, $Q$ the **recording tableau**.

**Theorem 8.18.** There is a bijection between $\mathbb{N}$-matrices $A = (a_{ij})$ of finite support and ordered pairs $(P, Q)$ of SSYT of the same shape. The number $j$ occurs in $P$ exactly $\sum_i a_{ij}$ times, and $i$ occurs in $Q$ exactly $\sum_j a_{ij}$ times. (This means that the type of $P$ is $\mathrm{col}(A)$ and the type of $Q$ is $\mathrm{row}(A)$, where $\mathrm{col}(A)$ is the vector of all the sums of columns of $A$, and similarly for $\mathrm{row}(A)$).

**Example 8.19.** Using the generalized permutation from Example 8.17, the RSK algorithm generates the following SSYT.

| $P(i)$ | $Q(i)$ |
|---|---|
| 1 | 1 |
| 1 3 | 1 1 |
| 1 3 3 | 1 1 1 |
| 1 2 3<br>3 | 1 1 1<br>2 |
| 1 2 2<br>3 3 | 1 1 1<br>2 2 |
| 1 1 2<br>2 3<br>3 | 1 1 1<br>2 2<br>3 |
| 1 1 2 2<br>2 3<br>3 | 1 1 1 3<br>2 2<br>3 |

## 8.4 Cauchy Identity

**Theorem 8.20** (Littlewood-Cauchy 1950)**.**

$$\sum_{\lambda} s_\lambda(x) s_\lambda(y) = \prod_{i,j \geqslant 1} \frac{1}{1 - x_i y_j}$$

*Proof of Theorem 8.20.* Write

$$\prod_{i,j}(1-x_iy_j)^{-1} = \prod_{i,j\geqslant 1}\left(\sum_{a_{ij}\geqslant 0}(x_iy_j)^{a_{ij}}\right) \tag{19}$$

The term $x^\alpha y^\beta$ obtained by choosing a matrix $M^T = (a_{ij})^T$ of finite support with $\mathrm{row}(M) = \alpha$ and $\mathrm{col}(M) = \beta$, where $\mathrm{row}(M)$ is the sum across the rows of $M$, and col($M$) is the sum across the columns. Hence, the coefficient of $x^\alpha y^\beta$ is $N_{\alpha\beta}$, which is the number of such $\mathbb{N}$-matrices $M$ of finite support with $\mathrm{row}(M) = \alpha$, $\mathrm{col}(M) = \beta$. But the coefficient of $x^\alpha y^\beta$ in

$$\sum_\lambda s_\lambda(x)s_\lambda(y)$$

is the number of pairs $(P, Q)$ of SSYT of the same shape $\lambda$ such that the type of $P$ is $\alpha$ and the type of $Q$ is $\beta$. The RSK algorithm (Theorem 8.18) sets up a bijection between these matrices $M$ and pairs of tableaux $(P, Q)$. □

**Remark 8.21.** Stanley uses this to deduce that the Schur functions form an orthonormal basis for the algebra $\Lambda[e_1, e_2, \ldots]$ of symmetric functions (generated by the elementary symmetric functions $e_i$) (Stanley, 7.12.2).

**Remark 8.22** (Exam Stuff)**.** There are six questions on the exam. You're supposed to do four of them. Do not believe the reputation I have for setting impossible exams. If you've been to class, written it all down, and done the examples sheets, it's "dead easy" (. . . ).

# 9 Extra Material

We could define column insertion of $x$ into $P$ by replacing row by column as appropriate. If column insertion of $x$ into $P$ produces $P'$, write $c_x(P) = P'$ (before we used $r_x(P) = P'$). In fact row and column operators commute (Sagan 3.2). For a NYT $P$, distinct $x, y \notin P$, $c_y r_x(P) = r_x c_y(P)$ (Sagan 3.2.2).

More generally, we could ask about the effects that various changes on the permutation $\pi$ have on a pair $(P, Q)$, for example, Schutzenberger's result that if $P(\pi) = P$, then $P(\pi^{\mathrm{rev}}) = P^T$, where $\pi^{\mathrm{rev}}(i) = \pi(n + 1 - i)$ and $P^T$ denotes the transpose of $P$.

**Theorem 9.1** (Symmetry Theorem)**.** Let $A$ be an $\mathbb{N}$-matrix of finite support and suppose that $A \xmapsto{\mathrm{RSK}} (P, Q)$. Then $A^T \xmapsto{RSK} (Q, P)$. So $A$ is symmetric if and only if $P = Q$. In particular, for the RS correspondence, if $\pi \xmapsto{\mathrm{RS}} (P, Q)$, where $P, Q$ are SYT, then $\pi^{-1} \xmapsto{\mathrm{RS}} (Q, P)$.

