NOTES FOR MATH 740 (SYMMETRIC FUNCTIONS)

STEVEN V SAM

Contents

The first part of these notes follows Stanley's exposition in [Sta, Chapter 7]. However, our perspective will be from the character theory of the general linear group rather than a combinatorial one. So we insert remarks throughout that connect the relevant results to representation theory.

1. Definition and motivation

1.1. **Definitions.** Let $x_1, \dots, x_n$ be a finite set of indeterminates. The symmetric group on n letters is Σ_n. It acts on $\mathbf{Z}[x_1, \dots, x_n]$, the ring of polynomials in n variables and integer coefficients, by substitution of variables:

$$\sigma \cdot f(x_1, \dots, x_n) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

The **ring of symmetric polynomials** is the set of fixed polynomials:

$$\Lambda(n) := \{f \in \mathbf{Z}[x_1, \dots, x_n] \mid \sigma \cdot f = f \text{ for all } \sigma \in \Sigma_n\}.$$

This is a subring of $\mathbf{Z}[x_1, \dots, x_n]$.

We will also treat the case $n = \infty$. Let $x_1, x_2, \dots$ be a countably infinite set of indeterminates. Let Σ_∞ be the group of all permutations of $\{1, 2, \dots\}$. Let R be the ring of power series in $x_1, x_2, \dots$ of *bounded degree*. Hence, elements of R can be infinite sums, but only in a finite number of degrees.

Write $\pi_n \colon \Lambda \to \Lambda(n)$ for the homomorphism which sets $x_{n+1} = x_{n+2} = \cdots = 0$.

Date: April 27, 2017.

Remark 1.1.1. (For those familiar with inverse limits.) There is a ring homomorphism $\pi_{n+1,n}\colon \Lambda(n+1) \to \Lambda(n)$ obtained by setting $x_{n+1} = 0$. Furthermore, $\Lambda(n) = \bigoplus_{d \ge 0} \Lambda(n)_d$ where $\Lambda(n)_d$ is the subgroup of homogeneous symmetric polynomials of degree d. The map $\pi_{n+1,n}$ restricts to a map $\Lambda(n+1)_d \to \Lambda(n)_d$; set

$$\Lambda_d = \varprojlim_n \Lambda(n)_d.$$

Then

$$\Lambda = \bigoplus_{d \ge 0} \Lambda_d.$$

Note that we aren't saying that Λ is the inverse limit of the $\Lambda(n)$; the latter object includes infinite sums of unbounded degree. $\square$

Then Σ_∞ acts on R, and we define the **ring of symmetric functions**

$$\Lambda := \{f \in R \mid \sigma \cdot f = f \text{ for all } \sigma \in \Sigma_\infty\}.$$

Again, this is a subring of R.

Example 1.1.2. Here are some basic examples of elements in Λ (we will study them more soon):

$$\begin{aligned} p_k &:= \sum_{i \ge 1} x_i^k \\ e_k &:= \sum_{i_1 < i_2 < \cdots < i_k} x_{i_1} x_{i_2} \cdots x_{i_k} \\ h_k &:= \sum_{i_1 \le i_2 \le \cdots \le i_k} x_{i_1} x_{i_2} \cdots x_{i_k}. \end{aligned}$$

$\square$

Sometimes, we want to work with rational coefficients instead of integer coefficients. In that case, we'll write $\Lambda_{\mathbf{Q}}$ or $\Lambda(n)_{\mathbf{Q}}$ to denote the appropriate rings.

The main topic of this course is the study of Λ and some of its bases, and how they interact with representation theory (of symmetric groups and general linear groups) and algebraic geometry (Schubert calculus, i.e., intersection theory on Grassmannians).

1.2. Polynomial representations of general linear groups.

Let $\mathbf{GL}_n(\mathbf{C})$ denote the group of invertible $n \times n$ complex matrices.

A **polynomial representation** of $\mathbf{GL}_n(\mathbf{C})$ is a homomorphism $\rho\colon \mathbf{GL}_n(\mathbf{C}) \to \mathbf{GL}(V)$ where V is a $\mathbf{C}$-vector space, and the entries of ρ can be expressed in terms of polynomials (as soon as we pick a basis for V).

A simple example is the identity map $\rho\colon \mathbf{GL}_n(\mathbf{C}) \to \mathbf{GL}_n(\mathbf{C})$. Slightly more sophisticated is $\rho\colon \mathbf{GL}_2(\mathbf{C}) \to \mathbf{GL}(\mathrm{Sym}^2(\mathbf{C}^2))$ where $\mathrm{Sym}^2(\mathbf{C}^2)$ is the space of degree 2 polynomials in x, y (which is a basis for $\mathbf{C}^2$). The homomorphism can be defined by linear change of coordinates, i.e.,

$$\rho(g)(ax^2 + bxy + cy^2) = a(gx)^2 + b(gx)(gy) + c(gy)^2.$$

If we pick the basis x^2, xy, y^2 for $\mathrm{Sym}^2(\mathbf{C}^2)$, this can be written in coordinates as

$$\mathbf{GL}_2(\mathbf{C}) \to \mathbf{GL}_3(\mathbf{C})$$

$$\begin{pmatrix} g_{1,1} & g_{1,2} \\ g_{2,1} & g_{2,2} \end{pmatrix} \mapsto \begin{pmatrix} g_{1,1}^2 & g_{1,1}g_{1,2} & g_{1,2}^2 \\ 2g_{1,1}g_{2,1} & g_{1,1}g_{2,2} + g_{1,2}g_{2,1} & 2g_{1,2}g_{2,2} \\ g_{2,1}^2 & g_{2,1}g_{2,2} & g_{2,2}^2 \end{pmatrix}. \tag{1.2.1}$$

More generally, we can define $\rho\colon \mathbf{GL}_n(\mathbf{C}) \to \mathbf{GL}(\mathrm{Sym}^d(\mathbf{C}^n))$ for any n, d. Another important example uses exterior powers instead of symmetric powers, so we have $\rho\colon \mathbf{GL}_n(\mathbf{C}) \to \mathbf{GL}(\wedge^d(\mathbf{C}^n))$.

An important invariant of a polynomial representation ρ is its **character**: define

$$\mathrm{char}(\rho)(x_1, \dots, x_n) := \mathrm{Tr}(\rho(\mathrm{diag}(x_1, \dots, x_n))),$$

where $\mathrm{diag}(x_1, \dots, x_n)$ is the diagonal matrix with entries $x_1, \dots, x_n$ and Tr denotes trace.

Lemma 1.2.2. $\mathrm{char}(\rho)(x_1, \dots, x_n) \in \Lambda(n)$.

Proof. Each permutation $\sigma \in \Sigma_n$ corresponds to a permutation matrix $M(\sigma)$: this is the matrix with a 1 in row $\sigma(i)$ and column i for $i = 1, \dots, n$ and 0's everywhere else. Then

$$M(\sigma)^{-1}\mathrm{diag}(x_1, \dots, x_n)M(\sigma) = \mathrm{diag}(x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

Now use that the trace of a matrix is invariant under conjugation:

$$\begin{aligned} \mathrm{char}(\rho)(x_1, \dots, x_n) &= \mathrm{Tr}(\rho(\mathrm{diag}(x_1, \dots, x_n))) \\ &= \mathrm{Tr}(\rho(M(\sigma))^{-1}\rho(\mathrm{diag}(x_1, \dots, x_n))\rho(M(\sigma))) \\ &= \mathrm{Tr}(\rho(M(\sigma)^{-1}\mathrm{diag}(x_1, \dots, x_n)M(\sigma))) \\ &= \mathrm{Tr}(\rho(\mathrm{diag}(x_{\sigma(1)}, \dots, x_{\sigma(n)}))) \\ &= \mathrm{char}(\rho)(x_{\sigma(1)}, \dots, x_{\sigma(n)}). \end{aligned}$$

□

Example 1.2.3.
- The character of the identity representation is $x_1 + x_2 + \cdots + x_n$.
- The character of the representation $\rho\colon \mathbf{GL}_n(\mathbf{C}) \to \mathbf{GL}(\mathrm{Sym}^d(\mathbf{C}^n))$ is

$$h_n(x_1, \dots, x_n) = \sum_{1 \le i_1 \le \cdots \le i_d \le n} x_{i_1} \cdots x_{i_d}.$$

- The character of the representation $\rho\colon \mathbf{GL}_n(\mathbf{C}) \to \mathbf{GL}(\wedge^d(\mathbf{C}^n))$ is

$$e_n(x_1, \dots, x_n) = \sum_{1 \le i_1 < \cdots < i_d \le n} x_{i_1} \cdots x_{i_d}.$$

□

A few remarks that aren't easy to see right now (though we may revisit):
- The set of characters in $\Lambda(n)$ generates all of $\Lambda(n)$ as an abelian group.
- If we are more careful about how to define characters in infinite-dimensional settings, we get that characters of polynomial representations of $\mathbf{GL}_\infty(\mathbf{C})$ are elements of Λ.
- The character determines the representation up to isomorphism: if $\mathrm{char}(\rho) = \mathrm{char}(\rho')$, then ρ and ρ' define isomorphic representations: one can be obtained from the other by a suitable isomorphism of the underlying vector spaces V and V'.

If we take these remarks as fact for now, this gives one motivation for studying Λ. This representation-theoretic interpretation of Λ will clarify various definitions and constructions we will encounter. A few basic ones that we can see now:

- If $\rho_i \colon \mathbf{GL}_n(\mathbf{C}) \to \mathbf{GL}(V_i)$ are polynomial representations for $i = 1, 2$, we can form the direct sum representation $\rho_1 \oplus \rho_2 \colon \mathbf{GL}_n(\mathbf{C}) \to \mathbf{GL}(V_1 \oplus V_2)$ via
$$(\rho_1 \oplus \rho_2)(g) = \begin{pmatrix} \rho_1(g) & 0 \\ 0 & \rho_2(g) \end{pmatrix}$$
and
$$\mathrm{char}(\rho_1 \oplus \rho_2)(x_1, \dots, x_n) = \mathrm{char}(\rho_1)(x_1, \dots, x_n) + \mathrm{char}(\rho_2)(x_1, \dots, x_n).$$
- There's also a multiplicative version using tensor product. If $\rho_i \colon \mathbf{GL}_n(\mathbf{C}) \to \mathbf{GL}(V_i)$ are polynomial representations for $i = 1, 2$, we can form the tensor product representation $\rho_1 \otimes \rho_2 \colon \mathbf{GL}_n(\mathbf{C}) \to \mathbf{GL}(V_1 \otimes V_2)$ via (assuming $\rho_1(g)$ is $N \times N$):
$$(\rho_1 \otimes \rho_2)(g) = \begin{pmatrix} \rho_1(g)_{1,1}\rho_2(g) & \rho_1(g)_{1,2}\rho_2(g) & \cdots & \rho_1(g)_{1,N}\rho_2(g) \\ \rho_1(g)_{2,1}\rho_2(g) & \rho_1(g)_{2,2}\rho_2(g) & \cdots & \rho_1(g)_{2,N}\rho_2(g) \\ \vdots & & & \\ \rho_1(g)_{N,1}\rho_2(g) & \rho_1(g)_{N,2}\rho_2(g) & \cdots & \rho_1(g)_{N,N}\rho_2(g) \end{pmatrix}$$
(here we are multiplying $\rho_2(g)$ by each entry of $\rho_1(g)$ and creating a giant block matrix) and
$$\mathrm{char}(\rho_1 \otimes \rho_2)(x_1, \dots, x_n) = \mathrm{char}(\rho_1)(x_1, \dots, x_n) \cdot \mathrm{char}(\rho_2)(x_1, \dots, x_n).$$

Note that subtraction will not have any natural interpretation, and in general, the difference of two characters need not be a character. In general, the elements of $\Lambda(n)$ or Λ can be thought of as virtual characters since every element is the difference of two characters.

1.3. **Partitions.** A **partition** of a nonnegative integer n is a sequence $\lambda = (\lambda_1, \dots, \lambda_k)$ such that $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_k \ge 0$ and $\lambda_1 + \cdots + \lambda_k = n$. We will consider two partitions the same if their nonzero entries are the same. And for shorthand, we may omit the commas, so the partition $(1, 1, 1, 1)$ of 4 can be written as 1111. As a further shorthand, the exponential notation is used for repetition, so for example, 1^4 is the partition $(1, 1, 1, 1)$. We let $\mathrm{Par}(n)$ be the set of partitions of n, and denote the size by $p(n) = |\mathrm{Par}(n)|$. By convention, $\mathrm{Par}(0)$ consists of exactly one partition, the empty one.

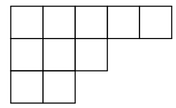
Example 1.3.1.
$$\begin{aligned} \mathrm{Par}(1) &= \{1\}, \\ \mathrm{Par}(2) &= \{2, 1^2\}, \\ \mathrm{Par}(3) &= \{3, 21, 1^3\}, \\ \mathrm{Par}(4) &= \{4, 31, 22, 21^2, 1^4\}, \\ \mathrm{Par}(5) &= \{5, 41, 32, 31^2, 2^21, 21^3, 1^5\}. \end{aligned}$$
□

If λ is a partition of n, we write $|\lambda| = n$ (**size**). Also, $\ell(\lambda)$ is the number of nonzero entries of λ (**length**). For each i, $m_i(\lambda)$ is the number of entries of λ that are equal to i.

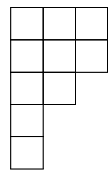
It will often be convenient to represent partitions graphically. This is done via **Young diagrams**, which is a collection of left-justified boxes with λ_i boxes in row i.[1] For example,

[1] In the **English convention**, row i sits above row $i + 1$, in the **French convention**, it is reversed. There is also the **Russian convention**, which is obtained from the French convention by rotating by 45 degrees counter-clockwise.

the Young diagram



corresponds to the partition $(5, 3, 2)$. Flipping across the main diagonal gives another partition $\lambda^\dagger$, called the **transpose**. In our example, flipping gives



So $(5, 3, 2)^\dagger = (3, 3, 2, 1, 1)$. In other words, the role of columns and rows has been interchanged. This is an important involution of $\mathrm{Par}(n)$ which we will use later.

We will use several different partial orderings of partitions:

- $\lambda \subseteq \mu$ if $\lambda_i \le \mu_i$ for all i.
- The **dominance order**: $\lambda \le \mu$ if $\lambda_1 + \cdots + \lambda_i \le \mu_1 + \cdots + \mu_i$ for all i. Note that if $|\lambda| = |\mu|$, then $\lambda \le \mu$ if and only if $\lambda^\dagger \ge \mu^\dagger$. So transpose is an order-reversing involution on the set of partitions of a fixed size.
- The **lexicographic order**: for partitions of the same size, $\lambda \le^R \mu$ if $\lambda = \mu$, or otherwise, there exists i such that $\lambda_1 = \mu_1, \dots, \lambda_{i-1} = \mu_{i-1}$, and $\lambda_i < \mu_i$. This is a total ordering.

2. Bases

Given an infinite sequence $(\alpha_1, \alpha_2, \dots)$ with finitely many nonzero entries, we use x^α as a convention for $\prod_{i \ge 1} x_i^{\alpha_i}$.

The following lemma will be useful, so we isolate it here.

Lemma 2.1. *Let $a_{\lambda,\mu}$ be a set of integers indexed by partitions of a fixed size n. Assume that $a_{\lambda,\lambda} = 1$ for all λ and that $a_{\lambda,\mu} \ne 0$ implies that $\mu \le \lambda$ (dominance order). For any ordering of the partitions, the matrix $(a_{\lambda,\mu})$ is invertible (i.e., has determinant ± 1).*

The same conclusion holds if instead we assume that $a_{\lambda,\lambda^\dagger} = 1$ for all λ and that $a_{\lambda,\mu} \ne 0$ implies that $\mu \le \lambda^\dagger$.

Proof. Note that if $\mu \le \lambda$, then $\mu \le^R \lambda$ (lexicographic order): suppose that $\lambda_1 = \mu_1, \dots, \lambda_i = \mu_i$ but $\lambda_{i+1} \ne \mu_{i+1}$. Using the dominance order, $\lambda_1 + \cdots + \lambda_{i+1} > \mu_1 + \cdots + \mu_{i+1}$, so we conclude that $\lambda_{i+1} > \mu_{i+1}$.

Now write down the matrix $(a_{\lambda,\mu})$ so that both the rows and columns are ordered by lexicographic ordering. Then this matrix has 1's on the diagonal and is lower-triangular. In particular, its determinant is 1, so it is invertible. Any other choice of ordering amounts to conjugating by a permutation matrix, which only changes the sign of the matrix.

In the second case, write down the matrix $(a_{\lambda,\mu})$ with respect to the lexicographic ordering $\lambda(1), \lambda(2), \dots, \lambda(p(n))$ for rows, but with respect to the ordering $\lambda(1)^\dagger, \lambda(2)^\dagger, \dots, \lambda(p(n))^\dagger$ for columns. This matrix again has 1's on the diagonal and is upper-triangular, so has determinant 1. If we want to write down the matrix with the same ordering for both rows and columns, we just need to permute the columns which changes the determinant by a sign only. $\square$

2.1. **Monomial symmetric functions.** Given a partition $\lambda = (\lambda_1, \lambda_2, \dots)$, define the **monomial symmetric function** by

$$m_\lambda = \sum_\alpha x^\alpha$$

where the sum is over all *distinct* permutations α of λ. This is symmetric by definition. So for example, $m_1 = \sum_{i \ge 1} x_i$ since all of the distinct permutations of $(1, 0, 0, \dots)$ are integer sequences with a single 1 somewhere and 0 elsewhere. By convention, $m_0 = 1$. Some other examples:

$$m_{1,1} = \sum_{i<j} x_i x_j$$
$$m_{3,2,1} = \sum_{\substack{i,j,k \\ i \neq j,\ j \neq k,\ i \neq k}} x_i x_j x_k.$$

In general, $m_{1^k} = e_k$ and $m_k = p_k$.

Theorem 2.1.1. *As we range over all partitions, the m_λ form a basis for Λ.*

Proof. They are linearly independent since no two m_λ have any monomials in common. Clearly they also span: given $f \in \Lambda$, we can write $f = \sum_\alpha c_\alpha x^\alpha$ and $c_\alpha = c_\beta$ if both are permutations of each other, so this can be rewritten as $f = \sum_\lambda c_\lambda m_\lambda$ where the sum is now over just the partitions. □

Corollary 2.1.2. *Λ_d has a basis given by $\{m_\lambda \mid |\lambda| = d\}$, and hence is a free abelian group of rank $p(d) = |\mathrm{Par}(d)|$.*

Theorem 2.1.3. *$\Lambda(n)_d$ has a basis given by $\{m_\lambda(x_1, \dots, x_n) \mid |\lambda| = d,\ \ell(\lambda) \le n\}$.*

2.2. **Elementary symmetric functions.** Recall that we defined

$$e_k = \sum_{i_1 < i_2 < \cdots < i_k} x_{i_1} x_{i_2} \cdots x_{i_k}.$$

For a partition $\lambda = (\lambda_1, \dots, \lambda_k)$, define the **elementary symmetric function** by

$$e_\lambda = e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_k}.$$

Note $e_\lambda \in \Lambda_{|\lambda|}$.

Since the m_μ form a basis for Λ (Theorem 2.1.1), we have expressions

$$e_\lambda = \sum_\mu M_{\lambda,\mu} m_\mu.$$

We can give an interpretation for these change of basis coefficients as follows. Given an (infinite) matrix A with finitely many nonzero entries, let $\mathrm{row}(A) = (\sum_{i \ge 1} A_{1,i}, \sum_{i \ge 1} A_{2,i}, \dots)$ be the sequence of row sums of A, and let $\mathrm{col}(A)$ be the sequence of column sums of A. A $(0,1)$-matrix is one whose entries are only 0 or 1.

Lemma 2.2.1. *$M_{\lambda,\mu}$ is the number of $(0,1)$-matrices A with $\mathrm{row}(A) = \lambda$ and $\mathrm{col}(A) = \mu$.*

Proof. To get a monomial in e_λ, we have to choose monomials from each e_{λ_i} to multiply. Each monomial of e_{λ_i} can be represented by a subset of $\{1, 2, \dots\}$ of size λ_i, or alternatively, as a sequence (thought of as a row vector) of 0's and 1's where a 1 is put in each place of the subset. Hence, we can represent each choice of multiplying out a monomial by concatenating these row vectors to get a matrix A. By definition, $\mathrm{row}(A) = \lambda$ and $\mathrm{col}(A) = \mu$ where the monomial we get is x^μ. $\square$

Corollary 2.2.2. $M_{\lambda,\mu} = M_{\mu,\lambda}$.

Proof. Take the transpose of each $(0,1)$-matrix to get the desired bijection. $\square$

Theorem 2.2.3. *If $M_{\lambda,\mu} \neq 0$, then $\mu \leq \lambda^\dagger$. Furthermore, $M_{\lambda,\lambda^\dagger} = 1$. In particular, the e_λ form a basis of Λ.*

Proof. Suppose $M_{\lambda,\mu} \neq 0$. Then there is $(0,1)$-matrix with $\mathrm{row}(A) = \lambda$ and $\mathrm{col}(A) = \mu$. Now let A' be obtained from A by left-justifying all of the 1's in each row (i.e., move all of the 1's in row i to the first λ_i positions). Note that $\mathrm{col}(A') = \lambda^\dagger$. Also, the number of 1's in the first i columns of A' is as least as many as the number of 1's in the first i columns of A, so $\lambda^\dagger_1 + \cdots + \lambda^\dagger_i \geq \mu_1 + \cdots + \mu_i$, i.e., $\lambda^\dagger \geq \mu$. Moreover, if $\mu = \lambda^\dagger$, A' is the only $(0,1)$-matrix with $\mathrm{row}(A') = \lambda$ and $\mathrm{col}(A') = \lambda^\dagger$.

The second statement follows from Lemma 2.1. $\square$

Theorem 2.2.4. *The set $\{e_\lambda(x_1, \dots, x_n) \mid \lambda_1 \leq n,\ |\lambda| = d\}$ is a basis of $\Lambda(n)_d$.*

Proof. If $\lambda_1 > n$, then $e_{\lambda_1}(x_1, \dots, x_n) = 0$, so $e_\lambda(x_1, \dots, x_n) = 0$. Hence under the map $\pi_n \colon \Lambda \to \Lambda(n)$, the proposed e_λ span the image. The number of such e_λ in degree d is $|\{\lambda \mid \lambda_1 \leq n,\ |\lambda| = d\}|$, which is the same as $|\{\lambda \mid \ell(\lambda) \leq n,\ |\lambda| = d\}|$ via the transpose $\dagger$, and this is the rank of $\Lambda(n)_d$, so the e_λ form a basis. $\square$

Remark 2.2.5. The previous two theorems say that the elements $e_1, e_2, e_3, \dots$ are algebraically independent in Λ, and that the elements $e_1, \dots, e_n$ are algebraically independent in $\Lambda(n)$. This is also known as the "fundamental theorem of symmetric functions". $\square$

2.3. **The involution ω.** Since the e_i are algebraically independent, we can define a ring homomorphism $f \colon \Lambda \to \Lambda$ by specifying $f(e_i)$ arbitrarily.[2] Define

$$\omega \colon \Lambda \to \Lambda$$

by $\omega(e_i) = h_i$, where recall that $h_k = \sum_{i_1 \leq \cdots \leq i_k} x_{i_1} \cdots x_{i_k}$.

Theorem 2.3.1. *ω is an involution, i.e., $\omega^2 = 1$. Equivalently, $\omega(h_i) = e_i$.*

Proof. Consider the ring $\Lambda[\![t]\!]$ of power series in t with coefficients in Λ. Define two elements of $\Lambda[\![t]\!]$:

$$E(t) = \sum_{n \geq 0} e_n t^n, \qquad H(t) = \sum_{n \geq 0} h_n t^n. \tag{2.3.2}$$

Note that $E(t) = \prod_{n \geq 1}(1 + x_i t)$ (by convention, the infinite product means we have to choose 1 all but finitely many times; if you multiply it out, the coefficient of t^n is all ways of getting a monomial $x_{i_1} \cdots x_{i_n}$ with $i_1 < \cdots < i_n$ and each one appears once, so it is e_n) and that $H(t) = \prod_{n \geq 1}(1 - x_i t)^{-1}$ (same as for $E(t)$ but use the geometric sum $(1 - x_i t)^{-1} = 1 + \sum_{d > 0} x_i^d t^d$).

[2] Every element is uniquely of the form $\sum_\lambda c_\lambda e_\lambda$; since f is a ring homomorphism, it sends this to $\sum_\lambda c_\lambda f(e_{\lambda_1}) f(e_{\lambda_2}) \cdots f(e_{\lambda_{\ell(\lambda)}})$.

This implies that $E(t)H(-t) = 1$. The coefficient of t^n on the left side of this identity is $\sum_{i=0}^n (-1)^{n-i} e_i h_{n-i}$. In particular, that sum is 0 for $n > 0$. Now apply ω to that sum and multiply by $(-1)^n$ to get

$$\sum_{i=0}^n (-1)^i h_i \omega(h_{n-i}) = 0.$$

This shows that $\sum_{n \ge 0} \omega(h_n) t^n = H(-t)^{-1} = E(t)$, so $\omega(h_n) = e_n$. $\square$

Furthermore, we can define a finite analogue of ω, the ring homomorphism $\omega_n \colon \Lambda(n) \to \Lambda(n)$, given by $\omega_n(e_i) = h_i$ for $i = 1, \dots, n$.

Theorem 2.3.3. *$\omega_n^2 = 1$, and ω_n is invertible. Equivalently, $\omega_n(h_i) = e_i$ for $i = 1, \dots, n$.*

Proof. In Λ, we have expressions $h_i = \sum_{|\lambda|=i} c_{i,\lambda} e_\lambda$. We know that $\omega(h_i) = e_i$, so we also get $e_i = \sum_{|\lambda|=i} c_{i,\lambda} \omega(e_\lambda)$.

Using the first relation, we also get $\omega_n(\pi_n(h_i)) = \sum_{|\lambda|=i} c_{i,\lambda} \omega_n(\pi_n(e_\lambda))$. By definition, $\omega_n(\pi_n(e_\lambda)) = \pi_n(\omega(e_\lambda))$ when $\lambda_1 \le n$. This condition is guaranteed if $i \le n$, so we can rewrite it as $\omega_n(\pi_n(h_i)) = \sum_{|\lambda|=i} c_{i,\lambda} \pi_n(\omega(e_\lambda))$. Finally, applying π_n to the second relation above, we get $\pi_n(e_i) = \sum_{|\lambda|=i} c_{i,\lambda} \pi_n(\omega(e_\lambda))$, so we conclude that $\pi_n(e_i) = \omega_n(\pi_n(h_i))$, as desired. $\square$

2.4. **Complete homogeneous symmetric functions.** For a partition $\lambda = (\lambda_1, \dots, \lambda_k)$, define the **complete homogeneous symmetric functions** by

$$h_\lambda = h_{\lambda_1} \cdots h_{\lambda_k}.$$

Theorem 2.4.1. *The h_λ form a basis for Λ.*

Proof. Since ω is a ring homomorphism, $\omega(e_\lambda) = h_\lambda$. Now use the fact that the e_λ form a basis (Theorem 2.2.3) and that ω is an isomorphism (Theorem 2.3.1). $\square$

Again, we can write h_λ in terms of m_μ:

$$h_\lambda = \sum_\mu N_{\lambda,\mu} m_\mu$$

and give an interpretation for the coefficients. This is similar to $M_{\lambda,\mu}$: the $N_{\lambda,\mu}$ is the number of matrices A with non-negative integer entries such that $\mathrm{row}(A) = \lambda$ and $\mathrm{col}(A) = \mu$ (not just $(0,1)$-matrices). The proof is similar to the $M_{\lambda,\mu}$ case. However, it does not satisfy any upper-triangularity properties, so it is not as easy to see directly (without using ω) that the h_λ are linearly independent.

Theorem 2.4.2. *$h_1, \dots, h_n$ are algebraically independent generators of $\Lambda(n)$, and the set $\{h_\lambda(x_1, \dots, x_n) \mid \lambda_1 \le n,\ |\lambda| = d\}$ is a basis of $\Lambda(n)_d$.*

Proof. Follows from Theorem 2.2.4 and Theorem 2.3.3. $\square$

2.5. **Power sum symmetric functions.** Recall we defined

$$p_k = \sum_{n \ge 1} x_n^k.$$

For a partition $\lambda = (\lambda_1, \dots, \lambda_k)$, the **power sum symmetric functions** are defined by

$$p_\lambda = p_{\lambda_1} \cdots p_{\lambda_k}.$$

We can write the p_λ in terms of the m_μ:

$$p_\lambda = \sum_\mu R_{\lambda,\mu} m_\mu.$$

Theorem 2.5.1. *If $R_{\lambda,\mu} \neq 0$, then $\lambda \leq \mu$. Also, $R_{\lambda,\lambda} \neq 0$. In particular, the p_λ are linearly independent.*

Proof. Every monomial in p_λ is of the form $x_{i_1}^{\lambda_1} \cdots x_{i_k}^{\lambda_k}$ for some choice of $i_1, \dots, i_k$. Suppose this is equal to x^μ for a partition μ. So we get μ be reordering the i_j and merging together equal indices. In other words, there is a decomposition $B_1 \amalg \cdots \amalg B_r = \{1, \dots, k\}$ so that $\mu_j = \sum_{i \in B_j} \lambda_i$. We claim this implies $\mu_1 + \cdots + \mu_i \geq \lambda_1 + \cdots + \lambda_i$ for all i. For each λ_j with $j \leq i$, if $j \notin B_1 \cup \cdots \cup B_i$, then $j \in B_{i'}$ for $i' > i$, and then $\mu_j \geq \mu_{i'} \geq \lambda_j$. Subtract μ_j from the sum $\mu_1 + \cdots + \mu_i$ and subtract λ_j from $\lambda_1 + \cdots + \lambda_i$. Then it suffices to show that $\sum_{s \in S} \mu_s \geq \sum_{s \in S} \lambda_s$ where S is the set of indices where $\lambda_s \in B_1 \cup \cdots \cup B_i$. But this is immediate from the definition of the B_j. $\square$

Remark 2.5.2. The p_λ do not form a basis for Λ. For example, in degree 2, we have

$$\begin{aligned} p_2 &= m_2 \\ p_{1,1} &= m_2 + 2m_{1,1} \end{aligned}$$

and the change of basis matrix has determinant 2, so is not invertible over $\mathbf{Z}$. However, they do form a basis for $\Lambda_{\mathbf{Q}}$. $\square$

Recall the definitions of $E(t)$ and $H(t)$ from (2.3.2):

$$E(t) = \sum_{n \geq 0} e_n t^n, \qquad H(t) = \sum_{n \geq 0} h_n t^n.$$

Define $P(t) \in \Lambda[\![t]\!]$ by

$$P(t) = \sum_{n \geq 1} p_n t^{n-1}$$

(note the unconventional indexing).

Lemma 2.5.3. *We have the following identities:*

$$P(t) = \frac{d}{dt} \log H(t), \qquad P(-t) = \frac{d}{dt} \log E(t).$$

Proof. We have

$$\begin{aligned} P(t) &= \sum_{n \geq 1} \sum_{i \geq 1} x_i^n t^{n-1} \\ &= \sum_{i \geq 1} \frac{x_i}{1 - x_i t} \\ &= \sum_{i \geq 1} \frac{d}{dt} \log \left(\frac{1}{1 - x_i t} \right) \\ &= \frac{d}{dt} \log \left(\prod_{i \geq 1} \frac{1}{1 - x_i t} \right) \\ &= \frac{d}{dt} \log H(t). \end{aligned}$$

The other identity is similar. ☐

Given a partition λ, recall that $m_i(\lambda)$ is the number of times that i appears in λ. Define

$$z_\lambda := \prod_{i \ge 1} i^{m_i(\lambda)} m_i(\lambda)!, \qquad \varepsilon_\lambda = (-1)^{|\lambda| - \ell(\lambda)}. \tag{2.5.4}$$

Theorem 2.5.5. *We have the following identities in* $\Lambda[\![t]\!]$*:*

$$E(t) = \sum_\lambda \varepsilon_\lambda z_\lambda^{-1} p_\lambda t^{|\lambda|}, \qquad H(t) = \sum_\lambda z_\lambda^{-1} p_\lambda t^{|\lambda|},$$
$$e_n = \sum_{|\lambda|=n} \varepsilon_\lambda z_\lambda^{-1} p_\lambda, \qquad h_n = \sum_{|\lambda|=n} z_\lambda^{-1} p_\lambda.$$

Proof. From Lemma 2.5.3, we have $P(t) = \frac{d}{dt} \log H(t)$. Integrate both sides (and get the boundary conditions right using that $\log H(0) = 0$) and apply the exponential map:

$$\begin{aligned} H(t) &= \exp\left(\sum_{n \ge 1} \frac{p_n t^n}{n}\right) \\ &= \prod_{n \ge 1} \exp\left(\frac{p_n t^n}{n}\right) \\ &= \prod_{n \ge 1} \sum_{d \ge 0} \frac{p_n^d t^{nd}}{n^d d!} \\ &= \sum_\lambda \frac{p_\lambda t^{|\lambda|}}{z_\lambda}. \end{aligned}$$

The identity for $E(t)$ is similar. Finally, the second row of identities comes from equating the coefficient of t^n in the first row of identities. ☐

Corollary 2.5.6. $\omega(p_\lambda) = \varepsilon_\lambda p_\lambda$*, i.e., the* p_λ *are a complete set of eigenvectors for* ω*.*

Proof. We prove this by induction on λ_1. When $\lambda_1 = 1$, this is clear since $p_{1^n} = p_1^n = e_1^n = h_1^n$ and $\varepsilon_{1^n} = 1$. So suppose we know that $\omega(p_\lambda) = \varepsilon_\lambda p_\lambda$ whenever $\lambda_1 < n$. Apply ω to the identity

$$e_n = \sum_{|\lambda|=n} \varepsilon_\lambda z_\lambda^{-1} p_\lambda,$$

to get

$$h_n = \sum_{|\lambda|=n} \varepsilon_\lambda z_\lambda^{-1} \omega(p_\lambda).$$

Every partition satisfies $\lambda_1 < n$ except for $\lambda = (n)$, so this can be simplified to

$$h_n = \varepsilon_n z_n^{-1} \omega(p_n) + \sum_{|\lambda|=n,\ \lambda_1<n} z_\lambda^{-1} p_\lambda.$$

Compare this to the identity

$$h_n = \sum_{|\lambda|=n} z_\lambda^{-1} p_\lambda$$

to conclude that $\varepsilon_n z_n^{-1}\omega(p_n) = z_n^{-1}p_n$; now multiply both sides by z_n. Now given any other partition with $\lambda_1 = n$, use the fact that ω is a ring homomorphism, that $\varepsilon_\lambda = \varepsilon_{\lambda_1}\cdots\varepsilon_{\lambda_k}$, and that $p_\lambda = p_{\lambda_1}\cdots p_{\lambda_k}$ to conclude that $\omega(p_\lambda) = \varepsilon_\lambda p_\lambda$. ☐

Theorem 2.5.7. *$p_1, \dots, p_n$ are algebraically independent generators of $\Lambda(n)_{\mathbf{Q}}$ and the set $\{p_\lambda(x_1,\dots,x_n) \mid \lambda_1 \le n,\ |\lambda| = d\}$ is a basis for $\Lambda(n)_{\mathbf{Q},d}$.*

Proof. By Theorem 2.5.5, we have $e_i(x_1,\dots,x_n) = \sum_{|\lambda|=i} \varepsilon_\lambda z_\lambda^{-1} p_\lambda(x_1,\dots,x_n)$ for all i. If $i \le n$, then this shows that $p_1,\dots,p_n$ are algebra generators for $\Lambda(n)_{\mathbf{Q}}$ since the $e_1,\dots,e_n$ are algebra generators. The space of possible monomials in the p_i of degree d is the number of λ with $\lambda_1 \le n$ and $|\lambda| = d$, which is $\dim_{\mathbf{Q}} \Lambda(n)_{\mathbf{Q},d}$, so there are no algebraic relations among them. ☐

The proof of Corollary 2.5.6 can be adapted to show that $\omega_n(p_\lambda) = \varepsilon_\lambda p_\lambda$ whenever $\lambda_1 \le n$.

2.6. A scalar product.

Define a bilinear form $\langle , \rangle \colon \Lambda \otimes \Lambda \to \mathbf{Z}$ by setting

$$\langle m_\lambda, h_\mu \rangle = \delta_{\lambda,\mu}$$

where δ is the Kronecker delta (1 if $\lambda = \mu$ and 0 otherwise). In other words, if $f = \sum_\lambda a_\lambda m_\lambda$ and $g = \sum_\mu b_\mu h_\mu$, then $\langle f, g\rangle = \sum_\lambda a_\lambda b_\lambda$ (well-defined since both m and h are bases). At this point, the definition looks completely unmotivated. However, this inner product is natural from the representation-theoretic perspective, which we'll mention in the next section (without proof).

In our setup, m and h are dual bases with respect to the pairing. We will want a general criteria for two bases to be dual to each other. To state this criterion, we need to work in two sets of variables x and y and in the ring $\Lambda \otimes \Lambda$ where the x's and y's are separately symmetric.

Lemma 2.6.1. *Let u_λ and v_μ be bases of Λ (or $\Lambda_{\mathbf{Q}}$). Then $\langle u_\lambda, v_\mu\rangle = \delta_{\lambda,\mu}$ if and only if*

$$\sum_\lambda u_\lambda(x)v_\lambda(y) = \prod_{i,j}(1 - x_i y_j)^{-1}.$$

Proof. Write $u_\lambda = \sum_\alpha a_{\lambda,\rho} m_\alpha$ and $v_\mu = \sum_\beta b_{\mu,\beta} h_\beta$. Pick an ordering of the partitions of a fixed size. Write $A = (a_{\lambda,\rho})$ and $B = (b_{\mu,\beta})$ in these orderings. First,

$$\langle u_\lambda, v_\mu\rangle = \sum_\gamma a_{\lambda,\gamma} b_{\mu,\gamma}.$$

Hence u_λ and v_μ are dual bases if and only if $\sum_\gamma a_{\lambda,\gamma} b_{\mu,\gamma} = \delta_{\lambda,\mu}$, or equivalently, $AB^T = I$ where T denotes transpose and I is the identity matrix. So A and B^T are inverses of each other, and so this is equivalent to $B^T A = I$, or $\sum_\gamma a_{\gamma,\lambda} b_{\gamma,\mu} = \delta_{\lambda,\mu}$. Finally, we have

$$\sum_\lambda u_\lambda(x)v_\lambda(y) = \sum_\lambda \sum_{\alpha,\beta} a_{\lambda,\alpha} b_{\lambda,\beta} m_\alpha(x) h_\beta(y) = \sum_{\alpha,\beta}\left(\sum_\lambda a_{\lambda,\alpha} b_{\lambda,\beta}\right) m_\alpha(x) h_\beta(y).$$

Since the $m_\alpha(x)h_\beta(y)$ are linearly independent, we see that $\sum_\gamma a_{\gamma,\lambda} b_{\gamma,\mu} = \delta_{\lambda,\mu}$ is equivalent to $\sum_\lambda u_\lambda(x)v_\lambda(y) = \sum_\lambda m_\lambda(x)h_\lambda(y)$. Now use Proposition 2.6.2 below. ☐

Proposition 2.6.2.

$$\sum_\lambda m_\lambda(x)h_\lambda(y) = \prod_{i,j}(1 - x_i y_j)^{-1}.$$

Proof. We have

$$\prod_i \prod_j (1 - x_i y_j)^{-1} = \prod_j \sum_{n \ge 0} h_n(y) x_j^n = \sum_\alpha h_\alpha(y) x^\alpha$$

where the sum is over all sequences $\alpha = (\alpha_1, \alpha_2, \dots)$ with finitely many nonzero entries, and $h_\alpha(y) = h_{\alpha_1}(y) h_{\alpha_2}(y) \cdots$; finally, grouping together terms α in the same Σ_∞-orbit, the latter sum simplifies to $\sum_\lambda m_\lambda(x) h_\lambda(y)$, where the sum is now over all partitions λ. □

Corollary 2.6.3. *The pairing is symmetric, i.e.,* $\langle f, g \rangle = \langle g, f \rangle$.

Proof. The condition above is the same if we interchange x and y, so $\langle m_\lambda, h_\mu \rangle = \langle h_\mu, m_\lambda \rangle$. Now use bilinearity. □

Proposition 2.6.4. *We have*

$$\sum_\lambda z_\lambda^{-1} p_\lambda(x) p_\lambda(y) = \prod_{i,j} (1 - x_i y_j)^{-1}.$$

In particular, $\langle p_\lambda, p_\mu \rangle = z_\lambda \delta_{\lambda,\mu}$, *and* p_λ *is an orthogonal basis of* $\Lambda_{\mathbf{Q}}$.

Proof. As in the first part of the proof of Theorem 2.5.5, we can write

$$\begin{aligned}
\prod_j \prod_i (1 - x_i y_j)^{-1} &= \prod_j \sum_{n \ge 0} h_n(x) y_j^n \\
&= \prod_j \exp\left(\sum_n \frac{p_n(x) y_j^n}{n} \right) \\
&= \exp\left(\frac{p_n(x) p_n(y)}{n} \right) \\
&= \sum_{d \ge 0} \frac{p_n(x)^d p_n(y)^d}{d! n^d} \\
&= \sum_\lambda z_\lambda^{-1} p_\lambda(x) p_\lambda(y)
\end{aligned}$$

where the final sum is over all partitions. □

Corollary 2.6.5. ω *is an isometry, i.e.,* $\langle f, g \rangle = \langle \omega(f), \omega(g) \rangle$.

Proof. By bilinearity, it suffices to show that this holds for any basis of $\Lambda_{\mathbf{Q}}$. We pick p_λ. Then

$$\langle \omega(p_\lambda), \omega(p_\mu) \rangle = \varepsilon_\lambda \varepsilon_\mu \langle p_\lambda, p_\mu \rangle = \varepsilon_\lambda \varepsilon_\mu \delta_{\lambda,\mu}.$$

by Corollary 2.5.6 and the previous result. The last expression is the same as $\delta_{\lambda,\mu}$ since $\varepsilon_\lambda^2 = 1$, so we see that $\langle \omega(p_\lambda), \omega(p_\mu) \rangle = \langle p_\lambda, p_\mu \rangle$ for all λ, μ. □

Corollary 2.6.6. *The bilinear form* $\langle , \rangle$ *is positive definite, i.e,* $\langle f, f \rangle > 0$ *for* $f \neq 0$.

Proof. Assume $f \neq 0$. Write $f = \sum_\lambda a_\lambda p_\lambda$ with some $a_\lambda \neq 0$. Then

$$\langle f, f \rangle = \sum_{\lambda,\mu} a_\lambda a_\mu \langle p_\lambda, p_\mu \rangle = \sum_\lambda z_\lambda a_\lambda^2.$$

Since $z_\lambda > 0$ and $a_\lambda^2 > 0$, we get the result. □

2.7. **Some representation theory.** From what we said before,

$$e_n = \mathrm{char}(\wedge^n(\mathbf{C}^\infty)), \qquad h_n = \mathrm{char}(\mathrm{Sym}^n(\mathbf{C}^\infty)).$$

Also, tensor products go to multiplication of characters, so

$$e_\lambda = \mathrm{char}(\wedge^{\lambda_1}(\mathbf{C}^\infty) \otimes \cdots \otimes \wedge^{\lambda_k}(\mathbf{C}^\infty)), \qquad h_\lambda = \mathrm{char}(\mathrm{Sym}^{\lambda_1}(\mathbf{C}^\infty) \otimes \cdots \otimes \mathrm{Sym}^{\lambda_k}(\mathbf{C}^\infty))$$

The symmetric functions m_λ and p_λ are not in general the characters of polynomial representations. For example, one can show that any polynomial complex representation of degree 2 (its character is homogeneous of degree 2) is a positive linear combination of h_2 and e_2, while $p_2 = h_2 - e_2$.

Remark 2.7.1. If we allow ourselves to work in positive characteristic, the situation changes dramatically. For example, in characteristic 2, p_2 is the character[3] of the Frobenius twist of the identity representation: this is the map $\mathbf{GL}_n(\mathbf{F}) \to \mathbf{GL}_n(\mathbf{F})$ (n finite or ∞) which squares every entry ($\mathbf{F}$ is some algebraically closed field of characteristic 2). The representation theory in positive characteristic is poorly understood, so we won't say much about it. $\square$

A representation-theoretic interpretation of p_λ will be made more clear when we discuss the connection between Λ and complex representations of the symmetric group.

In general the e_λ and h_λ are not characters of irreducible representations (those which do not have a nonzero proper subrepresentation), and part of the goal of the next section is to find the characters of the irreducibles.

The involution ω also has an algebraic interpretation. On the level of multilinear operations, it is swapping the exterior power and symmetric power functors. Recall that the kth exterior power of a vector space V (over a field of characteristic 0) can be constructed as

$$V^{\otimes k}/(\sigma(v_1 \otimes \cdots \otimes v_k) - \mathrm{sgn}(\sigma)(v_1 \otimes \cdots \otimes v_k))$$

where the relations range over all permutations $\sigma \in \Sigma_k$ and all $v_i \in V$. The symmetric power has a similar construction, except $\mathrm{sgn}(\sigma)$ does not appear. The two are related by the action of the symmetric group, and more precisely, the choice of a symmetry for the tensor product. The usual one is $V \otimes W \cong W \otimes V$ via $v \otimes w \mapsto w \otimes v$. However, there is another one defined by $v \otimes w \mapsto -w \otimes v$ (in categorical language, there are two different symmetric monoidal structures on the category of vector spaces that extend the monoidal structure given by tensor product). Abstractly, given a tensor product and a choice of symmetry, exterior powers and symmetric powers can be defined; if we change the symmetry, then the constructions get swapped. So ω is an incarnation of switching the symmetry of the tensor product.

Finally, the bilinear pairing $\langle , \rangle$ has an important representation-theoretic interpretation, which we will not prove. Given two polynomial representations V, W of $\mathbf{GL}_n(\mathbf{C})$, let $\mathrm{Hom}_{\mathbf{GL}_n(\mathbf{C})}(V, W)$ be the set of linear maps $f\colon V \to W$ such that $f(\rho_V(g)v) = \rho_W(g)f(v)$ for all $v \in V$ and $g \in \mathbf{GL}_n(\mathbf{C})$. When $n = \infty$, we have

$$\langle \mathrm{char}(V), \mathrm{char}(W) \rangle = \dim_{\mathbf{C}} \mathrm{Hom}_{\mathbf{GL}_\infty(\mathbf{C})}(V, W).$$

There is also a version for n finite if we define the analogue of $\langle , \rangle$ for $\Lambda(n)$ (in the same way).

[3]If we define the character as the trace, we get something valued in $\Lambda \otimes \mathbf{F}$, but it is possible to modify the definition so that it is valued in Λ.

The quantity $\prod_{i,j}(1 - x_i y_j)$ is the character of the $\mathbf{GL}_\infty(\mathbf{C}) \times \mathbf{GL}_\infty(\mathbf{C})$ action on the symmetric algebra $\bigoplus_{d \geq 0} \mathrm{Sym}^d(\mathbf{C}^\infty \otimes \mathbf{C}^\infty)$ where the first, respectively second, $\mathbf{GL}_\infty(\mathbf{C})$ acts on the first, respectively second, copy of $\mathbf{C}^\infty$.

3. Schur functions and the RSK algorithm

The goal of this section is to give several different definitions of Schur functions. This will appear unmotivated. However, at the end, we'll comment on their central role in the representation theory of general linear groups. The key point is that they are the characters of the irreducible representations.

3.1. **Semistandard Young tableaux.** Let λ be a partition. A **semistandard Young tableaux (SSYT)** T is an assignment of natural numbers to the Young diagram of λ so that the numbers are weakly increasing going left to right in each row, and the numbers are strictly increasing going top to bottom in each column.

Example 3.1.1. If $\lambda = (4, 3, 1)$, and we have the assignment

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

then, in order for this to be a SSYT, we need to have

- $a \leq b \leq c \leq d$,
- $e \leq f \leq g$,
- $a < e < h$,
- $b < f$, and
- $c < g$.

An example of a SSYT is $\begin{array}{|c|c|c|c|}\hline 1 & 1 & 3 & 5 \\ \hline 2 & 3 & 4 \\ \cline{1-3} 7 \\ \cline{1-1}\end{array}$. $\square$

The type of a SSYT T is the sequence $\mathrm{type}(T) = (\alpha_1, \alpha_2, \dots)$ where α_i is the number of times that i appears in T. We set

$$x^T = x_1^{\alpha_1} x_2^{\alpha_2} \cdots .$$

Given a pair of partitions $\mu \subseteq \lambda$, the Young diagram of λ/μ is the Young diagram of λ with the Young diagram of μ removed. We define a SSYT of shape λ/μ to be an assignment of natural numbers of this Young diagram which is weakly increasing in rows and strictly increasing in columns.

Example 3.1.2. If $\lambda = (5, 3, 1)$ and $\mu = (2, 1)$, then

$$\begin{array}{ccccc}\cline{3-5} & & \multicolumn{1}{|c|}{a} & \multicolumn{1}{c|}{b} & \multicolumn{1}{c|}{c} \\ \cline{2-5} & \multicolumn{1}{|c|}{d} & \multicolumn{1}{c|}{e} \\ \cline{1-3} \multicolumn{1}{|c|}{f} \\ \cline{1-1}\end{array}$$

is a SSYT if

- $a \leq b \leq c$,
- $d \leq e$, and
- $a < e$. $\square$

We define the type of T and x^T in the same way.

Given a partition λ, the **Schur function** s_λ is defined by

$$s_\lambda = \sum_T x^T$$

where the sum is over all SSYT of shape λ. Similarly, given $\mu \subseteq \lambda$, the **skew Schur function** $s_{\lambda/\mu}$ is defined by

$$s_{\lambda/\mu} = \sum_T x^T$$

where the sum is over all SSYT of shape λ/μ. Note that this is a strict generalization of the first definition since we can take $\mu = \varnothing$, the unique partition of 0.

We can make the same definitions in finitely many variables $x_1, \dots, x_n$ if we restrict the sums to be only over SSYT that only use the numbers $1, \dots, n$.

Example 3.1.3. $s_{1,1}(x_1, x_2, \dots, x_n)$ is the sum over SSYT of shape $(1,1)$. This is the same as a choice of $1 \le i < j \le n$, so $s_{1,1}(x_1, \dots, x_n) = \sum_{1 \le i < j \le n} x_i x_j = e_2(x_1, \dots, x_n)$, and by the same reasoning, $s_{1,1} = e_2$ in infinitely many variables. More generally, $s_{1^k} = e_k$ for any k.

Also, $s_k = h_k$ since a SSYT of shape (k) is a choice of $i_1 \le i_2 \le \cdots \le i_k$.

For something different, consider $s_{2,1}(x_1, x_2, x_3)$. There are 8 SSYT that of shape $(2,1)$ that only use $1, 2, 3$:

$$\begin{ytableau} 1 & 1 \\ 2 \end{ytableau} \quad \begin{ytableau} 1 & 1 \\ 3 \end{ytableau} \quad \begin{ytableau} 1 & 2 \\ 2 \end{ytableau} \quad \begin{ytableau} 1 & 2 \\ 3 \end{ytableau} \quad \begin{ytableau} 1 & 3 \\ 2 \end{ytableau} \quad \begin{ytableau} 1 & 3 \\ 3 \end{ytableau} \quad \begin{ytableau} 2 & 2 \\ 3 \end{ytableau} \quad \begin{ytableau} 2 & 3 \\ 3 \end{ytableau}$$

From this, we can read off that $s_{2,1}(x_1, x_2, x_3)$ is a symmetric polynomial. Furthermore, it is $m_{2,1}(x_1, x_2, x_3) + 2m_{1,1,1}(x_1, x_2, x_3)$. $\square$

Theorem 3.1.4. *For any $\mu \subseteq \lambda$, the skew Schur function $s_{\lambda/\mu}$ is a symmetric function.*

Proof. We need to check that for every sequence α and every permutation σ, the number of SSYT of shape λ/μ and type α is the same as the number of SSYT of shape λ/μ and type $\sigma(\alpha)$. Since α has only finitely many nonzero entries, we can always replace σ by a permutation σ' that permutes finitely many elements of $\{1, 2, \dots\}$ so that $\sigma(\alpha) = \sigma'(\alpha)$. But then σ' can be written as a finite product of adjacent transpositions $(i, i+1)$. So it's enough to check the case when $\sigma = (i, i+1)$.

Let T be a SSYT of shape α. Do the following: take the set of columns that only contain exactly one of i or $i+1$. Now consider just the entries of these columns that contain i or $i+1$. The result is a series of isolated rows. In a given row, if there are a instances of i and b instances of $i+1$, then replace it by b instances of i and a instances of $i+1$. The result is still a SSYT, but the type is now $(i, i+1)(\alpha)$. This is reversible, so defines the desired bijection. $\square$

We now focus on Schur functions. Suppose λ is a partition of n. Let $K_{\lambda,\alpha}$ be the number of SSYT of shape λ and type α, this is called a **Kostka number**. The previous theorem says $K_{\lambda,\alpha} = K_{\lambda,\sigma(\alpha)}$ for any permutation σ, so it's enough to study the case when α is a partition. By the definition of Schur function, we have

$$s_\lambda = \sum_{\mu \vdash n} K_{\lambda,\mu} m_\mu.$$

An important special case is when $\mu = 1^n$. Then $K_{\lambda,1^n}$ is the number of SSYT that use each of the numbers $1, \dots, n$ exactly once. Such a SSYT is called a **standard Young tableau**, and $K_{\lambda,1^n}$ is denoted f^λ.

Theorem 3.1.5. *If $K_{\lambda,\mu} \neq 0$, then $\mu \le \lambda$ (dominance order). Also, $K_{\lambda,\lambda} = 1$. In particular, the s_λ form a basis for Λ.*

Proof. Suppose that $K_{\lambda,\mu} \neq 0$. Pick a SSYT T of shape λ and type μ. Each number k can only appear in the first k rows of T: otherwise, there is a column with entries $1 \le i_1 < i_2 < \cdots < i_r < k$ where $r \ge k$, which is a contradiction. This implies that $\mu_1 + \cdots + \mu_k \le \lambda_1 + \cdots + \lambda_k$, so $\mu \le \lambda$.

The only SSYT of shape λ and type λ is the one that fills row i with the number i.

Now the last statement follows from Lemma 2.1. □

Corollary 3.1.6. *$\{s_\lambda \mid |\lambda| = d\}$ is a basis for Λ_d.*

Corollary 3.1.7. *$\{s_\lambda(x_1, \dots, x_n) \mid |\lambda| = d,\ \ell(\lambda) \le n\}$ is a basis for $\Lambda(n)_d$.*

Proof. Note that if $\ell(\lambda) > n$, there are no SSYT only using $1, \dots, n$, so $s_\lambda(x_1, \dots, x_n) = 0$. Hence the set in question spans $\Lambda(n)_d$. Since $\Lambda(n)_d$ is free of rank equal to the size of this set, it must also be a basis. □

Remark 3.1.8. For each partition λ, there is a multilinear construction, known as a **Schur functor** $\mathbf{S}_\lambda$ which takes as input a vector space V and outputs another vector space $\mathbf{S}_\lambda(V)$, and given any linear map $f\colon V \to W$, a linear map $\mathbf{S}_\lambda(f)\colon \mathbf{S}_\lambda(V) \to \mathbf{S}_\lambda(W)$. Furthermore, it is *functorial*: $\mathbf{S}_\lambda(g \circ f) = \mathbf{S}_\lambda(g) \circ \mathbf{S}_\lambda(f)$. In particular, $g \mapsto \mathbf{S}_\lambda(g)$ gives a representation of $\mathbf{GL}_n(\mathbf{C})$ (which turns out is polynomial). The important facts:

- $s_\lambda(x_1, \dots, x_n)$ is the character of $\mathbf{S}_\lambda(\mathbf{C}^n)$; likewise, s_λ is the character of $\mathbf{S}_\lambda(\mathbf{C}^\infty)$.
- The Schur functors $\mathbf{S}_\lambda(\mathbf{C}^n)$ with $\ell(\lambda) \le n$ give a complete set of irreducible polynomial representations of $\mathbf{GL}_n(\mathbf{C})$.

We likely will not prove these (or even give the construction for $\mathbf{S}_\lambda$) in this course, but this is one serious reason why the Schur functions are so important. An important special case of Schur functors are symmetric powers ($\mathbf{S}_k = \mathrm{Sym}^k$) and exterior powers ($\mathbf{S}_{1^k} = \bigwedge^k$).

Some references on Schur functors are [Fu1, §8], [FH, §6], and [W, §2]. □

3.2. **RSK algorithm.** The RSK algorithm converts a matrix with non-negative integer entries into a pair of SSYT of the same shape. This has a number of remarkable properties which we can use to get identities for Schur functions. The basic step is called row insertion, which we now define.

Let T be a SSYT of a partition λ, and let $k \ge 1$ be an integer. The **row insertion**, denoted $T \leftarrow k$, is another tableau defined as follows:

- Find the largest index i such that $T_{1,i-1} \le k$ (if no such index exists, set $i = 1$).
- Replace $T_{1,i}$ with k and set $k' = T_{1,i}$; we say that k is **bumping** k'. If $i = \lambda_1 + 1$, we are putting k at the end of the row, and the result is $T \leftarrow k$.
- Otherwise, let T' be the SSYT obtained by removing the first row of T. Calculate $T' \leftarrow k'$ and then add the new first row of T back to the result to get $T \leftarrow k$.

Let $I(T \leftarrow k)$ be the set of coordinates of the boxes that get replaced; this is the **insertion path**.

Example 3.2.1. Let $T = \begin{ytableau} 1 & 2 & 4 & 5 & 5 & 6 \\ 3 & 3 & 6 & 6 & 8 \\ 4 & 6 & 8 \\ 7 \\ 9 \end{ytableau}$ and $k = 4$.

We list the steps below, each time bolding the entry that gets replaced.

$$\begin{ytableau} 1 & 2 & 4 & \mathbf{5} & 5 & 6 \\ 3 & 3 & 6 & 6 & 8 \\ 4 & 6 & 8 \\ 7 \\ 8 \end{ytableau} \leftarrow \begin{ytableau} 4 \end{ytableau} = \begin{ytableau} 1 & 2 & 4 & 4 & 5 & 6 \\ 3 & 3 & \mathbf{6} & 6 & 8 \\ 4 & 6 & 8 \\ 7 \\ 8 \end{ytableau} \leftarrow \begin{ytableau} 5 \end{ytableau} = \begin{ytableau} 1 & 2 & 4 & 4 & 5 & 6 \\ 3 & 3 & 5 & 6 & 8 \\ 4 & 6 & \mathbf{8} \\ 7 \\ 8 \end{ytableau} \leftarrow \begin{ytableau} 6 \end{ytableau} = \begin{ytableau} 1 & 2 & 4 & 4 & 5 & 6 \\ 3 & 3 & 5 & 6 & 8 \\ 4 & 6 & 6 \\ 7 & 8 \\ 8 \end{ytableau}$$

The insertion path is $I(T \leftarrow 4) = \{(1,4), (2,3), (3,3), (4,2)\}$. □

Proposition 3.2.2. *$T \leftarrow k$ is a SSYT.*

Proof. By construction, the rows of $T \leftarrow k$ are weakly increasing. Also by construction, at each step, if we insert a into row i, then it can only bump a value b with $b > a$. We claim that if $(i,j), (i+1,j') \in I(T \leftarrow k)$, then $j \geq j'$. If not, then $T_{i+1,j} < T_{i,j}$ since b cannot bump the number in position $(i+1,j)$, but this contradicts that T is a SSYT. In particular, $b = T_{i,j} > T_{i,j'}$ and also $T_{i+2,j'} > T_{i+1,j'} > b$, so inserting b into position $(i+1,j')$ preserves the property of being a SSYT. □

Now we're ready to give the RSK (Robinson–Schensted–Knuth) algorithm. Let A be a matrix with non-negative integer entries (only finitely many of which are nonzero). Create a multiset of tuples (i,j) where the number of times that (i,j) appears is $A_{i,j}$. Now sort them by lexicographic order and put them as the columns of a matrix w_A with 2 rows.