**Corollary 9.2.** If $A = A^T$, then $A \xmapsto{\mathrm{RSK}} (P, P)$ and $\alpha = (\alpha_1, \alpha_2, \ldots)$ is a composition for some number, $\alpha_i \in \mathbb{N}_0$ and $\sum \alpha_i < \infty$. Then $A \leftrightarrow P$ establishes a bijection between symmetric $\mathbb{N}$-matrices of finite support with $\mathrm{row}(A) = \alpha$ and SSYT of type $\alpha$.

**Corollary 9.3.** $\sum_{\lambda \vdash n} f_\lambda = \#\{w \in S_n \mid w^2 = 1\} = \#(\text{involutions in } S_n)$

*Proof.* If $w \in S_n$, $w \overset{RS}{\longmapsto} (P, Q)$, where $P, Q$ are SYT of same shape $\lambda$. The permutation matrix corresponding to $w$ is symmetric if and only if $w^2 = 1$. This is the case if and only if $P = Q$. $\square$

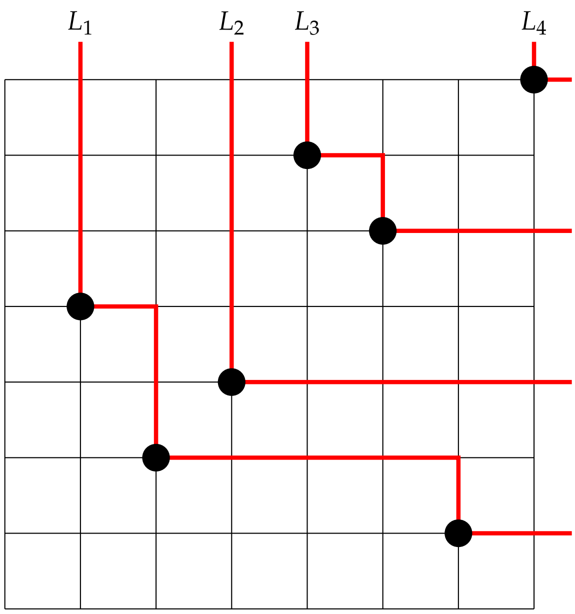
## 9.1 Viennot's geometric construction

(Sagan 3.6, Stanley 7.13). A permutation $\pi(i) = x_i$ can be represented by a box, with coordinates $(i, x_i)$. Light shines at $(0, 0)$, so each box casts a shadow with the boundaries parallel to coordinate axes. Consider the points of the permutation that are in the shadow of no other point.

Now we draw some shadow lines. The first shadow line, $L_1$ is the boundary of the combined shadows of these boxes that are not in the shadow of any other point (a broken line comprising line segments and exactly one horizontal and vertical ray).

The second shadow line $L_2$ is drawn by removing the boxes in $L_1$ and repeating the procedure.

**Example 9.4.** $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 2 & 3 & 6 & 5 & 1 & 7 \end{pmatrix}$




The points that are not in the shadow of any other point are $(1, 4)$, $(2, 2)$ and $(6, 1)$.

**Definition 9.5.** Given a permutation $\pi$, form shadow lines $L_1, L_2, \ldots$ as follows: assuming $L_1, L_2, \ldots, L_{i-1}$ have been constructed, remove all boxes on these lines. Let $L_i$ be the boundary of the shadow of the remaining boxes.

Continuing the shadow diagrams from last time. The $x$-coordinate of $L_i$ is denoted $x_{L_i}$ and defined as the $x$-coordinate of $L_i$'s vertical ray, and similarly the $y$-coordinate of $L_i$ is denoted $y_{L_i}$ and defined as the $y$-coordinate of $L_i$'s horizontal ray.

The shadow lines define the **shadow diagram** of $\pi$.

**Example 9.6.** Continuing Example 9.4, $x_{L_1} = 1, x_{L_2} = 3, x_{L_3} = 4, x_{L_4} = 7$ and $y_{L_1} = 1, y_{L_2} = 3, y_{L_3} = 4, y_{L_4} = 7$.

We can compare the coordinates of she shadow lines with their first row of the RS tableaux:

$$P(\pi) = \begin{array}{|c|c|c|c|}\hline 1 & 3 & 5 & 7 \\ \hline 2 & 6 \\ \cline{1-2} 4 \\ \cline{1-1}\end{array} \qquad Q(\pi) = \begin{array}{|c|c|c|c|}\hline 1 & 3 & 4 & 7 \\ \hline 2 & 5 \\ \cline{1-2} 6 \\ \cline{1-1}\end{array}$$

Is this a coincidence? Is $P_{1j} = y_{L_j}$ and $Q_{1y} = x_{L_j}$?