Example 3.2.3. If $A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, then $w_A = \begin{pmatrix} 1 & 1 & 1 & 2 & 2 & 2 & 2 & 3 \\ 1 & 1 & 3 & 2 & 2 & 2 & 3 & 3 \end{pmatrix}$. □

Given A, we're going to create a pair of tableaux (P,Q) by induction as follows. Start with $P(0) = \varnothing$, $Q(0) = \varnothing$. Assuming $P(t)$ and $Q(t)$ are defined, let $P(t+1) = (P(t) \leftarrow (w_A)_{2,t+1})$. This value gets added to some new box; add that same box to $Q(t)$ with value $(w_A)_{1,t+1}$ to get $Q(t+1)$. When finished, the result is P (the **insertion tableau**) and Q (the **recording tableau**).

Example 3.2.4. Continuing the previous example, we list $P(t), Q(t)$ in the rows of the following table.

$P(t)$	$Q(t)$
$\ytableaushort{1}$	$\ytableaushort{1}$
$\ytableaushort{11}$	$\ytableaushort{11}$
$\ytableaushort{113}$	$\ytableaushort{111}$
$\ytableaushort{112,3}$	$\ytableaushort{111,2}$
$\ytableaushort{1122,3}$	$\ytableaushort{1112,2}$
$\ytableaushort{11222,3}$	$\ytableaushort{11122,2}$
$\ytableaushort{112223,3}$	$\ytableaushort{111222,2}$
$\ytableaushort{1122233,3}$	$\ytableaushort{1112223,2}$

The last row has the tableaux P and Q. $\square$

Lemma 3.2.5. *P and Q are both SSYT.*

Proof. P is built by successive row insertions into SSYT, so is itself a SSYT by Proposition 3.2.2. The numbers are put into Q in weakly increasing order since the first row of w_A is weakly increasing. Hence the entries of Q are weakly increasing in each row and also in each column. So it suffices to check no column has repeated entries. Suppose this is the case, i.e., that $(w_A)_{1,k} = (w_A)_{1,k+1}$ and $(w_A)_{1,k+1}$ gets placed directly below $(w_A)_{1,k}$. But we have $(w_A)_{2,k} \le (w_A)_{2,k+1}$, so this contradicts the following lemma. $\square$

Lemma 3.2.6. *Let T be a SSYT and $j \le k$. Then $I(T \leftarrow j)$ is strictly to the left of $I((T \leftarrow j) \leftarrow k)$, i.e., if $(r,s) \in I(T \leftarrow j)$ and $(r,s') \in I((T \leftarrow j) \leftarrow k)$, then $s < s'$. Furthermore, $\#I(T \leftarrow j) \ge \#I((T \leftarrow j) \leftarrow k)$.*

Proof. When inserting k into the first row of $T \leftarrow j$, k must bump a number strictly larger than itself, so gets put in a position strictly to the right of whatever was bumped by j when computing $T \leftarrow j$. The numbers j' and k' that got bumped by j and k satisfy $j' \le k'$, so we can deduce the first statement by induction on the number of rows.

For the second statement, let $r = \#I(T \leftarrow j)$, so that the last move in computing $T \leftarrow j$ was to add an element to the end of row r. If $\#I((T \leftarrow j) \leftarrow k) \ge r$, then the bump in row r happens strictly to the right of row r, which means an element was added to the end, and hence $r = \#((T \leftarrow j) \leftarrow k)$ in this case. $\square$

Theorem 3.2.7. *The RSK algorithm gives a bijection between non-negative integer matrices A with finitely many nonzero entries and pairs of SSYT (P,Q) of the same shape. Furthermore, j appears in P exactly $\sum_i A_{i,j}$ times, while i appears in Q exactly $\sum_j A_{i,j}$ times.*

Proof. The last statement is clear from our construction, so it suffices to prove that RSK gives a bijection. We just give a sketch. First, we can recover w_A from (P,Q) as follows: the last entry in the first row of w_A is the largest entry in Q, and it was added wherever the rightmost occurrence of that entry is. Remove it to get Q'. Now, consider the number in the same position in P. That gives the last entry in the second row of w_A. We can undo the row insertion procedure to get this entry out of P and get a resulting P'. Now repeat to get the rest of the columns of w_A. So the RSK algorithm is injective. We can do this procedure to any pair (P, Q), though we don't know it leads to w_A for some matrix A; surjectivity amounts to proving this is the case. We will skip this check, which amounts to reversing some of the arguments presented above. □

Corollary 3.2.8 (Cauchy identity)**.**

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

where the sum is over all partitions.

Proof. Given a non-negative integer matrix A with finitely many nonzero entries, assign to it the monomial $m(A) = \prod_{i,j}(x_i y_j)^{A_{i,j}}$. The left hand side is then $\sum_A m(A)$ since the $A_{i,j}$ can be chosen arbitrarily. Via the RSK correspondence, A goes to a pair of SSYT (P, Q), and by Theorem 3.2.7, $m(A) = x^Q y^P$, and so $\sum_A m(A) = \sum_\lambda s_\lambda(x) s_\lambda(y)$. □

Remark 3.2.9. In finitely many variables $x_1, \dots, x_n, y_1, \dots, y_m$, we can think of $\prod_{i,j}(1 - x_i y_j)^{-1}$ as the character of $\mathbf{GL}_n(\mathbf{C}) \times \mathbf{GL}_m(\mathbf{C})$ acting on $\mathrm{Sym}(\mathbf{C}^n \otimes \mathbf{C}^m)$. A similar remark can be made with infinitely many variables. The Cauchy identity gives a decomposition into Schur functors:

$$\mathrm{Sym}(\mathbf{C}^n \otimes \mathbf{C}^m) \cong \bigoplus_\lambda \mathbf{S}_\lambda(\mathbf{C}^n) \otimes \mathbf{S}_\lambda(\mathbf{C}^m).$$

where the sum is over all partitions (or just those with $\ell(\lambda) \le \min(m, n)$). For an algebraic approach to this identity (avoiding symmetric functions), see [Ho, §2]. □

Corollary 3.2.10. *The Schur functions form an orthonormal basis with respect to $\langle,\rangle$, i.e., $\langle s_\lambda, s_\mu\rangle = \delta_{\lambda,\mu}$.*

Proof. Immediate from the Cauchy identity and Lemma 2.6.1. □

Corollary 3.2.11. *We have*

$$h_\mu = \sum_\lambda K_{\lambda,\mu} s_\lambda.$$

Proof. Write $h_\mu = \sum_\lambda a_{\lambda,\mu} s_\lambda$ for some coefficients a. By Corollary 3.2.10, $a_{\lambda,\mu} = \langle h_\mu, s_\lambda\rangle$. By definition, we have $s_\lambda = \sum_\nu K_{\lambda,\nu} m_\nu$. But also by definition, $\langle h_\mu, m_\nu\rangle = \delta_{\mu,\nu}$. Hence, $a_{\lambda,\mu} = K_{\lambda,\mu}$. □

Remark 3.2.12. Since h_μ is the character of $\mathrm{Sym}^{\mu_1} \otimes \cdots \otimes \mathrm{Sym}^{\mu_k}$, this formula explains how to decompose tensor products of symmetric powers into Schur functors. □

An important symmetry of the RSK algorithm is the following, but we omit the proof.

Theorem 3.2.13. *If $A \mapsto (P, Q)$ under RSK, then $A^T \mapsto (Q, P)$. In particular, RSK gives a bijection between symmetric non-negative integer matrices with finitely many nonzero entries and the set of all SSYT.*

3.3. **Dual RSK algorithm.** There is a variant of the RSK algorithm for $(0,1)$-matrices called **dual RSK**. The change occurs in the definition of row insertion: instead of k bumping the leftmost entry that is $> k$, it bumps the leftmost entry that is $\geq k$. This can be analyzed like the RSK algorithm, but we will omit this and state its consequences.

Theorem 3.3.1. *The dual RSK algorithm gives a bijection between $(0,1)$-matrices A with finitely many nonzero entries and pairs (P,Q) where P and Q are tableaux of the same shape, and $P^\dagger$ and Q are SSYT. Furthermore, the type of P is given by the column sums of A and the type of Q is given by the row sums of A.*

Corollary 3.3.2 (Dual Cauchy identity)**.**

$$\prod_{i,j}(1+x_iy_j) = \sum_\lambda s_\lambda(x)s_{\lambda^\dagger}(y).$$

Remark 3.3.3. In finitely many variables $x_1,\dots,x_n,y_1,\dots,y_m$, we can think of $\prod_{i,j}(1+x_iy_j)$ as the character of $\mathbf{GL}_n(\mathbf{C}) \times \mathbf{GL}_m(\mathbf{C})$ acting on the exterior algebra $\bigwedge^\bullet(\mathbf{C}^n \otimes \mathbf{C}^m)$. A similar remark can be made with infinitely many variables. The Cauchy identity gives a decomposition into Schur functors:

$$\bigwedge^\bullet(\mathbf{C}^n \otimes \mathbf{C}^m) \cong \bigoplus_\lambda \mathbf{S}_\lambda(\mathbf{C}^n) \otimes \mathbf{S}_{\lambda^\dagger}(\mathbf{C}^m).$$

where the sum is over all partitions (or just those with $\ell(\lambda) \leq n$ and $\lambda_1 \leq m$). ∎

Lemma 3.3.4. *Let ω_y be the action of ω on the second copy of Λ in $\Lambda \otimes \Lambda$. Then*

$$\prod_{i,j}(1+x_iy_j) = \omega_y \prod_{i,j}(1-x_iy_j)^{-1}.$$

Proof. We have

$$\begin{aligned}\omega_y \prod_{i,j}(1-x_iy_j)^{-1} &= \omega_y \sum_\lambda m_\lambda(x)h_\lambda(y)\\ &= \sum_\lambda m_\lambda(x)e_\lambda(y)\end{aligned}$$

where the first equality is Proposition 2.6.2 and the second is Theorem 2.3.1. Now we follow the proof of Proposition 2.6.2:

$$\prod_i\prod_j(1+x_iy_j) = \prod_i\sum_n e_n(y)x_i^n = \sum_\lambda m_\lambda(x)e_\lambda(y).$$

Combining these two gives the result. ∎

Corollary 3.3.5. $\omega(s_\lambda) = s_{\lambda^\dagger}$.

Proof.

$$\omega_y\sum_\lambda s_\lambda(x)s_\lambda(y) = \omega_y\prod_{i,j}(1-x_iy_j)^{-1} = \prod_{i,j}(1+x_iy_j) = \sum_\lambda s_\lambda(x)s_{\lambda^\dagger}(y).$$

The $s_\lambda(x)$ are linearly independent, so $\omega_y(s_\lambda(y)) = s_{\lambda^\dagger}(y)$. ∎

3.4. **Determinantal formula.** Let $\alpha = (\alpha_1, \dots, \alpha_n)$ be a non-negative integer sequence. Define

$$a_\alpha = \det(x_i^{\alpha_j})_{i,j=1}^n = \det \begin{pmatrix} x_1^{\alpha_1} & x_1^{\alpha_2} & \cdots & x_1^{\alpha_n} \\ x_2^{\alpha_1} & x_2^{\alpha_2} & \cdots & x_2^{\alpha_n} \\ \vdots & & & \vdots \\ x_n^{\alpha_1} & x_n^{\alpha_2} & \cdots & x_n^{\alpha_n} \end{pmatrix}.$$

Note that a_α is *skew-symmetric*: if we permute a_α by a permutation σ, then it changes by $\mathrm{sgn}(\sigma)$. Let $\rho = (n-1, n-2, \dots, 1, 0)$.

Lemma 3.4.1. (a) $\prod_{1 \le i < j \le n}(x_i - x_j)$ *divides every skew-symmetric polynomial in* $x_1, \dots, x_n$.
(b) $a_\rho = \prod_{1 \le i < j \le n}(x_i - x_j)$.

Proof. (a) Let $f(x_1, \dots, x_n)$ be skew-symmetric and let σ be the transposition (i, j). Then $\sigma f = -f$. However, σf and f are the same if we replace x_j by x_i, so this says that specializing x_j to x_i gives 0, i.e., f is divisible by $(x_i - x_j)$. This is true for any i, j, so this proves (a).

(b) a_ρ is divisible by $\prod_{1 \le i < j \le n}(x_i - x_j)$ since it is skew-symmetric. But also note that both are polynomials of degree $1 + 2 + \cdots + (n-1) = \binom{n}{2}$, so they are equal up to some integer multiple. The coefficient of $x_1^{n-1}x_2^{n-2}\cdots x_{n-1}$ for both is 1, so they are actually the same. $\square$

Define $\alpha + \beta = (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)$.

Given $\nu \subseteq \lambda$, let $K_{\lambda/\nu,\mu}$ be the number of SSYT of skew shape λ/ν and type μ.

Lemma 3.4.2. $a_{\nu+\rho}e_\mu = \sum_\lambda K_{\lambda^\dagger/\nu^\dagger,\mu} a_\lambda$.

Proof. First, we claim that given a partition μ, the coefficient of $x^{\lambda+\rho}$ in $a_{\nu+\rho}e_\mu$ is $K_{\lambda^\dagger/\nu^\dagger,\mu}$. To get a monomial in $a_\rho e_{\mu_1} \cdots e_{\mu_k}$, we pick a monomial x^β in $a_{\nu+\rho}$ and successively multiply it by monomials $x^{\alpha(1)}, \dots, x^{\alpha(k)}$ where $x^{\alpha(i)}$ is taken from e_{μ_i}. Note that each partial product $a_{\nu+\rho}e_{\mu_1} \cdots e_{\mu_r}$ is skew-symmetric, so each of its monomials have distinct exponents on all of the variables. So, we're only interested in choosing $x^{\alpha(r+1)}$ so that the product $x^\beta x^{\alpha(1)} \cdots x^{\alpha(r+1)}$ has all exponents distinct. Since $x^{\alpha(r+1)}$ is a product of distinct variables, the relative order of the exponents remains the same. Since we're interested in the coefficient of $x^{\lambda+\rho}$, whose exponents are strictly decreasing, we can only get to this if β is strictly decreasing and the $x^{\alpha(r+1)}$ is chosen so that the exponents of $x^\beta x^{\alpha(1)} \cdots x^{\alpha(r+1)}$ are strictly decreasing. The only β that works is $\nu + \rho$, and so the condition is the same as requiring that $\gamma(r+1) := \nu + \alpha(1) + \cdots + \alpha(r+1)$ is a partition for each r.

Note then we get a sequence of partitions

$$\nu = \gamma(0) \subseteq \gamma(1) \subseteq \gamma(2) \subseteq \cdots \subseteq \gamma(n) = \lambda$$

such that the difference $\gamma(r+1)/\gamma(r)$ only has boxes in different rows, and that conversely, given such a sequence, we can find a sequence of monomials that corresponds to this. However, this sequence is also equivalently encoding a labeling of the Young diagram of λ/ν which is weakly increasing in columns and strictly increasing in rows, i.e., taking the transpose gives a SSYT of $\lambda^\dagger/\nu^\dagger$ and type μ. So the claim is proven.

Finally, consider the difference

$$a_{\nu+\rho}e_\mu - \sum_\lambda K_{\lambda^\dagger/\nu^\dagger,\mu} a_{\lambda+\rho}.$$

If $\lambda' \neq \lambda$, then the coefficient of $x^{\lambda+\rho}$ in $a_{\lambda'+\rho}$ is 0, so the coefficient of each $x^{\lambda+\rho}$ of this difference is 0. However, any nonzero skew-symmetric function of degree $|\lambda| + \binom{n}{2}$ has a monomial of the form $x^{\lambda+\rho}$ for some partition λ, so we conclude then that the difference is 0. $\square$

Corollary 3.4.3. *Given a partition λ,*

$$s_\lambda(x_1, \dots, x_n) = \frac{a_{\lambda+\rho}}{a_\rho}.$$

Proof. Take $\nu = \varnothing$ in Lemma 3.4.2 and divide both sides by a_ρ to get

$$e_\mu = \sum_\lambda K_{\lambda^\dagger,\mu} \frac{a_{\lambda+\rho}}{a_\rho}.$$

However, we also have an expression

$$e_\mu = \sum_\lambda K_{\lambda^\dagger,\mu} s_\lambda$$

by applying ω to Corollary 3.2.11. The s_λ and the e_μ are both bases, so we can invert the matrix $(K_{\lambda^\dagger,\mu})$ to conclude that $s_\lambda = a_{\lambda+\rho}/a_\rho$. $\square$

Remark 3.4.4. (For those familiar with Lie theory.) The formula above is really an instance of the Weyl character formula for the Lie algebra $\mathfrak{gl}_n(\mathbf{C})$ (or actually, $\mathfrak{sl}_n(\mathbf{C})$ since it's semisimple, but we'll phrase everything in terms of $\mathfrak{gl}_n(\mathbf{C})$ because it's cleaner). To translate, first note that we can evaluate determinants by using a sum over all permutations, and in our case this gives

$$a_\alpha = \sum_{\sigma \in \Sigma_n} \mathrm{sgn}(\sigma)\sigma(x^\alpha).$$

In the context of $\mathfrak{gl}_n(\mathbf{C})$, (integral) weights are identified with elements of $\mathbf{Z}^n$, while dominant weights are the weakly decreasing ones. Also, ρ is used here to have the same meaning as in Lie theory: it is the sum of the fundamental dominant weights. Finally, Σ_n is the Weyl group of $\mathfrak{gl}_n(\mathbf{C})$, and $s_\lambda(x_1, \dots, x_n)$ is the character of the irreducible representation with highest weight λ. Then our formula becomes

$$s_\lambda(x_1, \dots, x_n) = \frac{\sum_{\sigma \in \Sigma_n} \mathrm{sgn}(\sigma)\sigma(x^{\lambda+\rho})}{\sum_{\sigma \in \Sigma_n} \mathrm{sgn}(\sigma)\sigma(x^\rho)}$$

which is the Weyl character formula, as might be found in [Hu, §24.3]. $\square$

3.5. **Multiplying Schur functions, Pieri rule.**

Corollary 3.5.1. $s_\nu e_\mu = \sum_\lambda K_{\lambda^\dagger/\nu^\dagger,\mu} s_\lambda$.

Proof. By Lemma 3.4.2, we have

$$a_{\nu+\rho} e_\mu = \sum_\lambda K_{\lambda^\dagger/\nu^\dagger,\mu} a_\lambda.$$

Divide both sides by a_ρ and use Corollary 3.4.3 to get the desired identity in finitely many variables $x_1, \dots, x_n$. Since it holds for every n, it also holds when $n = \infty$. $\square$

Corollary 3.5.2. $s_\nu h_\mu = \sum_\lambda K_{\lambda/\nu,\mu} s_\lambda$.

Proof. Apply ω to Corollary 3.5.1 and use Corollary 3.3.5 to get the desired identity with $\nu^\dagger$ and $\lambda^\dagger$ in place of ν and λ. But that's just an issue of indexing, so we get the desired identity. □

Theorem 3.5.3. *For any $f \in \Lambda$, we have*

$$\langle f s_\nu, s_\lambda \rangle = \langle f, s_{\lambda/\nu} \rangle.$$

Proof. Both sides of the equation are linear in f, so it suffices to prove this when f ranges over a particular basis, and we choose h_μ. By Corollary 3.5.2, $\langle h_\mu s_\nu, s_\lambda \rangle = K_{\lambda/\nu,\mu}$. This is the coefficient of m_μ in $s_{\lambda/\nu}$. Since $\langle h_\nu, h_\mu \rangle = \delta_{\nu,\mu}$, we conclude that $K_{\lambda/\nu,\mu} = \langle h_\mu, s_{\lambda/\nu} \rangle$. □

Of particular note is when $f = s_\mu$. Since the s_λ are a basis, we have unique expressions

$$s_\mu s_\nu = \sum_\lambda c^\lambda_{\mu,\nu} s_\lambda, \tag{3.5.4}$$

and the $c^\lambda_{\mu,\nu}$ are called **Littlewood–Richardson coefficients**. We will see some special cases soon and study this in more depth later. Since the s_λ are an orthonormal basis, we get

$$c^\lambda_{\mu,\nu} = \langle s_\mu s_\nu, s_\lambda \rangle = \langle s_\mu, s_{\lambda/\nu} \rangle. \tag{3.5.5}$$

In particular, we also have an identity

$$s_{\lambda/\nu} = \sum_\mu c^\lambda_{\mu,\nu} s_\mu.$$

From the definition, we have

$$c^\lambda_{\mu,\nu} = c^\lambda_{\nu,\mu}.$$

Applying ω to (3.5.4), we get

$$c^\lambda_{\mu,\nu} = c^{\lambda^\dagger}_{\nu^\dagger,\mu^\dagger}. \tag{3.5.6}$$

We can give an interpretation for the Littlewood–Richardson coefficients in the special case where μ (or ν) has a single part or all parts equal to 1. Say that λ/ν is a **horizontal strip** if no column in the skew Young diagram of λ/ν contains 2 or more boxes. Similarly, say that λ/ν is a **vertical strip** if no row in the skew Young diagram of λ/ν contains 2 or more boxes.

Theorem 3.5.7 (Pieri rule)**.** • *If $\mu = (1^k)$, then*

$$c^\lambda_{(1^k),\nu} = \begin{cases} 1 & \text{if } |\lambda| = |\nu| + k \text{ and } \lambda/\nu \text{ is a vertical strip} \\ 0 & \text{otherwise} \end{cases}.$$

In other words,

$$s_\nu s_{1^k} = \sum_\lambda s_\lambda$$

where the sum is over all λ such that λ/ν is a vertical strip of size k.

• *If $\mu = (k)$, then*

$$c^\lambda_{(k),\nu} = \begin{cases} 1 & \text{if } |\lambda| = |\nu| + k \text{ and } \lambda/\nu \text{ is a horizontal strip} \\ 0 & \text{otherwise} \end{cases}.$$

In other words,

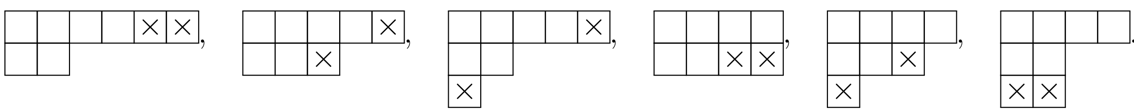
$$s_\nu s_k = \sum_\lambda s_\lambda$$

where the sum is over all λ such that λ/ν is a horizontal strip of size k.

Proof. Since $s_{1^k} = e_k$, we have $s_\nu s_{1^k} = \sum_\lambda K_{\lambda^\dagger/\nu^\dagger,k} s_\lambda$ by Corollary 3.5.1. So $c^\lambda_{1^k,\nu}$ is the number of SSYT of shape $\lambda^\dagger/\nu^\dagger$ using k 1's. There is at most one such SSYT, and it exists exactly when no two boxes of $\lambda^\dagger/\nu^\dagger$ are in the same column, i.e., λ/ν is a vertical strip.

The proof the second identity is similar, or can be obtained by using ω. $\square$

Example 3.5.8. To multiply s_λ by s_k, it suffices to enumerate all partitions that we can get by adding k boxes to the Young diagram of λ, no two of which are in the same column. For example, here we have drawn all of the ways to add 2 boxes to $(4,2)$:



So

$$s_{4,2}s_2 = s_{6,2} + s_{5,3} + s_{5,2,1} + s_{4,4} + s_{4,3,1} + s_{4,2,2}.$$ $\square$

Recall that for $|\lambda| = n$, $f^\lambda = K_{\lambda,1^n}$ is the number of standard Young tableaux of shape λ.

Corollary 3.5.9. $s_1^n = \sum_{\lambda \vdash n} f^\lambda s_\lambda$.

Proof. The Pieri rule says that to multiply s_1^n, we first enumerate all sequences $\lambda^{(1)} \subset \lambda^{(2)} \subset \cdots \subset \lambda^{(n)}$ where $|\lambda^{(i)}| = i$. Then the result is the sum of s_λ with multiplicity given by the number of sequences with $\lambda^{(n)} = \lambda$. But such sequences are in bijection with standard Young tableaux: label the unique box in $\lambda^{(i)}/\lambda^{(i-1)}$ with i. $\square$

Remark 3.5.10. From the interpretation of s_λ as the character of an irreducible representation $\mathbf{S}_\lambda$, and the fact that polynomial representations are direct sums of irreducible ones, we can reinterpret the Littlewood–Richardson coefficient as the multiplicity of $\mathbf{S}_\lambda$ in the decomposition of the tensor product of $\mathbf{S}_\mu \otimes \mathbf{S}_\nu$. From this, it is immediate that $c^\lambda_{\mu,\nu} \geq 0$. This non-negativity is not easy to see from our development so far otherwise.

The Pieri rule describes the decomposition of the tensor product of $\mathbf{S}_\lambda$ with an exterior power $\bigwedge^k$, respectively, a symmetric power Sym^k.

The decomposition of s_1^n can be interpreted as a decomposition of the tensor power of a vector space $(\mathbf{C}^d)^{\otimes n} = \bigoplus_{\substack{\lambda \vdash n \\ \ell(\lambda) \leq d}} \mathbf{S}_\lambda(\mathbf{C}^d)^{\oplus f^\lambda}$. Hence the multiplicity space of $\mathbf{S}_\lambda(\mathbf{C}^d)$ has dimension f^λ. When we discuss Schur–Weyl duality later, we will see that f^λ is the dimension of an irreducible representation of Σ_d. $\square$

Theorem 3.5.11. $\omega(s_{\lambda/\nu}) = s_{\lambda^\dagger/\nu^\dagger}$.

Proof. First, we have

$$\begin{aligned}
\langle s_{\mu^\dagger}, s_{\lambda^\dagger/\nu^\dagger}\rangle &= \langle s_{\mu^\dagger} s_{\nu^\dagger}, s_{\lambda^\dagger}\rangle && \text{(Theorem 3.5.3)}\\
&= \langle \omega(s_\mu s_\nu), \omega(s_\lambda)\rangle && \text{(Corollary 3.3.5)}\\
&= \langle s_\mu s_\nu, s_\lambda\rangle && \text{(Corollary 2.6.5)}\\
&= \langle s_\mu, s_{\lambda/\nu}\rangle && \text{(Theorem 3.5.3)}\\
&= \langle \omega(s_\mu), \omega(s_{\lambda/\nu})\rangle && \text{(Corollary 2.6.5)}\\
&= \langle s_{\mu^\dagger}, \omega(s_{\lambda/\nu})\rangle. && \text{(Corollary 3.3.5)}
\end{aligned}$$

The pairing is nondegenerate, so if we fix λ, ν and allow μ to vary, we get $s_{\lambda^\dagger/\nu^\dagger} = \omega(s_{\lambda/\nu})$. □

3.6. **Jacobi–Trudi identity.** Corollary 3.5.2 and Corollary 3.5.1 with $\nu = \varnothing$ explain how to rewrite the h_μ and e_μ bases in terms of the Schur basis using Kostka numbers. The Jacobi–Trudi identities go the other way around. We'll do something more general with skew Schur functions though.

Theorem 3.6.1. *Pick $\mu \subseteq \lambda$ with $\ell(\lambda) \le n$. Set $h_i = 0$ if $i < 0$. Then*

$$s_{\lambda/\mu} = \det(h_{\lambda_i-\mu_j-i+j})_{i,j=1}^n = \det\begin{pmatrix} h_{\lambda_1-\mu_1} & h_{\lambda_1-\mu_2+1} & h_{\lambda_1-\mu_3+2} & \cdots & h_{\lambda_1-\mu_n+n-1}\\ h_{\lambda_2-\mu_1-1} & h_{\lambda_2-\mu_2} & h_{\lambda_2-\mu_3+1} & \cdots & h_{\lambda_2-\mu_n+n-2}\\ \vdots & & & & \vdots\\ h_{\lambda_n-\mu_1-n+1} & h_{\lambda_n-\mu_2-n+2} & h_{\lambda_n-\mu_3-n+3} & \cdots & h_{\lambda_n-\mu_n}\end{pmatrix}$$

$$s_{\lambda/\mu} = \det(e_{\lambda_i^\dagger-\mu_j^\dagger-i+j})_{i,j=1}^n = \det\begin{pmatrix} e_{\lambda_1^\dagger-\mu_1^\dagger} & e_{\lambda_1^\dagger-\mu_2^\dagger+1} & e_{\lambda_1^\dagger-\mu_3^\dagger+2} & \cdots & e_{\lambda_1^\dagger-\mu_n^\dagger+n-1}\\ e_{\lambda_2^\dagger-\mu_1^\dagger-1} & e_{\lambda_2^\dagger-\mu_2^\dagger} & e_{\lambda_2^\dagger-\mu_3^\dagger+1} & \cdots & e_{\lambda_2^\dagger-\mu_n^\dagger+n-2}\\ \vdots & & & & \vdots\\ e_{\lambda_n^\dagger-\mu_1^\dagger-n+1} & e_{\lambda_n^\dagger-\mu_2^\dagger-n+2} & e_{\lambda_n^\dagger-\mu_3^\dagger-n+3} & \cdots & e_{\lambda_n^\dagger-\mu_n^\dagger}.\end{pmatrix}$$

Proof. First, note that if we move from n to $n+1$, the determinant does not change because the new row added is $(0, 0, \dots, 0, 1)$. Fix μ and work in $\Lambda(N) \otimes \Lambda(N)$ where $N \ge n$. We have

$$\begin{aligned}
\sum_\lambda s_{\lambda/\mu}(x) s_\lambda(y) &= \sum_\lambda \sum_\nu c^\lambda_{\mu,\nu} s_\nu(x) s_\lambda(y) && \text{(by (3.5.5))}\\
&= \sum_\nu s_\nu(x) s_\nu(y) s_\mu(y) && \text{(by (3.5.5))}\\
&= s_\mu(y) \prod_{i,j} (1 - x_i y_j)^{-1} && \text{(Corollary 3.2.8)}\\
&= s_\mu(y) \sum_\nu h_\nu(x) m_\mu(y). && \text{(Proposition 2.6.2)}
\end{aligned}$$

Multiply both sides by $a_\rho(y)$ to get

$$\begin{aligned}\sum_\lambda s_{\lambda/\mu}(x)a_{\lambda+\rho}(y) &= a_{\mu+\rho}(y)\sum_\nu h_\nu(x)m_\mu(y) && \text{(Corollary 3.4.3)}\\ &= \left(\sum_{\sigma\in\Sigma_N}\operatorname{sgn}(\sigma)\sigma(y^{\mu+\rho})\right)\left(\sum_{\alpha\in\mathbf{Z}^N_{\ge 0}} h_\alpha(x)y^\alpha\right)\\ &= \sum_{\sigma\in\Sigma_N}\sum_\alpha \operatorname{sgn}(\sigma)h_\alpha(x)y^{\alpha+\sigma(\mu+\rho)}.\end{aligned}$$

Now take the coefficient of $y^{\lambda+\rho}$. The left hand side gives $s_{\lambda/\mu}(x)$, while the right hand side gives

$$\sum_{\sigma\in\Sigma_N}\operatorname{sgn}(\sigma)h_{\lambda+\rho-\sigma(\mu+\rho)}(x) = \det(h_{\lambda_i-\mu_j-i+j})_{i,j=1}^N.$$

So we get the desired identity in N variables; let $N\to\infty$ to get it in general.

The second identity follows from the first by applying ω. □

Remark 3.6.2. It is possible to give an elegant combinatorial proof of the Jacobi–Trudi identity by interpreting SSYT as non-crossing lattice paths and using the Gessel–Viennot method of enumerating non-crossing lattice paths. The interested reader can find this argument in [Sta, §7.16]. □

Remark 3.6.3. The Jacobi–Trudi identity can be given an algebraic meaning. First, we expand it (not the same way used in the proof above):

$$\sum_{\sigma\in\Sigma_n}\operatorname{sgn}(\sigma)h_{\sigma(\lambda+\rho)-(\mu+\rho)}(x) = \det(h_{\lambda_i-\mu_j-i+j})_{i,j=1}^n.$$

However, we need a refinement of the left side. Given a permutation σ, define its length to be

$$\ell(\sigma) = \#\{i<j \mid \sigma(i)>\sigma(j)\}.$$

A basic fact is that $(-1)^{\ell(\sigma)} = \operatorname{sgn}(\sigma)$. So we can rewrite it as

$$\sum_{i\ge 0}\sum_{\sigma,\ \ell(\sigma)=i}(-1)^{\ell(\sigma)}h_{\sigma(\lambda+\rho)-(\mu+\rho)}(x) = s_{\lambda/\mu}$$

Recall that h_α is the character of $\mathrm{Sym}^\alpha := \mathrm{Sym}^{\alpha_1}\otimes\cdots\otimes\mathrm{Sym}^{\alpha_n}$. The left hand side can be interpreted as the (character version of the) Euler characteristic of a chain complex $\mathbf{F}_\bullet$ whose ith term is

$$\mathbf{F}_i = \bigoplus_{\ell(\sigma)=i}\mathrm{Sym}^{\sigma(\lambda+\rho)-(\mu+\rho)}.$$

The differentials of such a chain complex aren't determined by its character, but there exist a choice that makes $\mathbf{F}_\bullet$ exact except in degree 0, in which case it is a skew version of the Schur functor $\mathbf{S}_{\lambda/\mu}$. Hence, the Jacobi–Trudi identity is encoding the existence of a certain kind of resolution of $\mathbf{S}_{\lambda/\mu}$ by tensor products of symmetric powers (a version also exists using exterior powers). This complex was constructed by Akin [A] and Zelevinsky [Z] when $\mu=\varnothing$ (the construction of Zelevinsky can be adapted to handle arbitrary μ). □

4. Representation theory of the symmetric groups

4.1. **Complex representations of finite groups.** In this section, we recall the basic results about representation theory of finite groups. A good reference is Serre's book [Se].

Let G be a finite group. A **(complex) representation** is a homomorphism

$$\rho \colon G \to \mathbf{GL}(V)$$

for some complex vector space V, where $\mathbf{GL}(V)$ is the group of invertible linear operators on V. Equivalently, giving a representation is the same as giving a linear action of G on V, i.e., a multiplication $g \cdot v$ for $g \in G$ and $v \in V$ such that $g \cdot (v + v') = g \cdot v + g \cdot v'$ and $(gg') \cdot v = g \cdot (g' \cdot v)$, and $1 \cdot v = v$. A **subrepresentation** of V is a subspace $W \subseteq V$ such that $\rho(g)w \in W$ whenever $g \in G$ and $w \in W$. A nonzero representation V is **irreducible** if it has no nonzero subrepresentations other than itself.

We can define direct sums and isomorphisms of representations as we did before.

Theorem 4.1.1 (Maschke). *Every finite-dimensional representation of G is isomorphic to a direct sum of irreducible representations.*

The **character** of ρ is the function $\chi_\rho \colon G \to \mathbf{C}$ defined by $\chi_\rho(g) = \mathrm{Tr}(\rho(g))$. This is constant on conjugacy classes of g:

$$\chi_\rho(hgh^{-1}) = \mathrm{Tr}(\rho(h)\rho(g)\rho(h)^{-1}) = \mathrm{Tr}(\rho(g)) = \chi_\rho(g).$$

We let $\mathrm{CF}(G)$ denote the set of functions $G \to \mathbf{C}$ which are constant on conjugacy classes and define a bilinear pairing on $\mathrm{CF}(G)$:

$$(\varphi, \psi)_G = \frac{1}{|G|} \sum_{g \in G} \varphi(g)\overline{\psi(g)}$$

where the overline means complex conjugation. If we don't need to specify G, we'll just write $(,)$.

Theorem 4.1.2.
- *$(,)$ is an inner product on $\mathrm{CF}(G)$.*
- *The characters of the irreducible representations form an orthonormal basis for $\mathrm{CF}(G)$. In particular, the number of irreducible representations of G is equal to the number of conjugacy classes of G.*
- *If two representations have the same character, then they are isomorphic.*

Given a subgroup $H \subseteq G$, any representation ρ of G becomes a representation of H by restricting the map. This is called the **restriction** of ρ, and is denoted $\mathrm{Res}^G_H \rho$. In fact, restriction makes sense for any class functions.

On the other hand, given a representation ρ of H, one can define the **induced representation** $\mathrm{Ind}^G_H \rho$ which is a representation of G. The construction is probably clearest after introducing group algebras, but we will omit this. Again, it makes sense for any class function. The important fact is Frobenius reciprocity, which says that given $\varphi \in \mathrm{CF}(H)$ and $\psi \in \mathrm{CF}(G)$, we have

$$(\mathrm{Ind}^G_H \varphi, \psi)_G = (\varphi, \mathrm{Res}^G_H \psi)_H.$$

Finally, suppose we are given two groups G_1, G_2 and representations ρ_1, ρ_2 on vector spaces V_1, V_2. Then $G_1 \times G_2$ has a linear action on $V_1 \otimes V_2$ by

$$(g_1, g_2) \cdot \sum_i v(1)_i \otimes v(2)_i = \sum_i g_1 \cdot v(1)_i \otimes g_2 \cdot v(2)_i.$$

Hence we get a representation $\rho_1 \otimes \rho_2$ of $G_1 \times G_2$. Its character is given by

$$\chi_{\rho_1 \otimes \rho_2}(g_1, g_2) = \chi_{\rho_1}(g_1)\chi_{\rho_2}(g_2).$$

4.2. **Symmetric groups.** We now focus on the symmetric groups $G = \Sigma_n$. A nontrivial fact[4] is that all of its irreducible representations can be realized using only rational numbers, and hence the characters of its representations are always rational-valued (and hence, integer-valued[5]). Given this fact, we define CF_n to be the space of rational-valued class functions on Σ_n.

First, every permutation $\sigma \in \Sigma_n$ has a decomposition as a product of disjoint cycles (a cycle, denoted $(i_1, i_2, \dots, i_k)$, is the permutation which sends i_j to i_{j+1} for $j < k$ and i_k to i_1), and the lengths of these cycles (**cycle type**) arranged in decreasing order gives a partition, which we denote $t(\sigma)$.

Lemma 4.2.1.
- *The conjugacy classes of Σ_n are naturally indexed by partitions of n. In particular, the number of irreducible representations of Σ_n is the partition number $p(n)$.*
- *The conjugacy class assigned to λ has size $n!/z_\lambda$.*

Proof. For the first point, use the identity

$$\tau(i_1, i_2, \dots, i_k)\tau^{-1} = (\tau(i_1), \tau(i_2), \dots, \tau(i_k)).$$

Now we prove the second point. Given a permutation σ of cycle type λ, we claim that the centralizer of σ has size $z_\lambda = \prod_i m_i(\lambda)! i^{m_i(\lambda)}$. To see this, note that a cycle $(i_1, i_2, \dots, i_k)$ is equal to a cyclic shift $(i_r, i_{r+1}, \dots, i_k, i_1, \dots, i_{r-1})$. So any τ that sends a cycle of σ to any cyclic shift of another cycle of the same length is in the centralizer, and there are z_λ such τ. □

Given a partition λ, let 1_λ denote the class function which is 1 on all permutations with cycle type λ and 0 on all other permutations.

Corollary 4.2.2. *Given partitions λ, μ of n, we have $(1_\lambda, 1_\mu)_{\Sigma_n} = z_\lambda^{-1}\delta_{\lambda,\mu}$.*

Proof. If $\lambda \neq \mu$, then the definition of the pairing shows that $(1_\lambda, 1_\mu) = 0$. Otherwise, $(1_\lambda, 1_\lambda) = \frac{1}{n!}c$ where c is the size of the conjugacy class of cycle type λ, which we just said is $n!/z_\lambda$. □

Next, given positive integers n, m, we can think of $\Sigma_n \times \Sigma_m$ as a subgroup of Σ_{n+m} if we identify Σ_n with the subgroup which is the identity on $n+1, \dots, n+m$ and if we identify Σ_m with the subgroup which is the identity on $1, \dots, n$. Define an induction product

$$\circ \colon \mathrm{CF}_n \times \mathrm{CF}_m \to \mathrm{CF}_{n+m}$$
$$\varphi \circ \psi = \mathrm{Ind}_{\Sigma_n \times \Sigma_m}^{\Sigma_{n+m}}(\varphi \otimes \psi).$$

This turns $\bigoplus_{n \geq 0} \mathrm{CF}_n$ into a commutative ring (though it requires verification). Our next task is to show that it is isomorphic to $\Lambda_{\mathbf{Q}}$.

[4] which can be proven by providing an explicit construction, such as via Specht modules

[5] A general fact is that the character of a representation of a finite group is valued in algebraic integers, i.e., are roots of a monic integer polynomial equation: since $\rho(g)$ is finite order, all of its eigenvalues are roots of unity, and algebraic integers are closed under addition, so the trace is always an algebraic integer. The only algebraic integers in the rational numbers are integers.

4.3. **The characteristic map.** The **Frobenius characteristic map** is the linear function

$$\operatorname{ch}\colon \mathrm{CF}_n \to \Lambda_{\mathbf{Q},n}$$
$$\operatorname{ch}(\varphi) = \frac{1}{n!}\sum_{\sigma\in\Sigma_n} \varphi(\sigma) p_{t(\sigma)}$$

where recall that p is the power sum symmetric function. Alternatively, if we set $\varphi(\lambda)$ to be the value of φ on any permutation with cycle type λ, then $\operatorname{ch}(\varphi) = \sum_\lambda z_\lambda^{-1}\varphi(\lambda)p_\lambda$ by Lemma 4.2.1. Put these together to define a linear function

$$\operatorname{ch}\colon \bigoplus_{n\ge 0} \mathrm{CF}_n \to \Lambda_{\mathbf{Q}}.$$

Proposition 4.3.1. ch *is an isometry, i.e., given* $\varphi, \psi \in \mathrm{CF}_n$,

$$(\varphi,\psi)_{\Sigma_n} = \langle \operatorname{ch}(\varphi), \operatorname{ch}(\psi)\rangle.$$

Proof. Given that the conjugacy class of λ has size $n!/z_\lambda$, we have

$$\begin{aligned}\langle \operatorname{ch}(\varphi), \operatorname{ch}(\psi)\rangle &= \langle \sum_\lambda z_\lambda^{-1}\varphi(\lambda)p_\lambda, \sum_\mu z_\mu^{-1}\psi(\mu)p_\mu\rangle \\ &= \sum_\lambda z_\lambda^{-1}\varphi(\lambda)\psi(\lambda) \\ &= (\varphi,\psi)_{\Sigma_n},\end{aligned}$$

where the second equality is orthogonality of the p_λ (Proposition 2.6.4) and the third equality is the definition of the inner product for class functions (here we use that the class functions are rational-valued, and hence complex conjugation is unnecessary). ∎

Proposition 4.3.2. ch *is a ring isomorphism, i.e., given* $\varphi \in \mathrm{CF}_n$ *and* $\psi \in \mathrm{CF}_m$, *we have*

$$\operatorname{ch}(\varphi\circ\psi) = \operatorname{ch}(\varphi)\operatorname{ch}(\psi).$$

Proof. We claim that $1_\lambda \circ 1_\mu = \frac{z_{\lambda+\mu}}{z_\lambda z_\mu} 1_{\lambda+\mu}$. To see this, let ν be any partition of $n+m$. Then by Frobenius reciprocity and Corollary 4.2.2,

$$(\operatorname{Ind}_{\Sigma_n\times\Sigma_m}^{\Sigma_{n+m}} 1_\lambda\otimes 1_\mu, 1_\nu)_{\Sigma_{n+m}} = (1_\lambda\otimes 1_\mu, \operatorname{Res}_{\Sigma_n\times\Sigma_m}^{\Sigma_{n+m}} 1_\nu)_{\Sigma_n\times\Sigma_m} = \frac{\delta_{\lambda+\mu,\nu}}{z_\lambda z_\mu}.$$

Hence $1_\lambda\circ 1_\mu = c1_{\lambda+\mu}$ where $c = \frac{(1_\lambda\circ 1_\mu, 1_{\lambda+\mu})}{(1_{\lambda+\mu},1_{\lambda+\mu})} = \frac{z_{\lambda+\mu}}{z_\lambda z_\mu}$.

Next, $\operatorname{ch}(1_\lambda) = p_\lambda/z_\lambda$, so we see that $\operatorname{ch}(1_\lambda\circ 1_\mu) = \operatorname{ch}(1_\lambda)\operatorname{ch}(1_\mu)$. Since the 1_λ form a basis for CF, we conclude that ch is a ring homomorphism. Finally, the 1_λ map to a basis for Λ, so we also conclude that it is a bijection. ∎

We now wish to get a more refined statement. Let $R_n \subset \mathrm{CF}_n$ be the subspace of virtual characters, i.e., integer linear combinations of characters, and set $R = \bigoplus_{n\ge 0} R_n$. Our goal is to determine the irreducible characters.

Proposition 4.3.3. *Suppose* $\varphi_1, \varphi_2, \dots, \varphi_{p(n)} \in R_n$ *form an orthonormal basis with respect to* $(,)_{\Sigma_n}$. *Then the irreducible characters are* $\varepsilon_1\varphi_1, \varepsilon_2\varphi_2, \dots, \varepsilon_{p(n)}\varphi_{p(n)}$ *for some choices* $\varepsilon_i \in \{1,-1\}$.

Proof. By definition, the irreducible characters belong to R_n and they form an orthonormal basis by Theorem 4.1.2. Now write the φ_i as integer linear combinations of the irreducible characters. These coefficients give an orthogonal matrix (with respect to a positive definite form) with integer entries. The only orthonormal vectors with integer entries are standard basis vectors and their negatives, so each row is one of these. Since the matrix is invertible, we see that the matrix has exactly one nonzero entry in each row and column, and that entry is ± 1. ☐

With this in mind, the first step is to find an orthonormal basis of R_n. We know that the Schur functions form an orthonormal basis of Λ, so we define

$$\chi^\lambda = \mathrm{ch}^{-1}(s_\lambda).$$

Our goal now is show that $\chi^\lambda \in R$ and that they in fact are the irreducible characters.

The trivial representation 1_{Σ_n} is the trivial homomorphism $\Sigma_n \to \mathbf{GL}_1(\mathbf{C})$ which sends everything to 1. Its character just assigns 1 to every permutation, so $\mathrm{char}(1_{\Sigma_n}) = \sum_\lambda 1_\lambda$. For every partition $\alpha = (\alpha_1, \dots, \alpha_k)$, define

$$\eta^\alpha = 1_{\alpha_1} \circ \cdots \circ 1_{\alpha_k}.$$

Lemma 4.3.4. $\mathrm{ch}(\eta^\alpha) = h_\alpha$.

Proof. First, $\mathrm{ch}(1_{\Sigma_n}) = \sum_\lambda z_\lambda^{-1} p_\lambda$, which by Theorem 2.5.5, is h_n. Now use that ch is a ring homomorphism. ☐

Corollary 4.3.5. *$\chi^\lambda \in R$ and* ch *restricts to an isomorphism $R \to \Lambda$.*

Proof. We have $\eta^\alpha \in R$ and $s_\lambda = \det(h_{\lambda_i - i + j})_{i,j=1}^n$ by the Jacobi–Trudi identity (Theorem 3.6.1). This expresses the Schur function as an integer linear combination of h_α, so χ^λ is also an integer linear combination of the η^α.

So up to a sign, the χ^λ are the irreducible characters. In particular, every element of R is an integer linear combination of the χ^λ. Since ch sends them to a basis of Λ, we get the second statement. ☐

Finally, we have to determine which of χ^λ and $-\chi^\lambda$ is actually the irreducible character. To do this, we evaluate on the identity element of Σ_n. The trace of the identity element is the dimension of the representation, so we just need to determine if this evaluation is positive or negative. It will turn out that $\chi^\lambda(1) > 0$, and more generally, the Murnaghan–Nakayama rule, to be studied next, will determine the evaluation at any permutation.

4.4. **Murnaghan–Nakayama rule.** Given partitions λ, μ, let $\chi^\lambda(\mu)$ denote the evaluation of χ^λ on any permutation with cycle type μ. Then by definition,

$$\chi^\lambda = \sum_\mu \chi^\lambda(\mu) 1_\mu.$$

Applying the characteristic map, we get

$$s_\lambda = \sum_\mu z_\mu^{-1} \chi^\lambda(\mu) p_\mu.$$

So to determine these evaluations, we need to determine how the Schur functions can be written in terms of the power sum symmetric functions. More generally, given $\nu \subseteq \lambda$, define

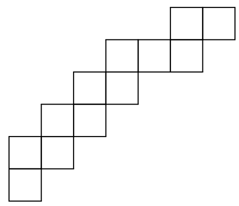
$$\chi^{\lambda/\nu} = \mathrm{ch}^{-1}(s_{\lambda/\nu}),$$

so that

$$s_{\lambda/\nu} = \sum_\mu z_\mu^{-1} \chi^{\lambda/\nu}(\mu) p_\mu. \tag{4.4.1}$$

We will first study the inverse problem of expressing the p's in terms of the Schur functions and get what we want using the scalar product.

Define a **border strip** to be a connected skew diagram with no 2×2 subdiagram. (Sharing only a corner is not considered connected, so $21/1$ is not connected.) Here is an example border strip:



The **height** of a border strip B is denoted $\mathrm{ht}(B)$, and is the number of rows minus 1. In the example above, the height is 5.

Theorem 4.4.2. *Given a positive integer r, we have*

$$s_\mu p_r = \sum_\lambda (-1)^{\mathrm{ht}(\lambda/\mu)} s_\lambda$$

where the sum is over all λ such that $\mu \subseteq \lambda$ and λ/μ is a border strip of size r.

Proof. It suffices to prove this in n variables where $n \gg 0$. Recall the definition of the determinant $a_\alpha = \det(x_i^{\alpha_j})_{i,j=1}^n$ where $\alpha = (\alpha_1, \dots, \alpha_n)$ is any sequence of non-negative integers. Recall $\rho = (n-1, n-2, \dots, 1, 0)$. Let ε_j be the sequence with a single 1 in position j and 0's elsewhere. By Corollary 3.4.3, we have $s_\lambda = a_{\lambda+\rho}/a_\rho$.

We have

$$a_{\mu+\rho} p_r = (\sum_{\sigma \in \Sigma_n} \mathrm{sgn}(\sigma) x^{\sigma(\mu+\rho)})(\sum_{j=1}^n x_j^r) = \sum_{j=1}^n \sum_{\sigma \in \Sigma_n} \mathrm{sgn}(\sigma) x^{\sigma(\mu+\rho+r\varepsilon_j)} = \sum_{j=1}^n a_{\mu+\rho+r\varepsilon_j},$$

where in the second equality, we multiply $x^{\sigma(\mu+\rho)}$ by $x_{\sigma(j)}^r$ to get $x^{\sigma(\mu+\rho+r\varepsilon_j)}$.

Next, $a_\alpha = -a_\beta$ if $\beta = (\alpha_1, \dots, \alpha_i, \alpha_{i-1}, \dots, \alpha_n)$, i.e., β is obtained from α by swapping two consecutive entries. In particular, $a_{\alpha+\rho} = -a_{\gamma+\rho}$ where $\gamma = (\alpha_1, \dots, \alpha_i - 1, \alpha_{i-1} + 1, \dots, \alpha_n)$ and also $a_{\alpha+\rho} = 0$ if α has any repeating entries. We say that γ is obtained by a shifted transposition at position $i-1$.

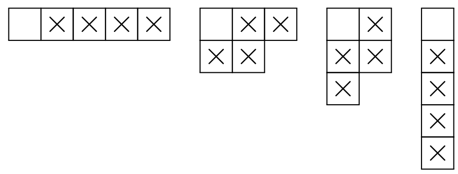
Suppose that $\mu + r\varepsilon_j$ has no repeating entries. Note that $\mu + r\varepsilon_j/\mu$ is a border strip of size r and height 0 ($\mu + r\varepsilon_j$ may not be a partition, but we will draw its diagram in the expected way). If $\mu + r\varepsilon_j$ is not a partition, then replace it by the shifted transposition at position $j-1$. What happens in rows $j-1$ and j is that $(\mu_{j-1}, \mu_j + r)$ gets replaced by $(\mu_j + r - 1, \mu_{j-1} + 1)$. This is a new shape that contains μ and the complement is a border strip with two rows, the first of length $\mu_j + r - 1 - \mu_{j-1}$ and the second of length $\mu_{j-1} + 1 - \mu_j$. If this is a partition, we stop, otherwise we apply another shifted transposition at position $j-2$, and so on.

The end result is a new partition containing μ whose complement is a border strip. The height of this border strip is precisely the number of shifted transpositions we applied, and all border strips arise in this way by taking j to be the row index of the last row of the border strip (we will not go into detail on this). Hence we get the formula

$$\sum_{j=1}^{n} a_{\mu+\rho+r\varepsilon_j} = \sum_{\lambda} (-1)^{\mathrm{ht}(\lambda/\mu)} a_{\lambda+\rho}$$

where the sum is over all λ containing μ such that λ/μ is a border strip of size r and $\ell(\lambda) \le n$. If $n \ge \ell(\mu) + r$, this accounts for all possible border strips. Now divide both sides by a_ρ to get the desired identity. $\square$

Example 4.4.3. $s_1 p_4 = s_5 - s_{3,2} + s_{2,2,1} - s_{1^5}$ corresponding to the following border strips:



In the notation of the proof, these come from $a_{1+\rho+4\varepsilon_j}$ for $j = 1, 2, 3, 5$. $\square$

Given a sequence of non-negative integers $\alpha = (\alpha_1, \dots, \alpha_k)$, a **border-strip tableau** of shape λ/μ is a sequence of partitions $\mu = \lambda^0 \subseteq \lambda^1 \subseteq \cdots \lambda^k = \lambda$ such that λ^i/λ^{i-1} is a border strip of size α_i. The height of this tableau is the sum of the heights of the border strips λ^i/λ^{i-1}.

Corollary 4.4.4.

$$s_\mu p_\alpha = \sum_\lambda \sum_T (-1)^{\mathrm{ht}(T)} s_\lambda$$

where the inner sum is over all border-strip tableaux T of shape λ/μ and type α. In particular,

$$p_\alpha = \sum_\lambda \sum_T (-1)^{\mathrm{ht}(T)} s_\lambda$$

where the inner sum is over all border-strip tableaux T of shape λ and type α.

Corollary 4.4.5. $\chi^{\lambda/\nu}(\mu) = \sum_T (-1)^{\mathrm{ht}(T)}$ *where the sum is over all border-strip tableaux T of shape λ/ν and type μ.*

Proof. By (4.4.1) and orthogonality properties of the p_μ, we get

$$\begin{aligned} \chi^{\lambda/\nu}(\mu) &= \langle s_{\lambda/\nu}, p_\mu \rangle \\ &= \langle s_\lambda, p_\mu s_\nu \rangle && \text{(Theorem 3.5.3)} \\ &= \sum_T (-1)^{\mathrm{ht}(T)} && \text{(Corollary 3.2.10)} \end{aligned}$$

where the sum is the one we want. $\square$

Corollary 4.4.6. *Let $n = |\lambda/\mu|$. Then $\chi^{\lambda/\mu}(1^n) = f^{\lambda/\mu}$, the number of standard Young tableaux of shape λ/μ. In particular, $\chi^\lambda(1^{|\lambda|}) > 0$, so χ^λ is the character of the symmetric group $\Sigma_{|\lambda|}$ of an irreducible representation of dimension f^λ.*

Proof. By definition, a border-strip tableau of type (1^n) is the same as a standard Young tableau, and its height is always 1. The last statement follows from the discussion at the end of the previous section. $\square$

Corollary 4.4.7. *The Littlewood–Richardson coefficient $c^\lambda_{\mu,\nu}$ is non-negative.*

Proof. Recall that $s_\nu s_\mu = \sum_\lambda c^\lambda_{\mu,\nu} s_\lambda$. Applying ch^{-1}, this becomes $\chi^\nu \circ \chi^\mu = \sum_\lambda c^\lambda_{\mu,\nu}\chi^\lambda$. The induction of the character of a representation is again the character of a representation, so the right hand side is the character of a representation. Since every character is a non-negative sum of the the irreducible ones, and the χ^λ are the irreducible characters, we conclude that $c^\lambda_{\mu,\nu} \ge 0$. □

A natural followup: since $c^\lambda_{\mu,\nu}$ is a non-negative integer, is it the cardinality of some combinatorially meaningful set? We will give some constructions of such sets coming from tableaux later.

4.5. **Schur–Weyl duality.** As we've remarked, the Schur functions are characters of irreducible representations of the general linear group, so the characteristic map connects the irreducible characters of symmetric groups with irreducible characters of the general linear group. Schur–Weyl duality offers another connection between the two.

Start with the natural representation $\mathbf{C}^n$ of $\mathbf{GL}_n(\mathbf{C})$. Taking tensor products allows us to build new representations from old. Here we will move away from matrices and use vector spaces. For our purposes, the tensor product of two vector spaces V and W is denoted $V \otimes W$ and consists of linear combinations of symbols $v \otimes w$ where $v \in V$ and $w \in W$ subject to the bilinearity conditions:

- $(v + v') \otimes w = v \otimes w + v' \otimes w$,
- $v \otimes (w + w') = v \otimes w + v \otimes w'$,
- $\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w)$ for $\lambda \in \mathbf{C}$.

If $e_1, \dots, e_n$ is a basis for V and $f_1, \dots, f_m$ is a basis for W, then $e_i \otimes f_j$ with $1 \le i \le n$ and $1 \le j \le m$ is a basis for $V \otimes W$.

So we have a representation $(\mathbf{C}^n)^{\otimes d}$ of $\mathbf{GL}_n(\mathbf{C})$ for all $d \ge 0$ (when $d = 0$, interpret $(\mathbf{C}^n)^{\otimes 0} = \mathbf{C}$ to be the trivial representation): the action is given by

$$g(\sum_i v_{i_1} \otimes \cdots \otimes v_{i_d}) = \sum_i g(v_{i_1}) \otimes \cdots \otimes g(v_{i_d}).$$

This space also has an action of the symmetric group Σ_d by permuting tensors, i.e.,

$$\sigma(\sum_i v_{i_1} \otimes \cdots \otimes v_{i_d}) = \sum_i v_{\sigma(i_1)} \otimes \cdots \otimes v_{\sigma(i_d)}.$$

From these formulas, it's clear that the action of Σ_d commutes with the action of $\mathbf{GL}_n(\mathbf{C})$, i.e., we can permute the tensor factors or do a change of coordinates in either order and get the same result. In particular, we have a well-defined action of $\mathbf{GL}_n(\mathbf{C}) \times \Sigma_d$ on $(\mathbf{C}^n)^{\otimes d}$.

Since both groups are semisimple, i.e., their finite-dimensional representations decompose as direct sums of irreducible ones, we can decompose this space as a sum of irreducible representations of $\mathbf{GL}_n(\mathbf{C}) \times \Sigma_d$. In fact, we get the following:

$$(\mathbf{C}^n)^{\otimes d} = \bigoplus_{\substack{\lambda \vdash d \\ \ell(\lambda) \le n}} \mathbf{S}_\lambda(\mathbf{C}^n) \otimes \mathbf{M}_\lambda,$$

where $\mathbf{S}_\lambda(\mathbf{C}^n)$ is an irreducible representation (Schur module) of $\mathbf{GL}_n(\mathbf{C})$ whose character is the Schur polynomial $s_\lambda(x_1, \dots, x_n)$ and $\mathbf{M}_\lambda$ is an irreducible representation (Specht module) of Σ_d whose character is χ^λ.

The same will happen if we take $n = \infty$, so we can get Schur functions instead of Schur polynomials. In Remark 3.5.10, we said that the multiplicity of $\mathbf{S}_\lambda(\mathbf{C}^n)$ in $(\mathbf{C}^n)^{\otimes d}$ is f^λ by the Pieri rule. Here we see that f^λ is the dimension of $\mathbf{M}_\lambda$ and the multiplicity space has the structure of a Σ_d-representation.

Schur–Weyl duality can be used to connect a number of coefficients defined for the symmetric group and those defined for the general linear group. One such are the Kronecker coefficients: these describe the multiplicity of χ^λ in the tensor product $\chi^\mu\chi^\nu$ (this is the usual product of characters, not the induction product). We may return to this later.

5. Schubert calculus

We have seen symmetric functions as they arise in the representation theory of general linear groups and symmetric groups. There are several more incarnations of symmetric functions, and we now explore one in algebraic geometry. For a more detailed exposition, see [Fu1] or [Ma].

5.1. **Grassmannians.**

Given an n-dimensional vector space V and a positive integer r, the Grassmannian is denoted $\mathbf{Gr}_r(V)$ and is the set of r-dimensional subspaces of V. We will primarily work with complex vector spaces for simplicity, though some nice combinatorics also comes out of considering vector spaces over finite fields.

Consider the space $\mathrm{Hom}(\mathbf{C}^r, V)$ of linear maps $\mathbf{C}^r \to V$. These can be encoded as matrices if we choose a basis for V. This space naturally inherits a topology from the fact that it is just $\mathbf{C}^{rn}$. Depending on your preferences, this is either the standard Euclidean topology, or if you are more algebraically inclined, the Zariski topology (everything we say will work for either). In either topology, the subset $\mathrm{Hom}^\circ(\mathbf{C}^r, V)$ of injective maps is an open set (and thus also inherits a topology). To see that it is open, note that a matrix is not injective if and only if all of its maximal size square submatrices have determinant 0. So the space of matrices which are not injective is the preimage of 0 under a map

$$\mathrm{Hom}(\mathbf{C}^r, V) \to \mathbf{C}^{\binom{n}{r}}$$

which sends a matrix to all such determinants, and hence is closed (these determinants are just polynomials and hence are continuous).