**Lemma 9.7.** Let the shadow diagram of $\pi$ be constructed as above, where $\pi(i) = x_i$. Suppose the vertical line $x = k$ intersects $i$ shadow lines. Let $y_j$ be the $y$-coordinate of the lowest point of intersection with $L_j$. The first row of $P_k = P(x_1, \ldots, x_k)$ is $R_1 = y_1 \ldots y_i$.

*Proof.* Proof by induction on $k$. Assume that this is true for $x = k$. Consider $x = k + 1$. There are two cases

(a) $x_{k+1} > y_i$.

Then the box $(k + 1, x_{k+1})$ starts a new shadow line. So none of the values $y_1, \ldots, y_i$ change, and obtain a new intersection $y_{i+1} = x_{k+1}$. Hence the $(k + 1)$-st intersection causes $x_{k+1}$ to be at the end of the first row (without bumping another element). So the result holds.

(b) $y_1 < \ldots < y_{j-1} < x_{k+1} < y_j < \ldots < y_i$.

Then $(k + 1, x_{k+1})$ is added to $L_j$. So the lowest coordinate on $L_j$ becomes $y'_j = x_{k+1}$, and al the other $y$ values stay the same. Now the first row of $P_{k+1}$ is $y_1 \ldots y_{j-1} y'_j y_j \ldots y_i$, as predicted. □

This lemma says that the shadow diagram is a timeline recording of the construction of the tableaux $P(\pi)$ and $Q(\pi)$, reading left-to-right. At the $k$-th stage, the line $x = k$ intersects one shadow line in a ray or a line segment, and all the rest in single points. In terms of the first row of $P_k$, a ray corresponds to placing an element at the end, a line segment corresponds to bumping an element, and points corresponds to elements that are unchanged.

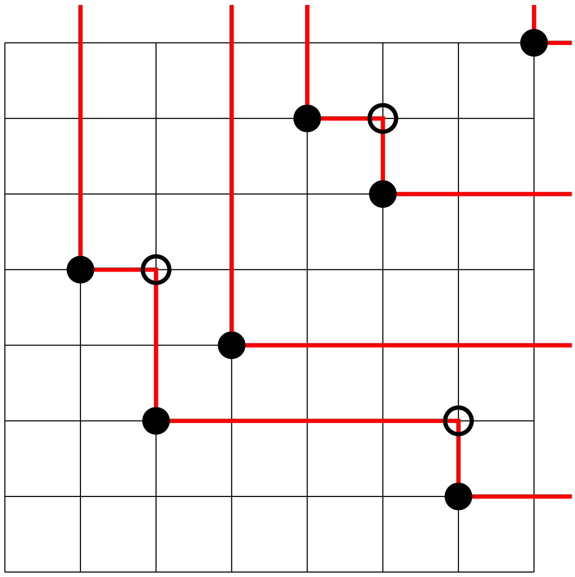
**Corollary 9.8** (Viennot 1976)**.** If $\pi$ has RS tableaux $(P, Q)$, and shadow lines $L_j$, then for all $j$, $P_{1,j} = y_{L_j}$ and $Q_{1,j} = x_{L_j}$.

*Proof.* The statement for $P$ is the case $k = n$ of Lemma 9.7. For $Q$, entry $k$ is added to $Q$ in box $(1, j)$ when $x_k >$ every element of the first row of $P_{k-1}$. The proof of Lemma 9.7 shows this happens precisely when the line $x = k$ intersects $L_j$ in a vertical ray, that is, $y_{L_j} = k = Q_{1,j}$. □

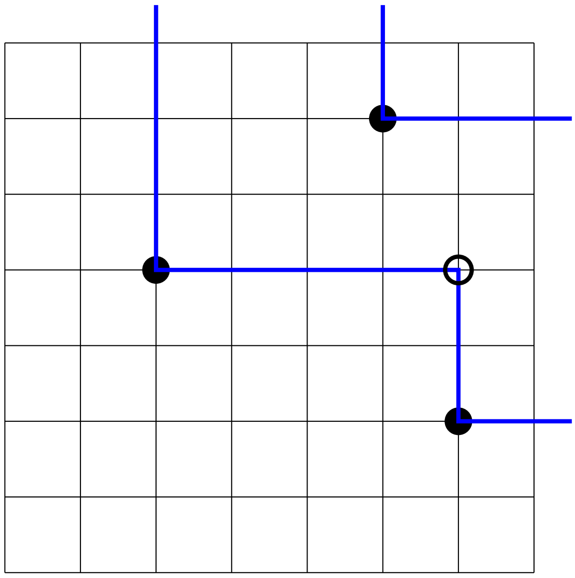
## 9.2 The $i$-skeleta

Focus on the NE corners of the shadow lines. If a corner has coordinate $(k, x')$, then by Lemma 9.7, $x'$ is displaced from the first row of $P_{k-1}$ by insertion of $x_k$. So circles correspond to elements inserted into the second row in the construction of $P$. So we can get the rest of tableaux by iterating the shadow diagram construction of $P$.