Now, we can think of $\mathbf{GL}_r(\mathbf{C})$ as acting on $\mathrm{Hom}^\circ(\mathbf{C}^r, V)$ by column operations. Since the linear maps are injective, this action is actually free, i.e., no map has a nontrivial stabilizer. Any two maps in the same $\mathbf{GL}_r(\mathbf{C})$ orbit have the same image in V, so we have an identification

$$\mathbf{Gr}_r(V) = \mathrm{Hom}^\circ(\mathbf{C}^r, V)/\mathbf{GL}_r(\mathbf{C}).$$

The right hand side gets the quotient topology, so this can be used to turn $\mathbf{Gr}_r(V)$ into a topological space.

Our object of study is the cohomology ring[6] of the Grassmannian, i.e., $\mathrm{H}^*(\mathbf{Gr}_r(V))$. We won't need the general definition of cohomology here, but will define it in an ad hoc way for the Grassmannian using the Schubert decomposition.

[6] If you prefer the Zariski topology, then you need to use the Chow ring.

5.2. **Schubert cells.** The general linear group $\mathbf{GL}(V)$ acts on $\mathbf{Gr}_r(V)$: given a subspace of dimension r, applying g to all of its vectors gives another subspace of dimension r. We now want to pick an ordered basis $e_1, \dots, e_n$ for V, so we can identify it with $\mathbf{C}^n$. Let $B \subset \mathbf{GL}_n(\mathbf{C})$ be the subgroup of upper-triangular matrices. We are interested in the action of B on $\mathbf{Gr}_r(\mathbf{C}^n)$.

Go back to the identification $\mathbf{Gr}_r(\mathbf{C}^n) = \mathrm{Hom}^\circ(\mathbf{C}^r, \mathbf{C}^n)/\mathbf{GL}_r(\mathbf{C})$. From this, we want to find a "standard" way to represent each subspace as an $n \times r$ matrix. Note that we are identifying two matrices if they differ by column operations. Start with such a matrix φ. Since φ is injective, some row must be nonzero, let i_r be the largest index so that the i_rth row is nonzero. Up to column operations, we can turn this row into the standard length r row vector $(0, 0, \dots, 0, 1)$. Next, let i_{r-1} be the largest index so that the i_{r-1}th row is not in the span of $(0, 0, \dots, 0, 1)$. Using column operations on just the first $r-1$ columns, we can turn this row vector into $(0, 0, \dots, 0, 1, *)$ and then using another column operation using just the last two columns, turn it into $(0, 0, \dots, 0, 1, 0)$ (the i_rth row remains the same). Now repeat: in general, we let i_k be the largest index so that the i_kth row is not in the span of the rows after it and turn it into $(0, 0, \dots, 1, 0, \dots, 0)$ where the 1 is in the kth entry.

Ultimately, there will be unique indices $i_1 < i_2 < \cdots < i_r$ so that row i_j is the jth standard basis vector and any row vector in between positions i_j and i_{j-1} (convention: $i_0 = 0$) has 0's in positions $< j$. Also, it is clear that this is a unique way to parametrize injective matrices up to column operations.

Example 5.2.1. If $r = 2$ and $n = 4$, all matrices are one of the following form up to column operations:

$$\begin{pmatrix}1&0\\0&1\\0&0\\0&0\end{pmatrix} \quad \begin{pmatrix}1&0\\0&*\\0&1\\0&0\end{pmatrix} \quad \begin{pmatrix}1&0\\0&*\\0&*\\0&1\end{pmatrix} \quad \begin{pmatrix}*&*\\1&0\\0&1\\0&0\end{pmatrix} \quad \begin{pmatrix}*&*\\1&0\\0&*\\0&1\end{pmatrix} \quad \begin{pmatrix}*&*\\*&*\\1&0\\0&1\end{pmatrix}$$

The $*$ are shorthand for an arbitrary complex number. ☐

Given a collection of indices $\mathbf{i} = (i_1 < i_2 < \cdots < i_r)$, we let $X_{\mathbf{i}}^\circ$ denote the set of subspaces whose matrix representatives have indices $i_1 < \cdots < i_r$ as above. From the above discussion, we see that the space $X_{\mathbf{i}}^\circ$ is naturally isomorphic to $\mathbf{C}^{d(\mathbf{i})}$ where $d(\mathbf{i}) = \sum_{j=1}^r (i_j - j)$.

There is an alternative indexing which is convenient, although appears unnatural at first. Given $\mathbf{i}$, let $\lambda_j = n - i_j - r + j$. Then $n - r \geq \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r \geq 0$, so λ is a partition sitting inside the $r \times (n-r)$ rectangular partition, and in this case $X_\lambda^\circ \cong \mathbf{C}^{r(n-r)-|\lambda|}$.

We summarize what we've discussed.

Proposition 5.2.2. *The Grassmannian has a decomposition*

$$\mathbf{Gr}_r(\mathbf{C}^n) = \amalg_{\lambda \subseteq r \times (n-r)} X_\lambda^\circ$$

where $X_\lambda^\circ \cong \mathbf{C}^{r(n-r)-|\lambda|}$. *Furthermore, each* X_λ° *is an orbit under the action of* B.

The largest cell has dimension $r(n-r)$, so we can justify saying $\dim \mathbf{Gr}_r(\mathbf{C}^n) = r(n-r)$. Hence we see that X_λ° has *codimension* $|\lambda|$, which is advantageous when discussing cohomology.

This decomposition is known as the **Schubert decomposition** of the Grassmannian.

Remark 5.2.3. The Schubert decomposition makes sense for an arbitrary field. If we use a finite field with q elements, then we get

$$|\mathbf{Gr}_r(\mathbf{F}_q^n)| = \sum_{\lambda \subseteq r \times (n-r)} q^{r(n-r)-|\lambda|} = \sum_{\lambda \subseteq r \times (n-r)} q^{|\lambda|}.$$

The second equality follows from the involution which sends a partition in a rectangle to its complementary shape (which is also a partition). The last sum can also be interpreted as the q-binomial coefficient $\begin{bmatrix} n \\ r \end{bmatrix}_q$. □

Proposition 5.2.4. *The closure of X_λ° is the union of all X_μ° such that $\mu \supseteq \lambda$.*

Example 5.2.5. Before proving this, let's give an example. Set $r = 2$ and $n = 4$ and take $\mu = (1, 0)$ and $\lambda = (0, 0)$. Take the subspace in $X_{1,0}^\circ$ represented by the matrix

$$\begin{pmatrix} a & b \\ 1 & 0 \\ 0 & c \\ 0 & 1 \end{pmatrix}.$$

Now consider the family of subspaces (indexed by ε)

$$\begin{pmatrix} a & b \\ 1 & 0 \\ \varepsilon & c \\ 0 & 1 \end{pmatrix}.$$

If $\varepsilon \neq 0$, then this is equivalent, up to column operations, to the matrix

$$\begin{pmatrix} a/\varepsilon & b - ac/\varepsilon \\ 1/\varepsilon & -c/\varepsilon \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

and so belongs to $X_{0,0}^\circ$. If we take the limit $\varepsilon \to 0$, then this becomes the matrix we started with, so it's in the closure. □

Proof. The idea of the proof is illustrated by the example. It suffices to show that any matrix in X_μ° is in the closure when $|\mu| = |\lambda| + 1$. Then say μ differs from λ in the jth entry. Take a matrix representing a subspace in X_μ°. Then the entry in the jth column and the $(i_j + 1)$th row is 0, replace it with ε. When $\varepsilon \neq 0$, this is a matrix representing a subspace in X_λ°; now take the limit $\varepsilon \to 0$. □

The closure of X_λ° is denoted X_λ and is a **Schubert variety**.

However, note that everything we did depended on a choice of basis. In fact, it depends on a little less than that and we want to make this precise now. A **complete flag** in V is a choice of subspaces $F_1 \subset F_2 \subset \cdots \subset F_{n-1} \subset V$ such that $\dim F_i = i$ for all i. In our discussion above, F_i is the span of $e_1, \dots, e_i$.

Recall from above that

$$i_j = n - \lambda_j - r + j.$$

Lemma 5.2.6. *We have*

$$X_\lambda = \{W \in \mathbf{Gr}_r(\mathbf{C}^n) \mid \dim(W \cap \operatorname{span}(e_1, \dots, e_{i_j})) \geq j \text{ for } j = 1, \dots, r\}.$$

Proof. Given a subspace W, the first j columns of the standard form for its representing matrix are in the span of $e_1, \dots, e_{i_j}$ if and only if $\dim(W \cap \mathrm{span}(e_1, \dots, e_{i_j})) \ge j$. The first condition is equivalent to $W \in X^\circ_\mu$ where $\mu_{r-j+1} \ge \lambda_{r-j+1}$. Taking all j at once, this is equivalent to $\mu \supseteq \lambda$. □

Motivated by this, given any flag $F_\bullet$, we define

$$X(F_\bullet)_\lambda = \{W \in \mathbf{Gr}_r(\mathbf{C}^n) \mid \dim(W \cap F_{i_j}) \ge j \text{ for } j = 1, \dots, r\}.$$

Lemma 5.2.7. *If $\lambda = (k)$ has a single nonzero entry, then*

$$X(F_\bullet)_k = \{W \in \mathbf{Gr}_r(\mathbf{C}^n) \mid W \cap F_{n-r-k+1} \neq 0\}.$$

Proof. From what we've shown, $X(F_\bullet)_k$ is the set of W such that $\dim(W \cap F_{n-r-k+1}) \ge 1$ and $\dim(W \cap F_{n-r+j}) \ge j$ for $j \ge 2$. The first condition is equivalent to $W \cap F_{n-r-k+1} \neq 0$. The latter conditions are automatic: recall that given two subspaces $E, F \subset \mathbf{C}^n$, we have

$$\dim(E) + \dim(F) = \dim(E + F) + \dim(E \cap F).$$

Take $E = W$ and $F = F_{n-r+j}$ and use the fact that $\dim(E + F) \le n$ always to get

$$\dim(W \cap F_{n-r+j}) \ge r + (n - r + j) - n = j.$$ □

5.3. **Cohomology ring.** We will not assume knowledge of algebraic topology, so we'll just state the facts about the cohomology ring that we'll be using. In particular, for each Grassmannian $\mathbf{Gr}_r(\mathbf{C}^n)$, we have a graded ring $\mathrm{H}^*(\mathbf{Gr}_r(\mathbf{C}^n))$, and it has the following important properties:

(1) Each Schubert variety $X(F_\bullet)_\lambda$ corresponds to an element in $\mathrm{H}^{2|\lambda|}(\mathbf{Gr}_r(\mathbf{C}^n))$ and this element does not depend on $F_\bullet$. We just call it σ_λ.

(2) $\{\sigma_\lambda \mid \lambda \subseteq r \times (n-r)\}$ is an integral basis for $\mathrm{H}^*(\mathbf{Gr}_r(\mathbf{C}^n))$. So as an additive group $\mathrm{H}^*(\mathbf{Gr}_r(\mathbf{C}^n))$ is free abelian of rank $\binom{n}{r}$ and $\mathrm{H}^i = 0$ if i is odd.

   Furthermore, $\mathrm{H}^{2r(n-r)}(\mathbf{Gr}_r(\mathbf{C}^n)) \cong \mathbf{Z}$ with distinguished basis element $\sigma_{r\times(n-r)}$. Given $f \in \mathrm{H}^*(\mathbf{Gr}_r(\mathbf{C}^n))$, we will let $\int f$ be the coefficient of $\sigma_{r\times(n-r)}$.

(3) The multiplication is commutative.

(4) If $|\lambda| + |\mu| = r(n-r)$, and $X(F_\bullet)_\lambda \cap X(G_\bullet)_\mu$ consists of a finite collection of c points (counted with multiplicity – which won't be an issue in our computations), then $\int \sigma_\lambda\sigma_\mu = c$ for some choice of flags $F_\bullet, G_\bullet$.

Given a partition $\lambda \subseteq r \times (n-r)$, let $\lambda^\vee$ be the complementary partition, i.e., $\lambda^\vee_i = n - r - \lambda_{r+1-i}$. Visually, it just means that we can rotate the Young diagram of $\lambda^\vee$ 180 degrees to get $(r \times (n-r))/\lambda$.

Throughout, we'll use the standard flag $E_i = \mathrm{span}(e_1, \dots, e_i)$ (the span of the first i basis vectors) as well as the dual flag $E^\vee_i = \mathrm{span}(e_{n+1-i}, \dots, e_{n-1}, e_n)$ (the span of the last i basis vectors).

Proposition 5.3.1. *If $X(E_\bullet)_\lambda \cap X(E^\vee_\bullet)_\mu \neq \varnothing$, then $\mu_{r+1-j} \le n - r - \lambda_j$ for $j = 1, \dots, r$.*

If $|\lambda| + |\mu| = r(n-r)$, then

$$\int \sigma_\lambda\sigma_\mu = \delta_{\mu,\lambda^\vee}.$$

Proof. Consider the intersection $X(E_\bullet)_\lambda \cap X(E^\vee_\bullet)_\mu$. If it is nonempty, let W be a subspace in both sets. By our description, for all $1 \le j \le r$, we have

$$\dim(W \cap E_{n-\lambda_j-r+j}) \ge j, \qquad \dim(W \cap E^\vee_{n-\mu_{r+1-j}+1-j}) \ge r + 1 - j.$$

Set $F = E_{n-\lambda_j-r+j}$ and $F' = E^\vee_{n-\mu_{r+1-j}+1-j}$. Since $\dim((W\cap F)+(W\cap F')) \le \dim W \le r$, we have

$$\dim(W\cap F\cap F') = \dim(W\cap F)+\dim(W\cap F')-\dim((W\cap F)+(W\cap F')) \ge j+(r+1-j)-r = 1.$$

In particular, $F\cap F' \ne 0$. However, F is the span of the first $n-\lambda_j-r+j$ basis vectors while F' is the span of the last $n-\mu_{r+1-j}+1-j$ basis vectors, so this implies that $\mu_{r+1-j}+j \le n-\lambda_j-r+j$, or just $\mu_{r+1-j} \le n-r-\lambda_j$.

Hence if $|\mu|+|\lambda| = r(n-r)$ and $\mu \ne \lambda^\vee$, then $\int \sigma_\lambda\sigma_\mu = 0$. Furthermore, if $\mu = \lambda^\vee$, the discussion above implies that W contains the basis vector $e_{n-\lambda_j-r+j}$ for all j, which means $W = \mathrm{span}\{e_{n-\lambda_j-r+j} \mid j = 1,\dots,r\}$, so $X(E_\bullet)_\lambda \cap X(E^\vee_\bullet)_{\lambda^\vee}$ consists of one point (in fact, it has multiplicity one, but we didn't define that, so we omit the verification). ☐

Note that since the σ_λ form a basis, the product of two Schubert classes is a sum of other Schubert classes. The above result tells us that the coefficient of σ_ν in $\sigma_\lambda\sigma_\mu$ is given by $\int \sigma_\lambda\sigma_\mu\sigma_{\nu^\vee}$.

Theorem 5.3.2 (Pieri formula)**.** *For each integer $0 \le k \le n-r$, we have*

$$\sigma_\lambda\sigma_k = \sum_\mu \sigma_\mu$$

where the sum is over all $\mu \supseteq \lambda$ such that μ/λ is a horizontal strip of size k and $\mu \subseteq r\times(n-r)$.

Proof. We need to calculate $\int \sigma_\lambda\sigma_k\sigma_{\mu^\vee}$. To do that, we will count the number of points in the intersection of $X(E_\bullet)_\lambda \cap X(E^\vee_\bullet)_{\mu^\vee} \cap X(L_\bullet)_k$ for some flag $L_\bullet$ to be chosen later. We may assume that $\lambda_i \le \mu_i$ for all i; if not, then $X(E_\bullet)_\lambda \cap X(E^\vee_\bullet)_{\mu^\vee} = \varnothing$ and so $\int \sigma_\lambda\sigma_k\sigma_{\mu^\vee} = 0$. We may also assume that $|\mu| = |\lambda|+k$, otherwise the integral is 0 for degree reasons.

For $i = 1,\dots,r$, define

$$\begin{aligned}
A_i &= E_{n-r+i-\lambda_i} = \mathrm{span}(e_1, e_2, \dots, e_{n-r+i-\lambda_i}),\\
B_i &= E^\vee_{n-r+i-\mu^\vee_i} = \mathrm{span}(e_{r-i+\mu^\vee_i+1}, \dots, e_n),\\
C_i &= A_i \cap B_{r+1-i} = \mathrm{span}(e_{n-r+i-\mu_i}, \dots, e_{n-r+i-\lambda_i}).
\end{aligned}$$

By convention, set $\lambda_0 = n-r = \mu^\vee_0$ so that $A_0 = B_0 = 0$. By our assumption, $C_i \ne 0$ for all i. Set $C = C_1 + C_2 + \cdots + C_r$.

Claim 1. $C = \bigcap_{i=0}^r (A_i + B_{r-i})$.

First, $B_{r-i} = \mathrm{span}(e_{n-r-\mu_{i+1}+i+1}, \dots, e_n)$. We first show $C \subseteq \bigcap_{i=0}^r (A_i + B_{r-i})$. Pick $e_p \in C_j$ so that $n-r+j-\mu_j \le p \le n-r+j-\lambda_j$. If $i < j$, then

$$n-r-\mu_{i+1}+i+1 \le n-r-\mu_j+j \le p,$$

so $e_p \in B_{r-i}$. If $i \ge j$, then

$$n-r+i-\lambda_i \ge n-r+j-\lambda_j \le p,$$

so $e_p \in A_i$. In either case, $e_p \in A_i + B_{r-i}$, so $C \subseteq \bigcap_{i=0}^r (A_i + B_{r-i})$.

For the other inclusion, suppose $e_p \in \bigcap_{i=0}^r (A_i + B_{r-i})$. Pick j minimal so that $p \le n-r+j-\lambda_j$. In particular, $e_p \notin A_{j-1}$, so we must have $e_p \in B_{r-j+1}$. But also $e_p \in A_j$, so $e_p \in C_j$ and hence $e_p \in C$.

Claim 2. If $W \in X(E_\bullet)_\lambda \cap X(E^\vee_\bullet)_{\mu^\vee}$, then $W \subseteq C$.

First start with $\dim(W \cap A_i) \ge i$ and $\dim(W \cap B_{r-i}) \ge r - i$. If $A_i \cap B_{r-i} \neq 0$, then $A_i + B_{r-i} = \mathbf{C}^n$ since it contains all basis vectors, and hence $W \subset A_i + B_{r-i}$ trivially. Otherwise, if $A_i \cap B_{r-i} = 0$, we get that $\dim(W \cap (A_i + B_{r-i})) \ge i + (r-i) = r$, so $W \subseteq A_i + B_{i+1}$ again. In particular, $W \subseteq C$ by Claim 1.

Claim 3. The sum $C = C_1 + \cdots + C_r$ is a direct sum if and only if $\dim C = k + r$ if and only if $\lambda \subseteq \mu$ and μ/λ is a horizontal strip.

First, $\dim C_i = \mu_i - \lambda_i + 1$, so $\sum_i \dim C_i = |\mu| - |\lambda| + r = k + r$. Note that $\dim C = \sum_i \dim C_i$ if and only if $C_i \cap C_j = 0$ for all i, j, which is equivalent to

$$\mu_1 \ge \lambda_1 \ge \mu_2 \ge \lambda_2 \ge \cdots \ge \mu_r \ge \lambda_r.$$

This last set of inequalities is equivalent to saying that $\lambda \subseteq \mu$ and that μ/λ is a horizontal strip. This proves the claim.

Now we finish the proof. Suppose that μ/λ is not a horizontal strip. By Claim 3, $\dim C \le k + r - 1$. Let L be a subspace of dimension $n - k - r + 1$ such that $L \cap C = 0$ and let $L_\bullet$ be any flag such that $L_{n-k-r+1} = L$. Then $X(L_\bullet)_k = \{W \mid W \cap L \neq 0\}$, and so by Claim 2, $X(E_\bullet)_\lambda \cap X(E_\bullet^\vee)_{\mu^\vee} \cap X(L_\bullet)_k = \varnothing$. This implies $\int \sigma_\lambda \sigma_k \sigma_{\mu^\vee} = 0$.

Finally, suppose that μ/λ is a horizontal strip. Again by Claim 3, $\dim C = k + r$, and we choose L to be any subspace of dimension $n - k - r + 1$ such that $\dim(L \cap C) = 1$ (the minimal possible dimension) and such that a nonzero vector in $L \cap C$ has nonzero projection in each C_i. Suppose that $W \in X(E_\bullet)_\lambda \cap X(E_\bullet^\vee)_{\mu^\vee}$. Then

$$\begin{aligned} \dim(W \cap C_i) &= \dim(W \cap A_i \cap B_{r+1-i}) \\ &= \dim(W \cap A_i) + \dim(W \cap B_{r+1-i}) - \dim((W \cap A_i) + (W \cap B_{r+1-i})) \\ &\ge i + (r + 1 - i) - r = 1. \end{aligned}$$

Since C is a direct sum of the C_i, we must have $\dim(W \cap C_i) = 1$ for all i, let c_i be any nonzero vector in $W \cap C_i$. If $W \in X(L_\bullet)_k$ as well, then from Claim 2, $\dim(L \cap W) = 1$; let c span $L \cap W$. From above, the projection of c to C_i must be some nonzero multiple of c_i. In particular, we have shown that W is uniquely determined by c, it is the span of the projections of c to the various C_i. Hence $X(E_\bullet)_\lambda \cap X(E_\bullet^\vee)_{\mu^\vee} \cap X(L_\bullet)_k$ consists of 1 point (again, we omit verifying that the multiplicity is 1). $\square$

Define a linear map $\psi \colon \Lambda \to \mathrm{H}^*(\mathbf{Gr}_r(\mathbf{C}^n))$ by $\psi(s_\lambda) = \sigma_\lambda$ if $\lambda \subseteq r \times (n-r)$ and $\psi(s_\lambda) = 0$ otherwise.

Corollary 5.3.3. *ψ is a ring homomorphism.*

Proof. Since the Pieri rule holds in both Λ and $\mathrm{H}^*(\mathbf{Gr}_r(\mathbf{C}^n))$, we have $\psi(s_\lambda h_\nu) = \psi(s_\lambda)\psi(h_\nu)$ for any partitions λ, ν. The result now follows since both s_λ and h_ν are bases for Λ. $\square$

Corollary 5.3.4. *$\sigma_\lambda \sigma_\mu = \sum_\nu c^\nu_{\lambda,\mu} \sigma_\nu$ where the sum is over all $\nu \subseteq r \times (n-r)$ and $c^\nu_{\lambda,\mu}$ is the Littlewood–Richardson coefficient.*

Remark 5.3.5. This gives another proof that $c^\nu_{\lambda,\mu} \ge 0$: it is the number of intersection points in $X(E_\bullet)_\lambda \cap X(E_\bullet^\vee)_\mu \cap X(L_\bullet)_{\nu^\vee}$ for some generic choice of flag $L_\bullet$.

We have already interpreted $c^\nu_{\lambda,\mu}$ as the multiplicity of $\mathbf{S}_\nu(\mathbf{C}^n)$ in $\mathbf{S}_\lambda(\mathbf{C}^n) \otimes \mathbf{S}_\mu(\mathbf{C}^n)$ (where n is larger than the number of rows of λ, μ, ν). This can be reinterpreted as the dimension of $\mathrm{Hom}_{\mathbf{GL}_n(\mathbf{C})}(\mathbf{S}_\nu(\mathbf{C}^n), \mathbf{S}_\lambda(\mathbf{C}^n) \otimes \mathbf{S}_\mu(\mathbf{C}^n))$. See [MTV] for a connection between the two. $\square$

Corollary 5.3.6 (Giambelli's formula)**.**

$$\sigma_\lambda = \det(\sigma_{\lambda_i+j-i})_{i,j=1}^r$$

While this is a direct consequence of the Jacobi–Trudi identity, this identity was independently discovered in the context of Grassmannians (hence a different name).

5.4. **Chern classes.** Much of what we say below will apply to spaces X other than $\mathbf{Gr}(r, \mathbf{C}^n)$, but we won't use them, so we keep the discussion just to Grassmannians.

Consider the direct product $\mathbf{Gr}(r, \mathbf{C}^n) \times \mathbf{C}^n$ and the subspace

$$\mathcal{R} = \{(W, v) \mid v \in W\}.$$

This is called the **tautological subbundle** and it has a map $\mathcal{R} \to \mathbf{Gr}(r, \mathbf{C}^n)$ such that the preimage over W is just the space W. This is an example of a vector bundle of rank r (r being the dimension of the vector spaces), but we won't need the precise definition. For those unfamiliar with vector bundles, it suffices to know that one can do many of the things one does with vector spaces: we can take duals, take direct sums, and apply Schur functors. We will be interested in taking symmetric powers for our applications.

Associated to a rank N vector bundle E are **Chern classes** $c_k(E)$ which are elements in $\mathrm{H}^{2k}(\mathbf{Gr}_r(\mathbf{C}^n))$ for $k = 1, \dots, N$. We won't define them precisely, we just discuss how to calculate with them to get interesting applications.

Proposition 5.4.1. $c_k(\mathcal{R}^*) = \sigma_{1^k}$.

A **section** of a vector bundle $\pi \colon E \to \mathbf{Gr}_r(\mathbf{C}^n)$ is a function[7] $s \colon \mathbf{Gr}_r(\mathbf{C}^n) \to E$ such that $\pi \circ s$ is the identity. The **zero locus** of s is denoted $Z(s)$ and is the set of $W \in \mathbf{Gr}_r(\mathbf{C}^n)$ such that $s(W) = 0$. The *expected dimension* of $Z(s)$ is $\dim \mathbf{Gr}_r(\mathbf{C}^n) - \operatorname{rank} E$.

Proposition 5.4.2. *If* $\operatorname{rank} E = \dim \mathbf{Gr}_r(\mathbf{C}^n)$ *and* $Z(s)$ *is a finite collection of points, then the number of points (counted with multiplicity) is* $\int c_N(E)$ *where* $N = \operatorname{rank} E$.

A very useful device are **Chern roots**: for any vector bundle E of rank N, there exists a larger ring R containing $\mathrm{H}^*(\mathbf{Gr}_r(\mathbf{C}^n))$ and elements $x_1, \dots, x_N \in R$ such that $\sigma_k(E)$ is $e_k(x_1, \dots, x_N)$ the kth elementary symmetric function in the x's.

Proposition 5.4.3. *The Chern roots of* $\mathrm{Sym}^d E$ *are* $\{x_{i_1} + \cdots + x_{i_d} \mid i_1 \le i_2 \le \cdots \le i_d\}$.

The Chern roots of $\bigwedge^d E$ *are* $\{x_{i_1} + \cdots + x_{i_d} \mid i_1 < i_2 < \cdots < i_d\}$.

5.5. **Enumerative geometry.** We now illustrate a few types of calculations that can be done with our understanding of the cohomology ring of the Grassmannian.

In our context, projective space $\mathbf{P}^n$ is $\mathbf{Gr}(1, \mathbf{C}^{n+1})$. Every 2-dimensional subspace of $\mathbf{C}^{n+1}$ can be interpreted as a projective line in $\mathbf{P}^n$ by taking the set of 1-dimensional subspaces lying in it.

The following is the prototypical example of a Schubert calculus problem.

Example 5.5.1. Pick 4 sufficiently general lines $\ell_1, \dots, \ell_4$ in projective space $\mathbf{P}^3$. How many other lines ℓ satisfy $\ell \cap \ell_i \neq \varnothing$?

[7]What kind of function? Depends on the context. In topology, it would be a continuous function, in our context it will be "algebraic".

As mentioned above, each ℓ_i is the set of lines inside of some 2-dimensional space E_i in $\mathbf{C}^4$. A line ℓ as above is the same as a 2-dimensional space W such that $\dim(W \cap E_i) \geq 1$. The set

$$Z_i = \{W \in \mathbf{Gr}(2, \mathbf{C}^4) \mid \dim(W \cap E_i) \geq 1\}$$

is a Schubert variety of the form X_1. Hence we want to calculate σ_1^4. Using the Pieri rule, we see that $\sigma_1^4 = 2\sigma_{2,2}$, so the answer is 2.

Technically, this just says that $Z_1 \cap Z_2 \cap Z_3 \cap Z_4$ is 0-dimensional and has 2 points when counted with multiplicity, so it could just be a single double point. But if the ℓ_i are "generic enough" it will in fact be two different points. □

Another classical problem is to count lines on the cubic surface.

Example 5.5.2. A cubic surface is the solution set in $\mathbf{P}^3$ of a homogeneous cubic polynomial f in 4 variables. How many lines are contained in a sufficiently generic cubic surface? This can be rephrased as follows: a polynomial as above can be interpreted as a cubic function $f\colon \mathbf{C}^4 \to \mathbf{C}$. A line corresponds to a 2-dimensional subspace $W \subset \mathbf{C}^4$, and it is contained in the cubic surface if f is identically 0 when restricted to W.

The function f can be interpreted as cubic function on each subspace, and hence we can think of it as a section $f\colon \mathbf{Gr}(2, \mathbf{C}^4) \to \mathrm{Sym}^3(\mathcal{R}^*)$. Note that $\operatorname{rank} \mathrm{Sym}^3(\mathcal{R}^*) = 4 = \dim \mathbf{Gr}(2, \mathbf{C}^4)$. So the expected dimension of $Z(f)$ is 0. Note that points of $Z(f)$ can be identified with the lines inside of the cubic surface defined by f. Assuming there are finitely many, we can calculate the number of points (counted with multiplicity) of $Z(f)$ as $\int c_4(\mathrm{Sym}^3(\mathcal{R}^*))$ by Proposition 5.4.2.

To do this, let's use Chern roots. Let x, y be the Chern roots of $\mathcal{R}^*$. By Proposition 5.4.1, we know that $e_1(x, y) = c_1(\mathcal{R}^*) = \sigma_1$ and $e_2(x, y) = c_2(\mathcal{R}^*) = \sigma_{1,1}$. So in general, $s_\lambda(x, y) = \sigma_\lambda$. By Proposition 5.4.3, the Chern roots of $\mathrm{Sym}^3(\mathcal{R}^*)$ are $3x, 2x + y, x + 2y, 3y$, and $c_4(\mathrm{Sym}^3(\mathcal{R}^*))$ is their product. Let's simplify:

$$\begin{aligned} c_4(\mathrm{Sym}^3(\mathcal{R}^*)) &= (3x)(3y)(2x + y)(x + 2y) \\ &= 9\sigma_{1,1}(2x^2 + 5xy + 2y^2) \\ &= 9\sigma_{1,1}(2\sigma_2 + 3\sigma_{1,1}) \\ &= 27\sigma_{2,2}. \end{aligned}$$

So $\int c_4(\mathrm{Sym}^3(\mathcal{R}^*)) = 27$, and we see that the number of lines, if finite, is at most 27 since we have to take into account "multiplicity". For "sufficiently generic" cubics, the number is exactly 27. □

6. Combinatorial formulas

6.1. **Standard Young tableaux.** Given a partition $\lambda = (\lambda_1, \dots, \lambda_k)$ of n, recall that f^λ is the number of standard Young tableaux of shape λ. We gave a representation-theoretic interpretation in terms of symmetric groups: $f^\lambda = \chi^\lambda(1^n)$ is the dimension of an irreducible representation. Our goal now is to give some formulas for f^λ.

Theorem 6.1.1. *Pick $k \geq \ell(\lambda)$ and define $\ell_i = \lambda_i + k - i$. Then*

$$f^\lambda = \frac{n!}{\ell_1! \cdots \ell_k!} \prod_{1 \leq i < j \leq k} (\ell_i - \ell_j).$$

Proof. From §4.4, we have

$$p_\mu = \sum_\lambda \chi^\lambda(\mu) s_\lambda.$$

Now work in k variables $x_1, \dots, x_k$. Recall from §3.4 that $s_\lambda = a_{\lambda+\rho}/a_\rho$. Multiply both sides above by a_ρ to get

$$a_\rho p_\mu = \sum_\lambda \chi^\lambda(\mu) a_{\lambda+\rho}.$$

If λ, λ' are partitions, the coefficient of $x^{\lambda+\rho}$ in $a_{\lambda'+\rho}$ is $\delta_{\lambda,\lambda'}$. In particular, we conclude that $\chi^\lambda(\mu)$ is the coefficient of $x^{\lambda+\rho}$ in $a_\rho p_\mu$. We're interested in the case $\mu = 1^n$.

First, by definition, we have

$$a_\rho = \det(x_i^{k-j})_{i,j=1}^k = \sum_{\sigma \in \Sigma_k} \operatorname{sgn}(\sigma) x_1^{k-\sigma(1)} \cdots x_k^{k-\sigma(k)}$$

and

$$p_{1^n} = (x_1 + x_2 + \cdots + x_k)^n = \sum_{i_1,\dots,i_k} \binom{n}{i_1, \dots, i_k} x_1^{i_1} \cdots x_k^{i_k}$$

where the sum is over all integers $i_1, \dots, i_k \geq 0$ such that $i_1 + \cdots + i_k = n$ and

$$\binom{n}{i_1, \dots, i_k} = \frac{n!}{i_1! \cdots i_k!}$$

is the multinomial coefficient. Hence the coefficient of $x^{\lambda+\rho}$ in $a_\rho p_{1^n}$ is

$$\sum_{\sigma \in \Sigma_k} \operatorname{sgn}(\sigma) \binom{n}{\ell_1 - k + \sigma(1), \dots, \ell_k - k + \sigma(k)} = \sum_{\sigma \in \Sigma_k} \operatorname{sgn}(\sigma) \frac{n!}{(\ell_1 - k + \sigma(1))! \cdots (\lambda_k - k + \sigma(k))!}$$

where, by convention, the sum is over σ such that the binomial coefficients make sense (i.e., $\ell_i + \sigma(i) \geq k$ for all i). Define $(x)_r = x(x-1)\cdots(x-r+1)$ so that $(\ell_i - k + \sigma(i))! = \ell_i!/(\ell_i)_{k-\sigma(i)}$. Then we can further rewrite this as

$$\frac{n!}{\ell_1! \cdots \ell_k!} \sum_{\sigma \in \Sigma_k} \operatorname{sgn}(\sigma) \prod_{i=1}^k (\ell_i)_{k-\sigma(i)} = \frac{n!}{\ell_1! \cdots \ell_k!} \det \begin{pmatrix} (\ell_1)_{k-1} & (\ell_1)_{k-2} & \cdots & (\ell_1)_2 & \ell_1 & 1 \\ (\ell_2)_{k-1} & (\ell_2)_{k-2} & \cdots & (\ell_2)_2 & \ell_2 & 1 \\ \vdots & & & & & \vdots \\ (\ell_k)_{k-1} & (\ell_k)_{k-2} & \cdots & (\ell_k)_2 & \ell_k & 1 \end{pmatrix}$$

This determinant is in fact equal to $\prod_{1 \leq i < j \leq k} (\ell_i - \ell_j)$. To see this, we can either use column operations and reduce it to the matrix $a_\rho(\ell_1, \dots, \ell_k)$. Alternatively, replace the ℓ_i with variables x_i, and note that $(x_i - x_j)$ divides the determinant for all $i < j$ and that it has the same degree and leading coefficient as $\prod_{1 \leq i < j \leq k} (x_i - x_j)$. □

We can deduce another nice combinatorial formula from this one using the notion of hook lengths. Given a box (i,j) in the Young diagram of λ, its hook is the set of boxes to the right and below it (including itself). Its **hook length** $h(i,j)$ is the number of boxes in the book. Below, we list the hook lengths for the partition $(6,3,1)$:

8	6	5	3	2	1
4	2	1			
1					

Theorem 6.1.2 (Hook length formula)**.** *If λ is a partition of n with Young diagram $Y(\lambda)$, then*

$$f^\lambda = \frac{n!}{\prod_{(i,j)\in Y(\lambda)} h(i,j)}.$$

Example 6.1.3. Take $\lambda = (6,3,1)$. From the previous example, the hook length formula gives

$$f^{(6,3,1)} = \frac{10!}{8\cdot 6\cdot 5\cdot 3\cdot 2\cdot 4\cdot 2} = 315.$$

The formula in Theorem 6.1.1 gives

$$f^{(6,3,1)} = \frac{10!}{8!4!}(8-1)(8-4)(4-1) = 315. \qquad \square$$

Proof. Let $g^\lambda = \frac{n!}{\prod_{(i,j)\in Y(\lambda)} h(i,j)}$. We will show by induction on the number of columns of λ that $f^\lambda = g^\lambda$ using the formula $f^\lambda = \frac{n!}{\ell_1!\cdots\ell_k!}\prod_{1\le i<j\le k}(\ell_i-\ell_j)$ where $k=\ell(\lambda)$ and $\ell_i = \lambda_i + k - i$. If the number of columns is 1, then $g^\lambda = 1$ and $f^\lambda = 1$ by its definition as the number of standard Young tableaux.

In general, let μ be the partition obtained from λ by removing its first column. Note that the hook lengths in the first column of λ are $\ell_1, \ell_2, \dots, \ell_k$, but otherwise, the hook lengths in the other boxes in λ are the hook lengths of the boxes in μ. Hence,

$$g^\mu = \frac{(n-k)!}{n!}\ell_1\cdots\ell_k g^\lambda.$$

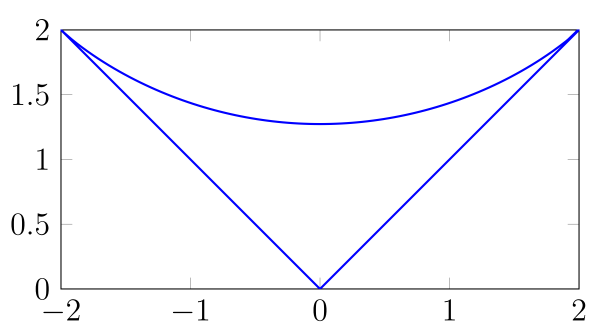
On the other hand, f^μ and f^λ satisfy the same relation, so by induction, we conclude that $f^\lambda = g^\lambda$. $\square$

Remark 6.1.4. The hook length statistic appears in an ad hoc way in the above derivation. For a more natural derivation that uses the hook lengths in an essential way, see [GNW]. $\square$

Remark 6.1.5. Note that $\sum_{\lambda\vdash n}(f^\lambda)^2 = n!$, so we put a probability measure on the partitions of n by choosing λ with probability $(f^\lambda)^2/n!$. Represent them by their Young diagram and normalize so that each box has area $1/n$. For the following, we will use the Russian convention. Work of Logan–Shepp and Vershik–Kerov show that there is a limiting curve for the boundary of our randomly chosen partition, and even give a formula for it:

$$\Omega(x) = \begin{cases} \frac{2}{\pi}(x\arcsin(\frac{x}{2}) + \sqrt{4-x^2}) & \text{if } |x|\le 2 \\ |x| & \text{if } |x| > 2 \end{cases}.$$

We plot Ω as the top curve below. The bottom portion is $|x|$ and represents the sides of other boundaries of the partition.



See [O] for a survey and further references. This also shows that the largest part of a random partition λ is $\lambda_1 \sim 2\sqrt{n}$ (and symmetrically, $\ell(\lambda) \sim 2\sqrt{n}$). $\square$

6.2. **Semistandard Young tableaux.** Now we derive a formula for the number of semistandard Young tableaux of shape λ using the numbers $1, \dots, k$, i.e., for the evaluation $s_\lambda(1, 1, \dots, 1)$ (k instances of 1). This is the dimension of an irreducible polynomial representation $\mathbf{S}_\lambda(\mathbf{C}^k)$ of $\mathbf{GL}_k(\mathbf{C})$.

Theorem 6.2.1.

$$\dim \mathbf{S}_\lambda(\mathbf{C}^k) = s_\lambda(1, \dots, 1) = \prod_{1 \le i < j \le k} \frac{\lambda_i - \lambda_j + j - i}{j - i}.$$

where there are k instances of 1 above.

Proof. Work in finitely many variables $x_1, \dots, x_k$ and use the determinantal formula in §3.4

$$s_\lambda(x_1, \dots, x_k) = \frac{\det(x_i^{\lambda_j + k - j})_{i,j=1}^k}{\det(x_i^{k-j})_{i,j=1}^k}.$$

We can't evaluate $x_i = 1$ directly since we'd get 0/0, but the following method let's us get around that. Let q be a new indeterminate and set $x_i = q^{i-1}$. Then

$$s_\lambda(1, q, \dots, q^{k-1}) = \frac{\det(q^{i(\lambda_j + k - j)})_{i,j=1}^k}{\det(q^{i(k-j)})_{i,j=1}^k}.$$

Now in fact, both determinants become Vandermonde matrices, so we can simplify:

$$s_\lambda(1, q, \dots, q^{k-1}) = \frac{\prod_{1 \le i < j \le k} (q^{\lambda_i + k - i} - q^{\lambda_j + k - j})}{\prod_{1 \le i < j \le k} (q^{k-i} - q^{k-j})} = \prod_{1 \le i < j \le k} q^{\lambda_j} \frac{q^{\lambda_i - \lambda_j + j - i} - 1}{q^{j-i} - 1}.$$

Now we can set $q = 1$ in the final expression (either by dividing the polynomials, or using l'Hôpital's rule) and get the desired formula. $\square$

Corollary 6.2.2. *Keep k fixed. The function $n \mapsto \dim \mathbf{S}_{n\lambda}(\mathbf{C}^k)$ is a polynomial in n whose degree is the number of pairs $i < j$ such that $\lambda_i \ne \lambda_j$.*

Remark 6.2.3. (For those who know some algebraic geometry) The function $n \mapsto \dim \mathbf{S}_{n\lambda}(\mathbf{C}^k)$ is actually the Hilbert function of a projective embedding of a partial flag variety. More specifically, let $1 \le i_1 < i_2 < \cdots < i_r$ be the indices such that $\lambda_{i_j} \ne \lambda_{i_j+1}$. Then the collection of subspaces $F_1 \subset \cdots \subset F_r \subset \mathbf{C}^k$ where $\dim F_j = i_j$ has the structure of a projective algebraic variety which admits an embedding into the projective space on $\mathbf{S}_\lambda(\mathbf{C}^k)$ giving rise to the Hilbert function we're talking about. $\square$

Given a box $(i, j) \in Y(\lambda)$, define its **content** to be $c(i, j) = j - i$.

Theorem 6.2.4 (Hook-content formula)**.**

$$\dim \mathbf{S}_\lambda(\mathbf{C}^k) = s_\lambda(1, \dots, 1) = \prod_{(i,j) \in Y(\lambda)} \frac{k + c(i, j)}{h(i, j)}$$

where there are k instances of 1 above.

Proof. Let $n = |\lambda|$ and set $\ell_i = \lambda_i + k - i$. Using the formulas for f^λ in the previous section, we have

$$\prod_{(i,j)\in Y(\lambda)} \frac{k + c(i,j)}{h(i,j)} = \frac{f^\lambda}{n!} \prod_{(i,j)\in Y(\lambda)} (k - i + j) = \frac{1}{\ell_1! \cdots \ell_k!} \prod_{1\le i<j\le k} (\ell_i - \ell_j) \prod_{(i,j)\in Y(\lambda)} (k - i + j).$$

Next, note that

$$\prod_{(i,j)\in Y(\lambda)} (k - i + j) = \prod_{i=1}^{k} \frac{\ell_i!}{(k-i)!},$$

so the above simplifies to

$$\frac{\prod_{1\le i<j\le k}(\ell_i - \ell_j)}{(k-1)!(k-2)!\cdots 2!} = \prod_{1\le i<j\le k} \frac{\lambda_i - \lambda_j - i + j}{j - i},$$

and the latter we have shown to be $\dim \mathbf{S}_\lambda(\mathbf{C}^k)$. $\square$

Corollary 6.2.5. *For each partition λ, the function $n \mapsto \dim \mathbf{S}_\lambda(\mathbf{C}^n)$ is a polynomial in n.*

The Jacobi–Trudi identity (Theorem 3.6.1) gives yet another formula. It is easy to see that $h_n(1, \dots, 1) = \binom{n+k-1}{n}$ (k 1's here, and this is the number of monomials of degree n with k variables).

Theorem 6.2.6. *If $n \ge \ell(\lambda)$, then*

$$\dim \mathbf{S}_\lambda(\mathbf{C}^k) = s_\lambda(1, \dots, 1) = \det\left(\binom{k + \lambda_i - i + j - 1}{k-1}\right)_{i,j=1}^{n}.$$

6.3. **Littlewood–Richardson coefficients.** We have encountered Littlewood–Richardson coefficients $c^\nu_{\lambda,\mu}$ in several contexts now:

- The multiplication of Schur functions: $s_\lambda s_\mu = \sum_\nu c^\nu_{\lambda,\mu} s_\nu$,
- The expansion of a skew Schur function: $s_{\nu/\mu} = \sum_\lambda c^\nu_{\lambda,\mu} s_\lambda$,
- The induction of symmetric group characters: $\chi^\lambda \circ \chi^\mu = \sum_\nu c^\nu_{\lambda,\mu} \chi^\nu$,
- The restriction of a symmetric group character: $\mathrm{Res}^{\Sigma_{n+m}}_{\Sigma_n\times\Sigma_m} \chi^\nu = \sum_{\lambda,\mu} c^\nu_{\lambda,\mu} (\chi^\lambda \otimes \chi^\mu)$,
- The multiplication of Schubert classes: $\sigma_\lambda \sigma_\mu = \sum_\nu c^\nu_{\lambda,\mu} \sigma_\nu$

We'll see another instance in the next section. Here we'll give one way to compute $c^\nu_{\lambda,\mu}$ (without proof). See [Sta, §7, Appendix 1] for details and more formulas.

Let $w = w_1 w_2 \cdots w_n$ be a sequence of positive integers and let $m_i(w)$ be the number of w_j equal to i. A prefix of w is any subsequence of the form $w_1 w_2 \cdots w_m$ for $m \le n$. We say that w is a **lattice permutation** (also called **Yamanouchi word** or **ballot sequence**) if, for every prefix v of w, we have $m_i(v) \ge m_{i+1}(v)$ for all i. Given a tableau T, its **reverse reading word** is the list of the entries of T in the following order: start with row 1 and list the entries from right to left, move to row 2 and list the entries from right to left, etc.

Call a Littlewood–Richardson tableau a SSYT of skew shape whose reverse reading word is a lattice permutation.

Theorem 6.3.1. *$c^\nu_{\lambda,\mu}$ is the number of Littlewood–Richardson tableaux of shape ν/μ and type λ.*

Recall that $c^\nu_{\lambda,\mu} = c^\nu_{\mu,\lambda}$, so there is a big asymmetry in this description for $c^\nu_{\lambda,\mu}$. This can sometimes be a good thing: one set of SSYT may be much easier to describe than the other, even though they must have the same size. It is also possible to give descriptions that are symmetric in μ and λ, but we will not discuss that here.

Example 6.3.2. Let $\lambda = (4, 2, 1)$, $\mu = (5, 2)$, $\nu = (6, 5, 2, 1)$. Then $c^\nu_{\lambda,\mu} = 3$; here are all of the SSYT of shape ν/μ of type λ whose reverse reading words are lattice permutations together with their reverse reading words:

$$\begin{ytableau} \none & \none & \none & \none & \none & 1 \\ \none & \none & 1 & 1 & 1 \\ 2 & 2 \\ 3 \end{ytableau} \qquad \begin{ytableau} \none & \none & \none & \none & \none & 1 \\ \none & \none & 1 & 1 & 2 \\ 1 & 2 \\ 3 \end{ytableau} \qquad \begin{ytableau} \none & \none & \none & \none & \none & 1 \\ \none & \none & 1 & 1 & 2 \\ 1 & 3 \\ 2 \end{ytableau}$$
$$1111223 \qquad 1211213 \qquad 1211312$$

Alternatively, we could count the number of SSYT of shape ν/λ of type μ whose reverse reading words are lattice permutations. We list them here:

$$\begin{ytableau} \none & \none & \none & \none & 1 & 1 \\ \none & \none & 1 & 1 & 2 \\ \none & 1 \\ 2 \end{ytableau} \qquad \begin{ytableau} \none & \none & \none & \none & 1 & 1 \\ \none & \none & 1 & 1 & 2 \\ \none & 2 \\ 1 \end{ytableau} \qquad \begin{ytableau} \none & \none & \none & \none & 1 & 1 \\ \none & \none & 1 & 2 & 2 \\ \none & 1 \\ 1 \end{ytableau}$$
$$1121112 \qquad 1121121 \qquad 1122111$$

□

Remark 6.3.3. If $\lambda = (d)$, then $c^\nu_{\lambda,\mu}$ can be computed by the Pieri rule. In fact, the description given above directly generalizes this: a SSYT of shape ν/μ of type (d) cannot have more than one box in a single column. But given any collection of boxes, no two in a single column, putting a 1 in each box gives a valid SSYT whose reverse reading word is a lattice permutation. So we conclude that it's 1 if ν/μ is a horizontal strip and 0 otherwise.

Similarly, if $\lambda = (1^d)$, then consider a SSYT of shape ν/μ of type (1^d) whose reverse reading word is a lattice permutation. The reverse reading word must then be $123 \cdots d$. But since it's a SSYT, no two of these entries can appear in the same row. So we conclude that it's 1 if ν/μ is a vertical strip and 0 otherwise. □

7. Hall algebras

Our goal is to discuss another appearance of Littlewood–Richardson coefficients. To do this, we study the classical Hall algebra of finite abelian p-groups. This notion has been generalized substantially and plays many important roles in representation theory, see [Sc] as a starting point. We will also discuss Hall–Littlewood symmetric functions, which are a certain important class of symmetric functions with a parameter. See [Mac, Chapters II, III] for a more detailed treatment.

7.1. **Counting subgroups of abelian groups.**

Fix a prime number p. A finite abelian p-group is one of the form $\bigoplus_{i=1}^k \mathbf{Z}/p^{\lambda_i}$ for some positive integers λ_i. We may as well assume that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$ so that its isomorphism class is determined by a partition λ, which we call its **type**.

Define the **length** $\ell(M)$ of an abelian p-group M to be its length as a $\mathbf{Z}$-module, i.e.,

$$\ell(M) = \sum_{i \geq 1} \dim_{\mathbf{Z}/p}(p^{i-1}M/p^iM).$$

Lemma 7.1.1. *Let M be an abelian p-group of type λ. Then $p^{i-1}M/p^iM$ is a $\mathbf{Z}/p$-vector space for each $i \geq 1$, let μ_i be its dimension. Then $(\mu_1, \mu_2, \dots) = \lambda^\dagger$. In particular, $\ell(M/p^iM) = \lambda_1^\dagger + \cdots + \lambda_i^\dagger$.*

Proof. If $M = \mathbf{Z}/p^k$, then $\mu = 1^k$. In particular, $\mu_i = \#\{\lambda_j \mid \lambda_j \geq i\} = \lambda_i^\dagger$. ☐

Let M be an abelian p-group of type λ and set $G^\lambda_{\mu,\nu}(p)$ to be the number of submodules $N \subseteq M$ of type μ such that M/N is of type ν. These numbers have remarkable properties which we summarize in the next result.

Define $n(\lambda) = \sum_{i \geq 1}(i-1)\lambda_i = \frac{1}{2}\sum_{i\geq 1} \lambda_i^\dagger(\lambda_i^\dagger - 1)$.

Theorem 7.1.2. *There is a unique polynomial $g^\lambda_{\mu,\nu}(t) \in \mathbf{Z}[t]$ such that $G^\lambda_{\mu,\nu}(p) = g^\lambda_{\mu,\nu}(p)$.*

If $c^\lambda_{\mu,\nu} = 0$, then $g^\lambda_{\mu,\nu}(t)$ is the 0 polynomial. Otherwise, it is a polynomial of degree $n(\lambda) - n(\mu) - n(\nu)$ and its leading coefficient is $c^\lambda_{\mu,\nu}$.

Remark 7.1.3. More generally, instead of finite abelian p-groups, we can consider finite length modules over a local PID. In our setting, we're considering finite length modules over $\mathbf{Z}_{(p)}$ (alternatively, over the p-adic integers $\mathbf{Z}_p$). The above theorem is still true, and the polynomial $g^\lambda_{\mu,\nu}(t)$ works for all local PID's, and one substitutes the size of the residue field instead of p in general. The proofs below go through with little change, but we have focused on this case for simplicity of exposition. ☐

Let's start with basic properties:

Lemma 7.1.4. (a) *$G^\lambda_{\mu,\nu}(p) = G^\lambda_{\nu,\mu}(p)$.*
(b) *If $G^\lambda_{\mu,\nu}(p) \neq 0$, then $\mu \subseteq \lambda$ and $\nu \subseteq \lambda$.*

Proof. (a) To show this, we use Pontryagin duality. Let S^1 be the circle group (the length 1 complex numbers under multiplication). Given a finite abelian group M, let $M^\vee$ be the group of homomorphisms $M \to \mathrm{S}^1$ (the multiplication is given by $(ff')(x) = f(x)f'(x)$ for $x \in M$). Then $(M_1 \oplus M_2)^\vee \cong M_1^\vee \oplus M_2^\vee$ and $(\mathbf{Z}/n)^\vee \cong \mathbf{Z}/n$ since every homomorphism has to send a generator of $\mathbf{Z}/n$ to some nth root of unity. This implies that $M \cong M^\vee$ always.

More interestingly, if $N \subseteq M$, then we get a surjection $M^\vee \to N^\vee$ by restricting homomorphisms from M to N, and the kernel is naturally $(M/N)^\vee$. In particular, if M is an abelian p-group of type λ, then Pontryagin duality gives a bijection between subgroups of type μ whose quotient is of type ν and subgroups of $M^\vee$ of type ν whose quotient is of type μ. But M and $M^\vee$ have the same type.

(b) Using (a), it suffices to show that $\nu \subseteq \lambda$. Let M have type λ and suppose $N \subseteq M$ has type μ and M/N has type ν. Then

$$\frac{p^{i-1}(M/N)}{p^i(M/N)} \cong \frac{p^{i-1}M + N}{p^iM + N} = \frac{p^{i-1}M + (p^iM + N)}{p^iM + N} \cong \frac{p^{i-1}M}{p^{i-1}M \cap (p^iM + N)},$$

where the second isomorphism is the "second isomorphism theorem" for modules. Since $p^iM \subseteq p^{i-1}M \cap (p^iM + N)$, it follows that

$$\ell\left(\frac{p^{i-1}(M/N)}{p^i(M/N)}\right) \leq \ell\left(\frac{p^{i-1}M}{p^iM}\right),$$

which, by Lemma 7.1.1, implies that $\nu_i^\dagger \leq \lambda_i^\dagger$. In particular, $\nu^\dagger \subseteq \lambda^\dagger$, and hence $\nu \subseteq \lambda$. ☐

Example 7.1.5. An abelian p-group of type 1^n is just $(\mathbf{Z}/p)^n$. In particular, $G^{1^n}_{\mu,\nu}(p) = 0$ unless $\mu = 1^k$ and $\nu = 1^{n-k}$ for some k. In that case, $G^{1^n}_{1^k,1^{n-k}}(p)$ is just the number of k-dimensional subspaces of the n-dimensional vector space $(\mathbf{Z}/p)^n$. By Remark 5.2.3, this is the p-binomial coefficient $\begin{bmatrix} n \\ k \end{bmatrix}_p$. Note that it is a polynomial in p. ☐

Here's a generalization of this example:

Proposition 7.1.6. *If λ/μ is not a vertical strip of size k, then $G^\lambda_{1^k,\mu}(p) = 0$.*

Otherwise, we have

$$G^\lambda_{1^k,\mu}(p) = p^d \prod_{i\ge 1} \begin{bmatrix} \lambda^\dagger_i - \lambda^\dagger_{i+1} \\ \lambda^\dagger_i - \mu^\dagger_i \end{bmatrix}_p$$

where

$$d = \sum_{i\ge 1} \left((\lambda^\dagger_i - \mu^\dagger_i) \left(\sum_{j\ge i+1} \mu^\dagger_j - \sum_{j \ge i+2} \lambda^\dagger_j \right) \right).$$

In particular, it is a polynomial in p of degree $n(\lambda) - n(\mu) - n(1^k)$.

Proof. Let M be an abelian group of type λ and let $N \subset M$ be a subgroup of type 1^k and let μ be the type of M/N. Define $S = \ker(M \xrightarrow{\cdot p} M)$. Then $N \subseteq S$, and so $M/S \cong (M/N)/(S/N)$. The type of M/S is easily seen to be $\widetilde{\lambda} = (\lambda_1 - 1, \lambda_2 - 1, \dots, \lambda_r - 1)$. By Lemma 7.1.4, $\widetilde{\lambda} \subseteq \mu \subseteq \lambda$, so λ/μ is a vertical strip.

Now we count the number of such submodules N. Let $S_i = S \cap p^i M$ and $N_i = N \cap S_i$. Then $\ell(S_i) = \#\{j \mid \lambda_j > i\} = \lambda^\dagger_{i+1}$.

Note that $N/N_i \cong (N + p^i M)/p^i M$ and hence

$$\begin{aligned}
\ell(N_{i-1}) - \ell(N_i) &= -\ell(N/N_{i-1}) + \ell(N/N_i) \\
&= -\ell\left(\frac{N + p^{i-1}M}{p^{i-1}M}\right) + \ell\left(\frac{N + p^i M}{p^i M}\right) \\
&= -\ell(N + p^{i-1}M) + \ell(p^{i-1}M) + \ell(N + p^i M) - \ell(p^i M) \\
&= \ell\left(\frac{p^{i-1}M}{p^i M}\right) - \ell\left(\frac{N + p^{i-1}M}{N + p^i M}\right) \\
&= \lambda^\dagger_i - \mu^\dagger_i.
\end{aligned}$$

So we see that M/N has type μ if and only if $\ell(N_{i-1}/N_i) = \lambda^\dagger_i - \mu^\dagger_i$. So to count the number of N, it suffices to count the number of submodules $N_i \subseteq S_i$ such that $N_{i-1} \cap S_i = N_i$ and $\ell(N_{i-1}/N_i) = \lambda^\dagger_i - \mu^\dagger_i$.

To do this, first choose subspaces $W_{i-1} \subseteq S_{i-1}/S_i$ of dimension $\lambda^\dagger_i - \mu^\dagger_i$. The set of choices for a given i is a Grassmannian, and hence has size

$$\begin{bmatrix} \lambda^\dagger_i - \lambda^\dagger_{i+1} \\ \lambda^\dagger_i - \mu^\dagger_i \end{bmatrix}_p.$$

We first set $N_{\lambda_1 - 1} = W_{\lambda_1 - 1}$. To get $N_{\lambda_1 - 2}$ from $W_{\lambda_1 - 2}$, we can take any preimages of a basis for $W_{\lambda_1 - 2}$ under the map $S_{\lambda_1 - 2} \to S_{\lambda_1 - 2}/S_{\lambda_1 - 1}$ and take its span with $N_{\lambda_1 - 1}$. However, any preimages that differ by elements in N_{λ_1-1} give the same result, so the number of unique choices is p raised to the following power:

$$\ell(W_{\lambda_1-2})\ell(S_{\lambda_1-1}/N_{\lambda_1-1}).$$

Similarly, once we've chosen N_i in general and W_{i-1}, to get N_{i-1}, we take any preimages of a basis for W_{i-1} under the map $S_{i-1} \to S_{i-1}/S_i$ and take its span with N_i. The number of unique choices is p raised to the following power:

$$\ell(W_{i-1})\ell(S_{i-1}/N_{i-1}) = (\lambda_i^\dagger - \mu_i^\dagger)\left(\sum_{j\geq i}\mu_j^\dagger - \sum_{j\geq i+1}\lambda_j^\dagger\right).$$

So we get a polynomial in p. Finally, we compute its degree. Set $L_i = \lambda_i^\dagger$ and $M_i = \mu_i^\dagger$. The binomial coefficient $\begin{bmatrix} n \\ k \end{bmatrix}_p$ has degree $k(n-k)$, so the degree is

$$\begin{aligned}
&\sum_{i\geq 1}(L_i - M_i)(M_i - L_{i+1}) + \sum_{i\geq 1}(L_i - M_i)(\sum_{j\geq i+1} M_j - \sum_{j\geq i+2} L_j) \\
&= \sum_{i\geq 1}(L_i - M_i)(\sum_{j\geq i} M_j - \sum_{j\geq i+1} L_j) \\
&= -\sum_{j>i} L_iL_j - \sum_{j\geq i} M_iM_j + \sum_{j\geq i} L_iM_j + \sum_{j>i} M_iL_j.
\end{aligned}$$

On the other hand,

$$n(\lambda) - n(\mu) - n(1^m) = \sum_{i\geq 1}\frac{L_i(L_i-1)}{2} - \sum_{i\geq 1}\frac{M_i(M_i-1)}{2} - \frac{m(m-1)}{2}.$$

Now use that $m = \sum_{i\geq 1}(L_i - M_i)$ to simplify:

$$\begin{aligned}
&= \frac{1}{2}\left(\sum_{i\geq 1} L_i(L_i-1) - \sum_{i\geq 1} M_i(M_i-1) - (\sum_{i\geq 1}(L_i-M_i))(-1+\sum_{i\geq 1}(L_i-M_i))\right) \\
&= \frac{1}{2}\left(\sum_{i\geq 1} L_i^2 - \sum_{i\geq 1} M_i^2 - (\sum_{i\geq 1}(L_i-M_i))^2\right) \\
&= -\sum_{j>i} L_iL_j - \sum_{j\geq i} M_iM_j + \sum_{j\geq i} L_iM_j + \sum_{j>i} M_iL_j,
\end{aligned}$$

which is what we wanted. □

Steinitz, and in later independent work, Hall, used the numbers $G^\lambda_{\mu,\nu}(p)$ to define an algebra. Let $H = H(p)$ be the free abelian group with basis u_λ, with λ ranging over all partitions. Define a multiplication on H by

$$u_\mu u_\nu = \sum_\lambda G^\lambda_{\mu,\nu}(p)u_\lambda.$$

Proposition 7.1.7. *The product just defined is commutative and associative.*

Proof. Commutativity is equivalent to the statement $G^\lambda_{\mu,\nu}(p) = G^\lambda_{\nu,\mu}(p)$ which we've shown.