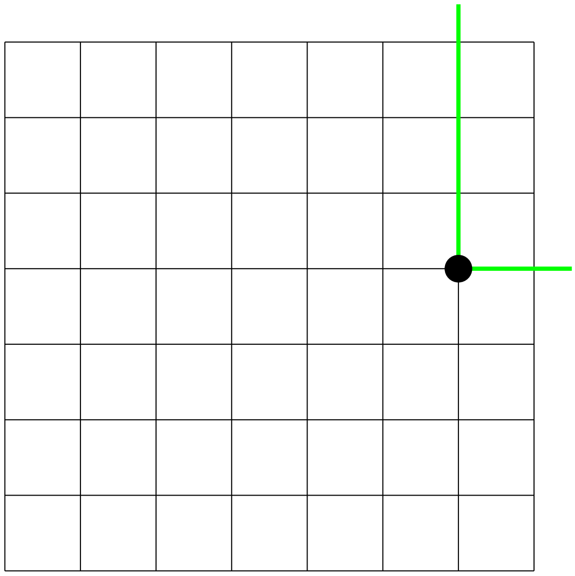
**Example 9.9.** Continuing Example 9.4 again. The NE corners of shadow lines are circled.



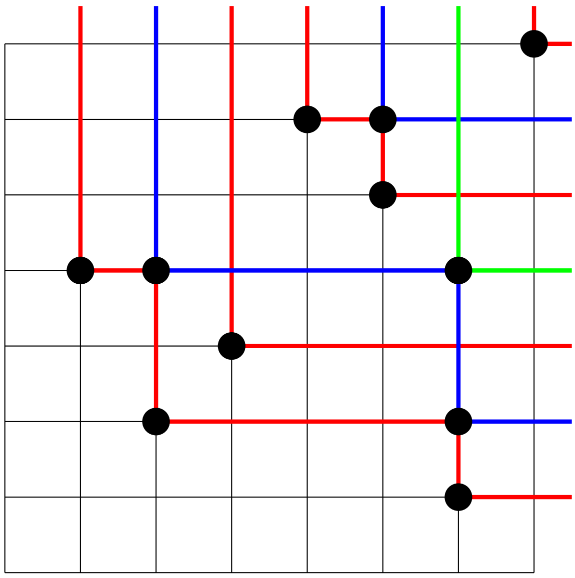
The second iteration of the shadow diagram consists of the blue lines.



We can iterate again to get the final stage



Drawing all of these lines together, we get:



Then notice that

$$P(\pi) = \begin{array}{|c|c|c|c|}\hline 1 & 3 & 5 & 7 \\ \hline 2 & 6 \\ \cline{1-2} 4 \\ \cline{1-1} \end{array} \qquad Q(\pi) = \begin{array}{|c|c|c|c|}\hline 1 & 3 & 4 & 7 \\ \hline 2 & 5 \\ \cline{1-2} 6 \\ \cline{1-1} \end{array}$$

where the first row is the $x$-coordinates of the red lines, and the second row are the $y$-coordinates of the $Y$-coordinates of the blue lines, and the last row is the $Y$-coordinate of the green line. Similarly, the first row of $Q(\pi)$ consists of

$x$-coordinates of the red lines, the second row is $x$-coordinates of the blue lines, and the third row is the $x$-coordinate of the green line.

**Definition 9.10.** The ***i*-th skeleton** of $\pi$, denoted $\pi^{(i)}$, is defined inductively by $\pi^{(1)} = \pi$ and

$$\pi^{(i)} = \begin{pmatrix} k_1 & \dots & k_m \\ \ell_1 & \dots & \ell_m \end{pmatrix}$$

where $(k_1, \ell_1), \dots, (k_m, \ell_m)$ are coordinates of the northeast corners of the shadow diagram of $\pi^{(i-1)}$. Shadow lines for $\pi^{(i)}$ are denoted $L_j^{(i)}$.

**Proposition 9.11** (Viennot)**.** Suppose $\pi \xmapsto{\text{RS}} (P, Q)$. Then $\pi^{(i)}$ is the permutation such that $\pi^{(i)} \xmapsto{\text{RS}} (P^{(i)}, Q^{(i)}$ where $P^{(i)}$ (resp. $Q^{(i)}$) comprises row $i$ and below of $P$ (resp. $Q$).

**Theorem 9.12** (Schutzenberger)**.** Given $\pi \in S_n$, then $P(\pi^{-1}) = Q(\pi)$ and $Q(\pi^{-1}) = P(\pi)$.

*Proof.* Taking the inverse of a permutation corresponds to reflecting the shadow diagram in the line $y = x$. Then apply Proposition 9.11. □