For associativity, consider the products $(u_\alpha u_\beta)u_\gamma$ and $u_\alpha(u_\beta u_\gamma)$. The coefficient of u_μ in each is

$$\sum_\lambda G^\mu_{\lambda,\gamma}(p)G^\lambda_{\alpha,\beta}(p), \qquad \sum_\lambda G^\mu_{\alpha,\lambda}(p)G^\lambda_{\beta,\gamma}(p).$$

Let M be an abelian p-group of type μ. Then both sums are counting the number of chains of subgroups $N_1 \subseteq N_2 \subseteq M$ such that

- N_1 has type α,
- N_2/N_1 has type β,
- M/N_2 has type γ.

So they are the same. In the first sum, λ is the type of N_2, while in the second sum it is the type of M/N_1. $\square$

H also has a multiplicative identity given by $u_\varnothing$. This is the **(classical) Hall algebra** or the **algebra of partitions**.

Proposition 7.1.8. *The elements $\{u_{1^r} \mid r \ge 1\}$ generate H as an algebra and are algebraically independent. Furthermore, there exist unique polynomials $g^\lambda_{\mu,\nu}(t)$ such that $G^\lambda_{\mu,\nu}(p) = g^\lambda_{\mu,\nu}(p)$.*

$g^\lambda_{\mu,\nu}(t)$ is called a **Hall polynomial**.

Proof. For a partition λ, set $v_\lambda = u_{1^{\lambda_1}} \cdots u_{1^{\lambda_k}}$. Then

$$v_\lambda = \sum_\mu A_{\lambda,\mu}(p)u_\mu$$

where $A_{\lambda,\mu}(p)$ is the number of chains of subgroups

$$M = M_0 \supset M_1 \supset \cdots \supset M_k = 0$$

such that M is an abelian p-group of type μ, and M_{i-1}/M_i is of type 1^{λ_i} for each i, i.e., $M_{i-1}/M_i \cong (\mathbf{Z}/p)^{\lambda_i}$. In particular, $pM_{i-1} \subseteq M_i$, and so $p^iM \subseteq M_i$ for all i.

In particular, if such a chain exists, i.e., $A_{\lambda,\mu}(p) \neq 0$, then it implies that $\ell(M/p^iM) \ge \ell(M/M_i)$, and so by Lemma 7.1.1, we conclude that $\mu^\dagger_1 + \cdots + \mu^\dagger_i \ge \lambda_1 + \cdots + \lambda_i$, i.e., $\mu^\dagger \ge \lambda$ (dominance order). If $\lambda = \mu^\dagger$, then we must have $M_i = p^iM$, and so $A_{\lambda,\lambda^\dagger}(p) = 1$. By Lemma 2.1 (though the inequality is reversed), the change of basis matrix between u_λ and v_λ is lower-unitriangular.

Furthermore, by Proposition 7.1.6, there exists a polynomial $a_{\lambda,\mu}(t)$ such that $A_{\lambda,\mu}(p) = a_{\lambda,\mu}(p)$ and $a_{\lambda,\lambda^\dagger}(t) = 1$. In particular, by the invertibility of the matrix above there exist polynomials, which when specialized at p, give the coefficients of u_λ written in terms of the v_μ. Now if we multiply out $u_\mu u_\nu$, we get a polynomial linear combination of v_λ, and hence a polynomial linear combination of u_λ. $\square$

In particular, we can define a **universal Hall algebra** $\mathbf{H} = H \otimes \mathbf{Z}[t]$ where we replace $G^\lambda_{\mu,\nu}(p)$ with the polynomial $g^\lambda_{\mu,\nu}(t)$. Each Hall algebra is obtained by specializing the variable t to the prime p.

Remark 7.1.9. Our proof of the existence of the Hall polynomials relied on one calculation but was otherwise conceptually simple. The downside is that it is difficult to deduce any properties about these polynomials. For that, see [Mac, Chapter II]. There, much more explicit information is worked out at the cost of heavier amounts of calculation and combinatorics. ☐

Theorem 7.1.10. *Let M be an abelian p-group of type λ, and let $N \subseteq M$ be a subgroup of type ν such that M/N has type μ. For each $i \geq 0$, let $\lambda^{(i)}$ be the type of M/p^iN. Then the sequence*

$$\lambda^{(0)\dagger} \subseteq \lambda^{(1)\dagger} \subseteq \cdots \subseteq \lambda^{(r)\dagger}$$

(where $p^rN = 0$) corresponds to a Littlewood–Richardson tableau of shape $\lambda^\dagger/\mu^\dagger$ and type $\nu^\dagger$.

Proof. Since M/p^iN is a quotient of $M/p^{i+1}N$, we have $\lambda^{(i)\dagger} \subseteq \lambda^{(i+1)\dagger}$ by Lemma 7.1.4 and Lemma 7.1.1. By definition, $\lambda^{(0)\dagger} = \mu^\dagger$ while $\lambda^{(r)\dagger} = \lambda^\dagger$. By Proposition 7.1.6, $\lambda^{(i)}/\lambda^{(i-1)}$ is a vertical strip and hence $\lambda^{(i)^\dagger}/\lambda^{(i-1)\dagger}$ is a horizontal strip. Hence we get a SSYT of shape $\lambda^\dagger/\mu^\dagger$ and type $\nu^\dagger$ (again by Lemma 7.1.1) by filling in i in the boxes of $\lambda^{(i)\dagger}/\lambda^{(i-1)\dagger}$.

Finally, we have to show that the reverse reading word of this SSYT is a lattice permutation. Since we already know it is a SSYT, we can reformulate this by saying that the number of instances of i in the first k rows is at least as many as the number of instances of $i+1$ in the first $k+1$ rows for all i and k. Equivalently, for all i and k, we have

$$\sum_{j=1}^{k}(\lambda_j^{(i)\dagger} - \lambda_j^{(i-1)\dagger}) \geq \sum_{j=1}^{k+1}(\lambda_j^{(i+1)\dagger} - \lambda_j^{(i)\dagger}).$$

To show this, first note that by Lemma 7.1.1,

$$\lambda_j^{(i)\dagger} = \ell(p^{j-1}(M/p^iN)/p^j(M/p^iN)).$$

Hence

$$\sum_{j=1}^{k}\lambda_j^{(i)\dagger} = \ell((M/p^iN)/p^k(M/p^iN)) = \ell(M/(p^iN + p^kM)).$$

Set $V_{ki} = (p^kM + p^{i-1}N)/(p^kM + p^iN)$ so that

$$\sum_{j=1}^{k}(\lambda_j^{(i)\dagger} - \lambda_j^{(i-1)\dagger}) = \ell(V_{ki}).$$

Finally, $\ell(V_{ki}) \geq \ell(V_{k+1,i+1})$ since multiplication by p gives a surjective map $V_{ki} \to V_{k+1,i+1}$. ☐

Corollary 7.1.11. *If $g^\lambda_{\mu,\nu}(t) \neq 0$, then $c^\lambda_{\mu,\nu} \neq 0$.*

Proof. Follows from Theorem 6.3.1. ☐

7.2. **Hall–Littlewood symmetric functions.** Let $x_1, \dots, x_n, t$ be variables and pick a partition $\lambda = (\lambda_1, \dots, \lambda_n)$. Let Σ_n^λ be the stabilizer of λ in Σ_n.

Set

$$v_n(t) = \sum_{\sigma \in \Sigma_n} \sigma\left(\prod_{1 \leq i < j \leq n} \frac{x_i - tx_j}{x_i - x_j}\right), \qquad v_\lambda(t) = \prod_{i \geq 0} v_{m_i(\lambda)}(t).$$

(Note that we include $m_0(\lambda)$; ordinarily this is 0, but for the proof below, we use $m_0(\lambda) = n - \ell(\lambda)$.) This is a polynomial in t which does not depend on $x_1, \dots, x_n$: the denominator of the σ summand is $\mathrm{sgn}(\sigma)$ times the Vandermonde determinant, and the expression

$$\sum_{\sigma\in\Sigma_n}\sigma\left(\prod_{1\le i<j\le n}(x_i-tx_j)\right)$$

is skew-symmetric (and of the same degree).

We will need another formula for $v_n(t)$ later. Given $\sigma\in\Sigma_n$, define

$$\ell(\sigma)=\#\{i<j\mid\sigma(i)>\sigma(j)\}.$$

Also, define $[i]_t=\frac{1-t^i}{1-t}$ and $[n]_t!=[n]_t[n-1]_t\cdots[2]_t[1]_t$.

Lemma 7.2.1. $v_n(t)=\sum_{\sigma\in\Sigma_n}t^{\ell(\sigma)}=[n]_t!$.

Proof. From above, we just need to calculate the coefficient of $x_1^{n-1}x_2^{n-2}\cdots x_{n-1}$ in

$$\sum_{\sigma\in\Sigma_n}\mathrm{sgn}(\sigma)\sigma\left(\prod_{1\le i<j\le n}(x_i-tx_j)\right).$$

We get one power of t for each σ, and it is clear from our definition that the power is $\mathrm{sgn}(\sigma)(-t)^{\ell(\sigma^{-1})}$ since we can equivalently find the coefficient of $x_{\sigma^{-1}(1)}^{n-1}\cdots x_{\sigma^{-1}(n-1)}$ in $\mathrm{sgn}(\sigma)\prod_{i<j}(x_i-tx_j)$. However, we also have $\ell(\sigma)=\ell(\sigma^{-1})$ by definition, and $(-1)^{\ell(\sigma)}=\mathrm{sgn}(\sigma)$. This proves the first equality.

We prove the second equality by induction on n; when $n=1$ both sides are just 1. We have $[n]_t!=(1+t+\cdots+t^{n-1})[n-1]_t!$, so it suffices to show that

$$\sum_{\sigma\in\Sigma_n}t^{\ell(\sigma)}=(1+t+\cdots+t^{n-1})\sum_{\tau\in\Sigma_{n-1}}t^{\ell(\tau)}.$$

Writing permutations in one-line notation $a_1\cdots a_n$, i.e., $\sigma(i)=a_i$, we can construct a permutation $\sigma\in\Sigma_n$ from a permutation in $\tau\in\Sigma_{n-1}$ by arbitrarily inserting n somewhere. If we insert if right before a_i then $\ell(\sigma)=\ell(\tau)+n-i$, and if we insert it at the end, $\ell(\sigma)=\ell(\tau)$. This proves the desired identity. $\square$

Lemma 7.2.2. *We have*

$$\frac{1}{v_\lambda(t)}\sum_{\sigma\in\Sigma_n}\sigma\left(x_1^{\lambda_1}\cdots x_n^{\lambda_n}\prod_{1\le i<j\le n}\frac{x_i-tx_j}{x_i-x_j}\right)=\sum_{\sigma\in\Sigma_n/\Sigma_n^\lambda}\sigma\left(x_1^{\lambda_1}\cdots x_n^{\lambda_n}\prod_{\substack{i,j\\ \lambda_i>\lambda_j}}\frac{x_i-tx_j}{x_i-x_j}\right).$$

Proof. Choose coset representatives $\alpha_1,\alpha_2,\dots,\alpha_d$ for $\Sigma_n/\Sigma_n^\lambda$. We have

$$\sum_{\beta\in\Sigma_n^\lambda}\beta\prod_{i<j}\frac{x_i-tx_j}{x_i-x_j}=\sum_{(\beta_{\lambda_1},\dots,\beta_0)}\prod_{i<j}\frac{x_i-tx_j}{x_i-x_j}=v_\lambda(t)\prod_{\lambda_i>\lambda_j}\frac{x_i-tx_j}{x_i-x_j}$$

where the second sum is over all tuples in $\Sigma_{m_{\lambda_1}(\lambda)}\times\cdots\times\Sigma_{m_0(\lambda)}$. To see this, we think of β_{λ_1} as acting on the first $m_{\lambda_1}(\lambda)$ variables, β_{λ_1-1} as acting on the next $m_{\lambda_1-1}(\lambda)$ variables, etc., and each copy of $v_{m_i(\lambda)}(t)$ that we pick up involves different x variables – but the point is that v is independent of the x's. Then

$$\frac{1}{v_\lambda(t)}\sum_{\sigma\in\Sigma_n}\sigma\left(x_1^{\lambda_1}\cdots x_n^{\lambda_n}\prod_{1\le i<j\le n}\frac{x_i-tx_j}{x_i-x_j}\right)=\frac{1}{v_\lambda(t)}\sum_{k=1}^{d}\sum_{\beta\in\Sigma_n^\lambda}\alpha_k(x_1^{\lambda_1}\cdots x_n^{\lambda_n})\alpha_k\beta\left(\prod_{i<j}\frac{x_i-tx_j}{x_i-x_j}\right)$$
$$=\sum_{\sigma\in\Sigma_n/\Sigma_n^\lambda}\sigma\left(x_1^{\lambda_1}\cdots x_n^{\lambda_n}\prod_{\substack{i,j\\ \lambda_i>\lambda_j}}\frac{x_i-tx_j}{x_i-x_j}\right). \qquad \square$$

Let $P_\lambda(x_1,\dots,x_n;t)$ denote either of these expressions. This is the **Hall–Littlewood polynomial**. Set $\Lambda(n)[t]=\Lambda(n)\otimes\mathbf{Z}[t]$ and $\Lambda[t]=\Lambda\otimes\mathbf{Z}[t]$.

Proposition 7.2.3. (a) *$P_\lambda(x_1,\dots,x_n;t)$ is an element of $\Lambda(n)[t]$.*
(b) *$P_\lambda(x_1,\dots,x_n;0)=s_\lambda(x_1,\dots,x_n)$.*
(c) *$P_\lambda(x_1,\dots,x_n;1)=m_\lambda(x_1,\dots,x_n)$.*
(d) *$P_\lambda(x_1,\dots,x_n,0;t)=P_\lambda(x_1,\dots,x_n;t)$.*

Proof. Consider the two expressions for P_λ given by Lemma 7.2.2.

(a) The left expression for $P_\lambda(x_1,\dots,x_n;t)$ exhibits it as a quotient of a skew-symmetric polynomial in $x_1,\dots,x_n$ by the Vandermonde determinant, now use Lemma 3.4.1. The right expression does not involve any division by t.

(b) First, $v_n(0)=1$ since it simplifies to the quotient of two polynomials which are both the Vandermonde determinant. Now plug in $t=0$ into the left expression to get the determinantal formula for $s_\lambda(x_1,\dots,x_n)$ from §3.4.

(c) This is clear by plugging in $t=1$ into the right expression.

(d) This is clear from the right expression. $\square$

Using (d) above, we can define $P_\lambda(x;t)\in\Lambda[t]$ by taking it to be the unique symmetric function such that $P_\lambda(x_1,\dots,x_n,0,0,\dots;t)=P_\lambda(x_1,\dots,x_n;t)$ for all $n\ge\ell(\lambda)$. This is the **Hall–Littlewood symmetric function**. From (b) and (c) above, we see that it interpolates between monomial symmetric functions and Schur functions.

We think of $\Lambda[t]$ as being a $\mathbf{Z}[t]$-module rather than an abelian group. As such, it has a basis given by the Schur functions. Hence, we have expressions

$$P_\lambda(x;t)=\sum_\mu w_{\lambda,\mu}(t)s_\mu(x)$$

for some polynomials $w_{\lambda,\mu}(t)\in\mathbf{Z}[t]$. The following can be done by going through our definitions carefully, but we omit it (see [Mac, §III.1] for proofs):

Proposition 7.2.4. *$w_{\lambda,\mu}(t)\neq 0$ implies that $|\lambda|=|\mu|$ and $\lambda\ge\mu$. Furthermore, $w_{\lambda,\lambda}(t)=1$.*

Corollary 7.2.5. *$\{P_\lambda(x;t)\}$ is a $\mathbf{Z}[t]$-basis for $\Lambda[t]$.*

In particular, the product of two Hall–Littlewood functions is a $\mathbf{Z}[t]$-linear combination of others:

$$P_\mu(x;t)P_\nu(x;t)=\sum_\lambda f^\lambda_{\mu,\nu}(t)P_\lambda(x;t).$$

Since $P_\mu(x;0)=s_\mu$, we immediately see that

$$f^\lambda_{\mu,\nu}(0)=c^\lambda_{\mu,\nu}.$$

Our goal now is to connect these polynomials to Hall polynomials. Since we have a Pieri formula for Hall polynomials, we need to establish the analogue for Hall–Littlewood symmetric functions.

Lemma 7.2.6. *$P_{1^m}(x;t) = e_m(x)$.*

Proof. Since $P_{1^m}(x;t)$ is a sum of Schur functions s_λ with $|\lambda| = m$, it suffices to check this in finitely many variables $x_1, \dots, x_m$. In that case, Lemma 7.2.2 gives

$$P_{1^m}(x_1, \dots, x_m; t) = x_1 \cdots x_m = e_m(x_1, \dots, x_m). \qquad \square$$

Lemma 7.2.7.

$$\begin{bmatrix} n \\ r \end{bmatrix}_q = \frac{[n]_q!}{[r]_q![n-r]_q!}.$$

Proof. Recall we said that the left hand side is the number of r-dimensional subspaces of $\mathbf{F}_q^n$. Note that the number of pairs $(W, (v_1, \dots, v_r))$ where W is r-dimensional and $v_1, \dots, v_r$ is an ordered basis for W is $(q^n - 1)(q^n - q) \cdots (q^n - q^{r-1})$. Each subspace W has $(q^r - 1)(q^r - q) \cdots (q^r - q^{r-1})$ ordered bases, so we conclude that

$$\begin{aligned} \begin{bmatrix} n \\ r \end{bmatrix}_q &= \frac{(q^n - 1)(q^n - q) \cdots (q^n - q^{r-1})}{(q^r - 1)(q^r - q) \cdots (q^r - q^{r-1})} \\ &= \frac{(q^n - 1)(q^{n-1} - 1) \cdots (q^{n-r+1} - 1)}{(q^r - 1)(q^{r-1} - 1) \cdots (q - 1)} = \frac{[n]_q!}{[r]_q![n-r]_q!}. \qquad \square \end{aligned}$$

Proposition 7.2.8. *If $|\lambda| = |\mu| + m$, then*

$$f^\lambda_{\mu,1^m}(t) = \prod_{i \geq 1} \begin{bmatrix} \lambda^\dagger_i - \lambda^\dagger_{i+1} \\ \lambda^\dagger_i - \mu^\dagger_i \end{bmatrix}_t.$$

Proof. By Lemma 7.2.6, we need to find the coefficient of $P_\lambda(x;t)$ in $P_\mu(x;t) e_m(x)$. It will suffice to work in finitely many variables n where $n \geq \ell(\mu) + m$. Set $k = \mu_1$ and write $\{x_1, \dots, x_n\} = X_0 \amalg \cdots \amalg X_k$ where $x_j \in X_i$ if $\mu_j = i$. In particular, $|X_i| = m_i(\mu)$. Let $e_r(X_i)$ be the rth elementary symmetric function in the variables in X_i. Then

$$e_m(x_1, \dots, x_n) = \sum_{\mathbf{r} = (r_0, \dots, r_k)} e_{r_0}(X_0) \cdots e_{r_k}(X_k).$$

where the sum is over all tuples of non-negative integers $\mathbf{r}$ such that $r_i \leq m_i(\mu)$ and $\sum_i r_i = m$. Note that $m_i(\mu) = \mu^\dagger_i - \mu^\dagger_{i+1}$. Hence, given $\mathbf{r}$ as above, if we define a partition $\lambda(\mathbf{r})$ by adding r_{i-1} boxes to the ith column of μ, then $\lambda(\mathbf{r})/\mu$ is a vertical strip of size m, and these are all possible ways of adding a vertical strip of size m to μ.

If $X_i = \{y_1, \dots, y_{r_i}\}$ (written in order), then

$$e_{r_i}(X_i) = P_{1^{r_i}}(X_i; t) = \frac{1}{v_{r_i}(t) v_{m_i(\mu) - r_i}(t)} \sum_{\sigma \in \mathrm{Aut}(\{1, \dots, r_i\})} \sigma \left(y_1 y_2 \cdots y_{r_i} \prod_{1 \leq i < j \leq m_i(\mu)} \frac{y_i - t y_j}{y_i - y_j} \right).$$

In particular,

$$e_{r_0}(X_0) \cdots e_{r_k}(X_k) = \left(\prod_{i=0}^k v_{r_i}(t) v_{m_i(\mu)}(t) \right)^{-1} \sum_{\sigma \in \Sigma^\mu_n} \sigma \left(x_1^{\lambda(\mathbf{r})_1 - \mu_1} \cdots x_n^{\lambda(\mathbf{r})_n - \mu_n} \prod_{\substack{i<j \\ \mu_i = \mu_j}} \frac{x_i - t x_j}{x_i - x_j} \right).$$

If we multiply this by

$$P_\mu(x;t) = \sum_{\sigma \in \Sigma_n/\Sigma_n^\mu} \sigma\left(x_1^{\mu_1} \cdots x_n^{\mu_n} \prod_{\substack{i,j \\ \mu_i > \mu_j}} \frac{x_i - tx_j}{x_i - x_j} \right),$$

we get

$$\begin{aligned} P_\mu(x;t)e_{r_0}(X_0)\cdots e_{r_k}(X_k) &= \left(\prod_{i=0}^k v_{r_i}(t) v_{m_i(\mu)}(t) \right)^{-1} \sum_{\sigma \in \Sigma_n} \sigma\left(x_1^{\lambda(\mathbf{r})_1} \cdots x_n^{\lambda(\mathbf{r})_n} \prod_{i<j} \frac{x_i - tx_j}{x_i - x_j} \right) \\ &= \left(\prod_{i=0}^k v_{r_i}(t) v_{m_i(\mu)}(t) \right)^{-1} v_{\lambda(\mathbf{r})}(t) P_{\lambda(\mathbf{r})}(x;t). \end{aligned}$$

This implies that $f^{\lambda(\mathbf{r})}_{\mu,1^m}(t) = \left(\prod_{i=0}^k v_{r_i}(t)v_{m_i(\mu)}(t)\right)^{-1} v_{\lambda(\mathbf{r})}(t)$. Finally, we note that

$$\begin{aligned} \left(\prod_{i=0}^k v_{r_i}(t) v_{m_i(\mu)}(t) \right)^{-1} v_{\lambda(\mathbf{r})}(t) &= \prod_{i \ge 0} \frac{v_{m_i(\lambda(\mathbf{r}))}(t)}{v_{r_i}(t) v_{m_i(\mu)}(t)} \\ &= \prod_{i \ge 1} \begin{bmatrix} \lambda(\mathbf{r})_i^\dagger - \lambda(\mathbf{r})_{i+1}^\dagger \\ \lambda(\mathbf{r})_i^\dagger - \mu_i^\dagger \end{bmatrix}_t . \end{aligned}$$ □

Recall that $n(\lambda) = \sum_{i \ge 1} (i-1)\lambda_i$.

Corollary 7.2.9. *We have*

$$g^\lambda_{\mu,1^m}(t) = t^{n(\lambda)-n(\mu)-n(1^m)} f^\lambda_{\mu,1^m}(t^{-1}).$$

Proof. Let $f(t) = f^\lambda_{\mu,1^m}(t)$ and $g(t) = g^\lambda_{\mu,1^m}(t)$. Let $D = \deg f(t)$ and $E = \deg g(t) = n(\lambda) - n(\mu) - n(1^m)$. Since f is a product of binomial coefficients, it is palindromic, i.e., $f(t^{-1}) = t^{-D} f(t)$. From Proposition 7.1.6, $g(t) = t^{E-D} f(t)$, so we're done. □

Now define a linear map

$$\begin{aligned} \psi \colon H(p) \otimes \mathbf{Q} &\to \Lambda_{\mathbf{Q}} \\ \psi(u_\lambda) &= p^{-n(\lambda)} P_\lambda(x; p^{-1}). \end{aligned}$$

Theorem 7.2.10. *ψ is a ring isomorphism.*

Proof. ψ maps a basis to a basis, so the isomorphism part is clear. It remains to show it is a ring homomorphism. Since the u_{1^r} freely generate $H(p)$ by Proposition 7.1.8, we can define a ring homomorphism $\psi' \colon H(p) \otimes \mathbf{Q} \to \Lambda_{\mathbf{Q}}$ by $\psi'(u_{1^r}) = p^{-n(1^r)} e_r$. We will show by induction on λ (via $|\lambda|$ and then by dominance order) and $\psi(u_\lambda) = \psi'(u_\lambda)$. When $\lambda = \varnothing$ this is clear.

If $\lambda \ne \varnothing$, let μ be obtained from λ by removing its first column, and let m be the number of boxes of this first column. It is clear that $G^\lambda_{1^m,\mu}(p) = 1$ since any submodule of type 1^m of a module of type λ must be the kernel of the multiplication by p map. Furthermore, if ν/μ is a vertical strip, then $\nu \le \lambda$ in dominance order. Hence,

$$u_\mu u_{1^m} = u_\lambda + \sum_{\nu < \lambda} G^\nu_{\mu,1^m}(p) u_\nu.$$

Apply ψ' to both sides and use that $\psi' = \psi$ on everything involved except possibly u_λ:
$$p^{-n(\mu)-n(1^m)} P_\mu(x;p^{-1}) e_m(x) = \psi'(u_\lambda) + \sum_{\nu < \lambda} G^\nu_{\mu,1^m}(p) p^{-n(\nu)} P_\nu(x;p^{-1}).$$
Now write
$$P_\mu(x;p^{-1}) e_m(x) = P_\lambda(x;p^{-1}) + \sum_\nu f^\nu_{\mu,1^m}(p^{-1}) P_\nu(x;p^{-1})$$
($f^\lambda_{\mu,1^m}(p^{-1}) = 1$ by Proposition 7.2.8). Comparing both equalities and using Corollary 7.2.9 gives $\psi'(u_\lambda) = p^{n(\mu)+n(1^m)} P_\lambda(x;p^{-1})$. From how we defined μ, we have $n(\mu)+n(1^m) = n(\lambda)$, so we're done. $\square$

Corollary 7.2.11. *For all primes p, we have*
$$G^\lambda_{\mu,\nu}(p) = p^{n(\lambda)-n(\mu)-n(\nu)} f^\lambda_{\mu,\nu}(p^{-1}).$$
In particular, $g^\lambda_{\mu,\nu}(t)$ is a polynomial of degree $\le n(\lambda)-n(\mu)-n(\nu)$ and the coefficient of $t^{n(\lambda)-n(\mu)-n(\nu)}$ is $c^\lambda_{\mu,\nu}$. Furthermore, if $c^\lambda_{\mu,\nu} = 0$, then $g^\lambda_{\mu,\nu}(t)$ is the 0 polynomial.

Proof. The first identity follows from the fact that ψ is a ring homomorphism. The second statement follows from the fact that $f^\lambda_{\mu,\nu}(0) = c^\lambda_{\mu,\nu}$. For the last statement, note that if $c^\lambda_{\mu,\nu} = 0$, then $G^\lambda_{\mu,\nu}(p) = 0$ for all p by Corollary 7.1.11. $\square$

Remark 7.2.12. We would like to say that $c^\lambda_{\mu,\nu} \ne 0$ if and only if $G^\lambda_{\mu,\nu}(p) \ne 0$ for all primes p, though the above approach leaves the possibility that $g^\lambda_{\mu,\nu}(t)$ could be nonzero but have prime roots. In fact, this does not happen, though I don't see a way to deduce that from what we've already proven. A much stronger property of the Hall polynomials is shown in [Mal]: it is a non-negative linear combination of powers of $t-1$. $\square$

We can now prove what's known as the "semigroup property" of Littlewood–Richardson coefficients:

Corollary 7.2.13. *If $c^\lambda_{\mu,\nu} \ne 0$ and $c^\alpha_{\beta,\gamma} \ne 0$, then $c^{\lambda+\alpha}_{\beta+\mu,\nu+\gamma} \ne 0$.*

Proof. By (3.5.6), $c^{\lambda^\dagger}_{\mu^\dagger,\nu^\dagger} \ne 0$ and $c^{\alpha^\dagger}_{\beta^\dagger,\gamma^\dagger} \ne 0$. Let p be prime which is not a root of $g^{\lambda^\dagger}_{\mu^\dagger,\nu^\dagger}(t)$ nor $g^{\alpha^\dagger}_{\beta^\dagger,\gamma^\dagger}(t)$ (by the above remark, any prime will do, though we don't need this fact). Then there exist abelian p-groups $M \subset L$ and $B \subset A$ such that $M, L, L/M, B, A, A/B$ have types $\mu^\dagger, \lambda^\dagger, \nu^\dagger, \beta^\dagger, \alpha^\dagger, \gamma^\dagger$, respectively. Now $M \times B \subset L \times A$ is a subgroup of type $\mu^\dagger \cup \beta^\dagger$ of a group of type $\lambda^\dagger \cup \alpha^\dagger$ whose quotient has type $\nu^\dagger \cup \gamma^\dagger$. Note that in general, $\delta^\dagger \cup \varepsilon^\dagger = (\delta+\varepsilon)^\dagger$. Hence we've shown that $c^{\lambda+\alpha}_{\mu+\beta,\nu+\gamma} \ne 0$. $\square$

Remark 7.2.14. This property is not clear from the rule we've given via Littlewood–Richardson tableaux. However, it did require a lot of work. If we allow ourselves to use some algebraic geometry, then it is possible to give a conceptually simpler proof using the representation-theoretic interpretation of Littlewood–Richardson coefficients as multiplicities of Schur functors inside of a tensor product of Schur functors.

One consequence of the semigroup property is that $c^\lambda_{\mu,\nu} \ne 0$ implies that $c^{N\lambda}_{N\mu,N\nu} \ne 0$ for any positive integer N. Conversely, one can ask if the Littlewood–Richardson coefficients are "saturated", that is, if $c^{N\lambda}_{N\mu,N\nu} \ne 0$ for some $N > 0$, does it imply that $c^\lambda_{\mu,\nu} \ne 0$? The answer is yes, and is closely related to "Horn's problem". For context, given a Hermitian $n \times n$ matrix A, its eigenvalues are all real, so we can rearrange them into a weakly decreasing sequence. In particular, we can ask for which triples of partitions λ, μ, ν of length at most n, does there exist Hermitian matrices A, B, C such that $A = B + C$ and the spectrum of A, B, C are λ, μ, ν, respectively? Surprisingly enough, the answer is that this is true if and only if $c^\lambda_{\mu,\nu} \neq 0$. See [Fu2] for a survey and further references. □

8. More on Hall–Littlewood functions

Let $\Lambda(t) = \Lambda[t] \otimes_{\mathbf{Z}[t]} \mathbf{Q}(t)$. Define

$$\varphi_r(t) = (1-t)(1-t^2)\cdots(1-t^r),$$

and for a partition λ, set

$$b_\lambda(t) = \prod_{i \geq 1} \varphi_{m_i(\lambda)}(t)$$

and

$$Q_\lambda(x;t) = b_\lambda(t) P_\lambda(x;t).$$

Define $q_0 = 1$; for each $r \geq 1$, we also define

$$q_r(x;t) = Q_r(x;t) = (1-t)P_r(x;t).$$

In finitely many variables, it will be convenient to use the expression

$$g_i = (1-t) \prod_{\substack{1 \leq j \leq n \\ j \neq i}} \frac{x_i - tx_j}{x_i - x_j}.$$

In this case, we have

$$q_r(x_1, \dots, x_n; t) = \sum_{i=1}^n g_i x_i^r = (1-t) \sum_{i=1}^n x_i^r \prod_{j \neq i} \frac{x_i - tx_j}{x_i - x_j} \tag{8.1}$$

where the equality is from the second expression for P_r in Lemma 7.2.2.

We have $q_r(x;0) = h_r(x)$ and $Q_r(x;0) = s_\lambda(x)$. The goal in this section is to prove t-analogues of various identities that we have established for symmetric functions.

8.1. t-analogue of Jacobi–Trudi.

Lemma 8.1.1.

$$\sum_{r \geq 0} q_r(x;t) u^r = \prod_{i \geq 1} \frac{1 - x_i t u}{1 - x_i u}.$$

We define $\underline{Q}(u) = \sum_{r \geq 0} q_r(x;t) u^r$.

Proof. First work in finitely many variables $x_1, \dots, x_n$ and set $z = 1/u$. Applying partial fraction decomposition with respect to the variable z, the right side becomes

$$\begin{aligned}
\prod_{i=1}^n \frac{z - tx_i}{z - x_i} &= 1 + \frac{\prod_{i=1}^n (z - tx_i) - \prod_{i=1}^n (z - x_i)}{\prod_{i=1}^n (z - x_i)} \\
&= 1 + \sum_{i=1}^n \frac{1}{z - x_i} \frac{\prod_{j=1}^n (x_i - tx_j)}{\prod_{j \ne i} (x_i - x_j)} \\
&= 1 + \sum_{i=1}^n \frac{(1-t)x_i}{z - x_i} \prod_{j \ne i} \frac{x_i - tx_j}{x_i - x_j} \\
&= 1 + (1-t) \sum_{i=1}^n \frac{ux_i}{1 - ux_i} \prod_{j \ne i} \frac{x_i - tx_j}{x_i - x_j}.
\end{aligned}$$

The coefficient of u^r for $r > 0$ is

$$(1-t) \sum_{i=1}^n x_i^r \prod_{j \ne i} \frac{x_i - tx_j}{x_i - x_j} = q_r(x_1, \dots, x_n; t),$$

where the equality is (8.1). Now take $n \to \infty$. □

Given a polynomial or formal power series f in variables x, let $f^{(i)}$ be the result of setting $x_i = 0$. The following generalizes the identity (8.1).

Lemma 8.1.2. *Given a partition λ, let $\mu = (\lambda_2, \lambda_3, \dots)$. We have*

$$Q_\lambda(x_1, \dots, x_n; t) = \sum_{i=1}^n x_i^{\lambda_1} g_i Q_\mu^{(i)}(x_1, \dots, x_n; t).$$

Proof. Set $k = \ell(\lambda)$. Then

$$\begin{aligned}
Q_\lambda(x_1, \dots, x_n; t) &= b_\lambda(t) P_\lambda(x_1, \dots, x_n; t) \\
&= \frac{b_\lambda(t)}{v_\lambda(t)} \sum_{\sigma \in \Sigma_n} \sigma \left(x_1^{\lambda_1} \cdots x_k^{\lambda_k} \prod_{1 \le i < j \le n} \frac{x_i - tx_j}{x_i - x_j} \right) \\
&= \frac{b_\lambda(t) v_{n-k}(t)}{v_\lambda(t)} \sum_{\sigma \in \Sigma_n / \Sigma_{n-k}} \sigma \left(x_1^{\lambda_1} \cdots x_k^{\lambda_k} \prod_{i=1}^k \prod_{j > i} \frac{x_i - tx_j}{x_i - x_j} \right)
\end{aligned}$$

where, in the last equality, Σ_{n-k} acts on the last $n-k$ variables, and we used the definition of $v_{n-k}(t)$. By Lemma 7.2.1 and the definition of $b_\lambda(t)$, the leading coefficient can be replaced by $(1-t)^k$. We can decompose the sum above by writing each coset representative σ as a product $\sigma_1\sigma_2$ where σ_1 is fixed coset representative of Σ_n/Σ_{n-1} (where Σ_{n-1} acts on $x_2, \dots, x_n$) and σ_2 is a fixed coset representative of $\Sigma_{n-1}/\Sigma_{n-k}$. Hence

$$\begin{aligned}
Q_\lambda(x_1, \dots, x_n; t) &= \sum_{\sigma \in \Sigma_n / \Sigma_{n-k}} \sigma_1 \left(x_1^{\lambda_1} g_1 \sigma_2 \left(x_2^{\lambda_2} \cdots x_k^{\lambda_k} \prod_{i=2}^k (1-t) \prod_{j > i} \frac{x_i - tx_j}{x_i - x_j} \right) \right) \\
&= \sum_{\sigma \in \Sigma_n / \Sigma_{n-k}} \sigma_1 \left(x_1^{\lambda_1} g_1 \sigma_2 \left(x_2^{\lambda_2} \cdots x_k^{\lambda_k} \frac{b_\mu(t) v_{n-k}(t)}{v_\mu(t)} \prod_{2 \le i < j \le n} \frac{x_i - tx_j}{x_i - x_j} \right) \right).
\end{aligned}$$

Fixing σ_1 and varying over all possibilities for σ_2, the inner expression is $b_\mu(t)P_\mu(x_2,\dots,x_n;t) = Q_\mu^{(1)}(x_1,\dots,x_n;t)$, so we get

$$Q_\lambda(x_1,\dots,x_n;t) = \sum_{\sigma\in\Sigma_n/\Sigma_{n-1}} \sigma\left(x_1^{\lambda_1} g_1 Q_\mu^{(1)}(x_1,\dots,x_n;t)\right) = \sum_{i=1}^n x_i^{\lambda_1} g_i Q_\mu^{(i)}(x_1,\dots,x_n;t). \qquad \square$$

For the next result, define

$$F(u) = \frac{1-u}{1-tu} = 1 + \sum_{r\ge 1}(t^r - t^{r-1})u^r.$$

Proposition 8.1.3. *Define*

$$\underline{Q}(u_1,u_2,\dots) = \prod_{i\ge 1}\underline{Q}(u_i)\prod_{i<j}F(u_i^{-1}u_j).$$

Then $Q_\lambda(x;t)$ is the coefficient of $u_1^{\lambda_1}u_2^{\lambda_2}\cdots$ in $\underline{Q}(u_1,u_2,\dots)$.

Proof. We prove this by induction on $\ell(\lambda)$. When $\ell(\lambda)=1$, we have $Q_r(x;t) = q_r(x;t)$ and the coefficient of u_1^r in $\underline{Q}(u_1,u_2,\dots)$ is the coefficient of u_1^r in $\underline{Q}(u_1)$, and so this follows from Lemma 8.1.1.

Now assume we have shown it for all partitions with length strictly smaller than λ. Set $\mu = (\lambda_2,\lambda_3,\dots)$, so that the result is true for μ. Again, by Lemma 8.1.1, we have

$$\underline{Q}^{(i)}(u) = F(x_iu)\underline{Q}(u),$$

and this implies

$$\underline{Q}^{(i)}(u_1,u_2,\dots) = \underline{Q}(u_1,u_2,\dots)\prod_{j\ge 1}F(x_iu_j).$$

By our induction hypothesis, $Q_\mu(x_1,\dots,x_n;t)$ is the coefficient of $u_2^{\lambda_2}u_3^{\lambda_3}\cdots$ in $\underline{Q}(u_2,u_3,\dots)$. Hence the same is true for $Q_\mu^{(i)}$ and $\underline{Q}^{(i)}$. Now by Lemma 8.1.2, $Q_\lambda(x_1,\dots,x_n;t)$ is the coefficient of $u_1^{\lambda_1}u_2^{\lambda_2}\cdots$ in

$$\sum_{\lambda_1\ge 0} u_1^{\lambda_1}\sum_{i=1}^n x_i^{\lambda_1} g_i \underline{Q}^{(i)}(u_2,u_3,\dots) = \underline{Q}(u_2,u_3,\dots)\sum_{\lambda_1\ge 0}u_1^{\lambda_1}\sum_{i=1}^n x_i^{\lambda_1}g_i\prod_{j\ge 2}F(x_iu_j)$$

where $g_i = (1-t)\prod_{\substack{1\le j\le n\\ j\ne i}}\frac{x_i - tx_j}{x_i - x_j}$. Write $\prod_{j\ge 2}F(x_iu_j) = \sum_{m\ge 0} f_m x_i^m$ where f_m is a polynomial in $t, u_2, u_3, \dots$. Then we have

$$\begin{aligned}
\sum_{\lambda_1\ge 0} u_1^{\lambda_1}\sum_{i=1}^n x_i^{\lambda_1} g_i \prod_{j\ge 2}F(x_iu_j) &= \sum_{\lambda_1,m\ge 0} u_1^{\lambda_1}\sum_{i=1}^n x_i^{\lambda_1+m} g_i f_m \\
&\overset{(8.1)}{=} \sum_{\lambda_1,m\ge 0} u_1^{\lambda_1} f_m q_{\lambda_1+m}(x_1,\dots,x_n;t) \\
&= \underline{Q}(u_1)\sum_{m\ge 0} f_m u_1^{-m} \\
&= \underline{Q}(u_1)\prod_{j\ge 2}F(u_1^{-1}u_j).
\end{aligned}$$

Combining everything, we conclude that $Q_\lambda(x_1,\dots,x_n;t)$ is the coefficient of $u_1^{\lambda_1}u_2^{\lambda_2}\cdots$ in

$$\underline{Q}(u_2,u_3,\dots)\underline{Q}(u_1)\prod_{j\ge 2}F(u_1^{-1}u_j)=\underline{Q}(u_1,u_2,\dots).$$

Now take $n\to\infty$. $\square$

At this point, it is convenient to introduce the formalism of raising operators so that we can reinterpret the above result in a more compact way. Given an integer sequence $\alpha=(\alpha_1,\alpha_2,\dots)$, and $i<j$, define

$$R_{ij}\alpha=\alpha+\varepsilon_i-\varepsilon_j=(\alpha_1,\dots,\alpha_i+1,\dots,\alpha_j-1,\dots)$$

A **raising operator** is any polynomial R in the R_{ij}.

For any integer sequence $\alpha=(\alpha_1,\alpha_2,\dots)$, set

$$q_\alpha(x;t)=q_{\alpha_1}(x;t)q_{\alpha_2}(x;t)\cdots=Q_{\alpha_1}(x;t)Q_{\alpha_2}(x;t)\cdots$$

with the convention that $q_\alpha=0$ if any part is negative.

Given a raising operator R, we define[8]

$$Rq_\alpha(x;t)=q_{R\alpha}(x;t).$$

Corollary 8.1.4. *We have*

$$Q_\lambda(x;t)=\left(\prod_{i<j}\frac{1-R_{ij}}{1-tR_{ij}}\right)q_\lambda(x;t).$$

Proof. The coefficient of u^α in

$$\prod_{i<j}F(u_i^{-1}u_j)\prod_{k\ge 1}\underline{Q}(u_k)=\prod_{i<j}\left(1+\sum_{r\ge 1}(t^r-t^{r-1})u_i^{-r}u_j\right)\sum_\beta q_\beta u^\beta$$

is

$$\prod_{i<j}\left(1+\sum_{r\ge 1}(t^r-t^{r-1})R_{ij}^r\right)q_\alpha=\left(\prod_{i<j}\frac{1-R_{ij}}{1-tR_{ij}}\right)q_\alpha.$$

Now use the previous result with $\alpha=\lambda$. $\square$

Remark 8.1.5. Recall that the Jacobi–Trudi identity (Theorem 3.6.1) says

$$s_\lambda=\det(h_{\lambda_i-i+j})_{i,j=1}^n$$

for any $n\ge \ell(\lambda)$. Set $\rho=(n-1,n-2,\dots,1,0)$. If we expand the determinant, we get

$$\sum_{\sigma\in\Sigma_n}\operatorname{sgn}(\sigma)h_{\lambda+\rho-\sigma(\rho)}=\left(\prod_{1\le i<j\le n}(1-R_{ij})\right)h_\lambda.$$

[8]**Warning:** this is not an action! For example, $R_{23}R_{12}=R_{13}$, and $R_{13}q_{0,0,1}=q_1$ while $R_{23}(R_{12}q_{0,0,1})=0$. When these are used, we will typically multiply out all of the polynomials, and then apply them to q_α.

To see this, first consider the following identity in $\mathbf{Z}[x_1^{\pm 1}, \dots, x_n^{\pm 1}]$:

$$\begin{aligned}\sum_{\sigma \in \Sigma_n} \mathrm{sgn}(\sigma) x^{\lambda+\rho-\sigma(\rho)} &= x^{\lambda+\rho} a_{-\rho} \\ &= x^{\lambda+\rho} \prod_{1 \le i < j \le n} (x_i^{-1} - x_j^{-1}) \\ &= x^{\lambda} \prod_{1 \le i < j \le n} (1 - x_i x_j^{-1}) \\ &= \left(\prod_{1 \le i < j \le n} (1 - R_{ij}) \right) x^{\lambda}\end{aligned}$$

where for a raising operator, we define $R x^{\lambda} := x^{R\lambda}$. Now apply the $\mathbf{Z}$-linear homomorphism $\mathbf{Z}[x_1^{\pm 1}, \dots, x_n^{\pm 1}] \to \Lambda$ given by $x^{\alpha} \mapsto h_{\alpha}$. This is also the result of setting $t = 0$ in the previous identity, so we can think of it as a t-deformation of the Jacobi–Trudi identity. □

Theorem 8.1.6. *There exist $a_{\lambda,\mu}(t) \in \mathbf{Z}[t]$ such that $a_{\lambda,\lambda}(t) = 1$ and $a_{\lambda,\mu}(t) \neq 0$ implies that $\lambda \le \mu$ such that*

$$Q_{\lambda}(x;t) = \sum_{\mu} a_{\lambda,\mu}(t) q_{\mu}(x;t).$$

Furthermore, $a_{\lambda,\mu}(t)$ is divisible by $(t-1)^{\ell(\lambda)-\ell(\mu)}$. In particular, $\{q_{\lambda}(x;t) \mid \lambda \in \mathrm{Par}\}$ is a $\mathbf{Q}(t)$-basis for $\Lambda(t)$.

Proof. From Corollary 8.1.4, we have

$$Q_{\lambda}(x;t) = \left(\prod_{i<j} (1 + (t-1)R_{ij} + (t^2 - t)R_{ij}^2 + \cdots) \right) q_{\lambda}(x;t)$$

which implies that $Q_{\lambda}(x;t) - q_{\lambda}(x;t)$ is a $\mathbf{Z}[t]$-linear combination of $q_{\mu}(x;t)$ with $\mu \ge \lambda$. Furthermore, the coefficient of $q_{\mu}(x;t)$ is divisible by one power of $(1-t)$ for each different nontrivial R_{ij} component used in the product above, and at least $\ell(\lambda) - \ell(\mu)$ of them have to be used, so we get the divisibility statement.

By Corollary 7.2.5, $P_{\lambda}(x;t)$ is a $\mathbf{Z}[t]$-basis for $\Lambda[t]$, and hence $Q_{\lambda}(x;t)$ is a $\mathbf{Q}(t)$-basis for $\Lambda(t)$ by their definition. We just showed that the change of basis matrix between Q and q is invertible over $\mathbf{Z}[t]$, so we get the desired result. □

8.2. **t-analogue of the inner product.** We define a $\mathbf{Q}(t)$-valued bilinear form on $\Lambda(t)$ by

$$\langle q_{\lambda}(x;t), m_{\mu}(x) \rangle_t = \delta_{\lambda,\mu}$$

and extending it to be $\mathbf{Q}(t)$-linear in each variable.

Proposition 8.2.1. *We have*

$$\sum_{\lambda} q_{\lambda}(x;t) m_{\lambda}(y) = \prod_{i,j} \frac{1 - t x_i y_j}{1 - x_i y_j} = \sum_{\lambda} m_{\lambda}(x) q_{\lambda}(y;t).$$

Furthermore, if $\{u_{\lambda}\}$ and $\{v_{\mu}\}$ are $\mathbf{Q}(t)$-linear bases of $\Lambda(t)$, then

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

if and only if $\langle u_{\lambda}, v_{\lambda} \rangle_t = \delta_{\lambda,\mu}$.

Proof. Use Lemma 8.1.1 with $u = y_j$ to get

$$\prod_{i,j} \frac{1 - tx_iy_j}{1 - x_iy_j} = \prod_j \sum_{r \geq 0} q_r(x;t) y_j^r = \sum_\lambda q_\lambda(x;t) m_\lambda(y).$$

Similarly, we can use Lemma 8.1.1 with $u = x_i$ to get

$$\prod_{i,j} \frac{1 - tx_iy_j}{1 - x_iy_j} = \prod_i \sum_{r \geq 0} q_r(y;t) x_i^r = \sum_\lambda q_\lambda(y;t) m_\lambda(x).$$

The second part is formally the same as Lemma 2.6.1, so we omit the details. ☐

Corollary 8.2.2. *The bilinear form $\langle , \rangle_t$ is symmetric, i.e., $\langle f, g \rangle_t = \langle g, f \rangle_t$ for all $f, g \in \Lambda(t)$.*

Proposition 8.2.3. *We have*

$$\sum_\lambda P_\lambda(x;t) Q_\lambda(y;t) = \prod_{i,j} \frac{1 - tx_iy_j}{1 - x_iy_j}.$$

In particular,

$$\langle P_\lambda(x;t), Q_\mu(x;t) \rangle_t = \delta_{\lambda,\mu}.$$

Proof. Given two bases u_λ and v_μ, write $M(u,v)$ for the change of basis matrix that records the coefficients of expanding u_λ in terms of the v_μ. Also, order the entries by some fixed total ordering that extends dominance order. Then $M(u,v)^{-1} = M(v,u)$ and if one is upper (resp., lower) triangular, then the same is true for the other.

Let $D = M(m,Q)^T M(q,Q)$. Note that $M(q,Q) = M(q,m)M(m,Q)$ and that $M(q,m)$ is a symmetric matrix by Proposition 8.2.1, and so the same is true for D. Furthermore, $M(Q,q)$, and hence $M(q,Q)$, is lower unitriangular by Theorem 8.1.6. We claim that $M(m,Q)^T$ is also lower triangular: this follows since

$$M(Q,m) = M(Q,P)M(P,s)M(s,m),$$

and $M(Q,P)$ is diagonal by definition, $M(P,s)$ is upper triangular by Proposition 7.2.4, and $M(s,m)$ is upper triangular by Theorem 3.1.5. In conclusion, D is symmetric and lower-triangular, so it is actually a diagonal matrix. It shares its diagonal entries with $M(P,Q)$ since all of the other change of basis matrices involved have 1's on the diagonal. This implies that

$$\sum_\lambda M(m,Q)_{\lambda,\mu} M(q,Q)_{\lambda,\nu} = b_\mu(t)^{-1} \delta_{\mu,\nu}.$$

Now we can finish:

$$\begin{aligned} \sum_\lambda q_\lambda(x;t) m_\lambda(y) &= \sum_{\lambda,\mu,\nu} M(m,Q)_{\lambda,\mu} M(q,Q)_{\lambda,\nu} Q_\mu(x;t) Q_\nu(y;t) \\ &= \sum_\mu b_\mu(t)^{-1} Q_\mu(x;t) Q_\mu(y;t) \\ &= \sum_\mu P_\mu(x;t) Q_\mu(y;t) = \prod_{i,j} \frac{1 - tx_iy_j}{1 - x_iy_j}. \end{aligned}$$

where the last equality is Proposition 8.2.1. ☐

Define

$$z_\lambda(t) = z_\lambda \prod_{i\ge 1} (1-t^{\lambda_i})^{-1} = \prod_{i \ge 1} \frac{m_i(\lambda)! i^{m_i(\lambda)}}{1-t^{\lambda_i}}.$$

Proposition 8.2.4. *We have*

$$\sum_\lambda z_\lambda(t)^{-1} p_\lambda(x) p_\lambda(y) = \prod_{i,j} \frac{1-tx_iy_j}{1-x_iy_j}.$$

In particular,

$$\langle p_\lambda(x), p_\mu(x)\rangle_t = z_\lambda(t)\delta_{\lambda,\mu}.$$

Proof. Start with

$$\begin{aligned}
\log \prod_{i,j} \frac{1-tx_iy_j}{1-x_iy_j} &= \sum_{i,j} (\log(1-tx_iy_j) - \log(1-x_iy_j)) \\
&= \sum_{i,j}\sum_{r\ge 1} \left(-\frac{(tx_iy_j)^r}{r} + \frac{(x_iy_j)^r}{r}\right) = \sum_{i,j}\sum_{r\ge 1} \frac{1-t^r}{r}(x_iy_j)^r \\
&= \sum_{r \ge 1} \frac{1-t^r}{r} p_r(x)p_r(y).
\end{aligned}$$

Now exponentiate both sides to get

$$\begin{aligned}
\prod_{i,j} \frac{1-tx_iy_j}{1-x_iy_j} &= \prod_{r\ge 1} \exp\left(\frac{1-t^r}{r} p_r(x)p_r(y)\right) \\
&= \prod_{r\ge 1}\sum_{d_r \ge 0} \left(\frac{(1-t^r)^{d_r}}{d_r! r^{d_r}} p_r(x)^{d_r} p_r(y)^{d_r}\right) \\
&= \sum_\lambda z_\lambda(t)^{-1} p_\lambda(x)p_\lambda(y).
\end{aligned}$$

∎

9. Schur Q-functions

9.1. **Hall–Littlewood functions at $t=-1$.** We now set $t=-1$ in the functions from the previous section and use the shorthand:

$$q_r = q_r(x;-1), \qquad q_\lambda = q_\lambda(x;-1), \qquad P_\lambda = P_\lambda(x;-1), \qquad Q_\lambda = Q_\lambda(x;-1).$$

Also define the generating function

$$Q(u) = \sum_{r \ge 0} q_r u^r.$$

Specializing $t=-1$ in Lemma 8.1.1, we get

$$Q(u) = \prod_{i \ge 1} \frac{1+x_iu}{1-x_iu}.$$

This implies in particular that $Q(u)Q(-u) = 1$, which gives the identity

$$\sum_{i=0}^n (-1)^i q_i q_{n-i} = 0$$

for $n \geq 1$. When $n = 2m$ is even, this implies

$$q_{2m} = \sum_{r=1}^{m-1} (-1)^{r-1} q_r q_{2m-r} + \frac{(-1)^{m-1}}{2} q_m^2 \tag{9.1.1}$$

which shows that $q_{2m} \in \mathbf{Q}[q_1, q_2, \dots, q_{2m-1}]$. By induction on m, we can simplify this to $q_{2m} \in \mathbf{Q}[q_1, q_3, q_5, \dots, q_{2m-1}]$.

Lemma 9.1.2. *We have $\mathbf{Q}[p_r \mid r \text{ odd}] = \mathbf{Q}[q_r \mid r \text{ odd}]$. Also, the q_r with r odd are algebraically independent.*

Proof. Recall $E(t) = \prod_{i \geq 1}(1 + x_i t)$ and $H(t) = \prod_{i \geq 1}(1 - x_i t)^{-1}$ and $P(t) = \sum_{n \geq 1} p_n t^{n-1}$ from §2.5 and that $H'(t)/H(t) = P(t)$ and $E'(t)/E(t) = P(-t)$. Then from our identity for $Q(u)$ above, we have $Q(u) = H(u)E(u)$, and hence

$$\frac{Q'(u)}{Q(u)} = \frac{H'(u)E(u) + H(u)E'(u)}{H(u)E(u)} = \frac{H'(u)}{H(u)} + \frac{E'(u)}{E(u)} = P(u) + P(-u) = 2\sum_{r \geq 0} p_{2r+1} t^{2r}.$$

Multiply both sides by $Q(u)$ and take the coefficient of u^{r-1} to get

$$r q_r = 2(p_1 q_{r-1} + p_3 q_{r-3} + p_5 q_{r-5} + \cdots),$$

where the sum ends with $p_{r-1} q_1$ if r is even and with p_r if r is odd. Hence $q_1 = 2p_1$ and by induction on r, we see that q_r is in the $\mathbf{Q}$-subalgebra generated by the odd p_i and also that p_r with r odd is in the $\mathbf{Q}$-subalgebra generated by the q_i. $\square$

We let Γ be the subring of Λ generated by $q_1, q_2, \dots$ and let $\Gamma_{\mathbf{Q}}$ be the $\mathbf{Q}$-subalgebra of $\Lambda_{\mathbf{Q}}$ generated by $q_1, q_2, \dots$. The previous result gives:

$$\Gamma_{\mathbf{Q}} = \mathbf{Q}[q_1, q_3, q_5, \dots] = \mathbf{Q}[p_1, p_3, p_5, \dots].$$

Let $\mathrm{OPar}(n)$ be the set of partitions of n such that all parts are odd, and let $\mathrm{DPar}(n)$ be the set of the partitions of n such that all parts are distinct. Set $\mathrm{OPar} = \bigcup_{n \geq 0} \mathrm{OPar}(n)$ and $\mathrm{DPar} = \bigcup_{n \geq 0} \mathrm{DPar}(n)$.

Lemma 9.1.3.

$$q_n = \sum_{\lambda \in \mathrm{OPar}(n)} z_\lambda^{-1} 2^{\ell(\lambda)} p_\lambda.$$

Proof. Earlier, we obtained the identity

$$\sum_{n \geq 0} q_n u^n = \prod_j \frac{1 + x_j u}{1 - x_j u}$$

by specializing $t = -1$ in Lemma 8.1.1. Now specialize $t = -1$, $y_1 = u$, and $y_i = 0$ for $i > 1$ in Proposition 8.2.4 to get

$$\sum_{\lambda \in \mathrm{OPar}} z_\lambda^{-1} 2^{\ell(\lambda)} p_\lambda(x) u^{|\lambda|} = \prod_j \frac{1 + x_j u}{1 - x_j u}.$$

Comparing these two identities and taking the coefficient of u^n gives the desired identity. $\square$

Lemma 9.1.4 (Euler)**.** $|\mathrm{OPar}(n)| = |\mathrm{DPar}(n)|$.

Proof. We have

$$\sum_{n\geq 0}|\mathrm{DPar}(n)|t^n = \prod_{i\geq 1}(1+t^i) = \prod_{i\geq 1}\frac{1-t^{2i}}{1-t^i} = \prod_{i\geq 1}\frac{1}{1-t^{2i+1}} = \sum_{n\geq 0}|\mathrm{OPar}(n)|t^n.$$

For the third equality, we cancel each numerator $1-t^{2i}$ with a corresponding denominator, and observe that the only remaining terms are what's written. ☐

Proposition 9.1.5. (a) $\{q_\lambda \mid \lambda \in \mathrm{OPar}\}$ *is a* $\mathbf{Q}$*-basis for* $\Gamma_{\mathbf{Q}}$.

(b) $\{q_\lambda \mid \lambda \in \mathrm{DPar}\}$ *is a* $\mathbf{Z}$*-basis for* Γ*. Furthermore, each* q_μ *can be written as a* $\mathbf{Z}$*-linear combination of* q_λ *with* λ *strict and* $\lambda \geq \mu$ *such that the coefficient of* q_μ *is divisible by* $2^{\ell(\lambda)-\ell(\mu)}$.

Proof. (a) was already shown in Lemma 9.1.2.

(b) We prove by descending induction on dominance order for all partitions of n that either μ is strict, or else q_μ is a $\mathbf{Z}$-linear combination of q_λ for λ strict with the required divisibility condition. The base case is $\mu = (n)$, which is strict. Now assume that the induction hypothesis holds for all $\nu > \mu$. If μ is not strict, pick i such that $\mu_i = \mu_{i+1} = m$. Use (9.1.1) to rewrite q_μ as a sum of $q_{\nu^1}, \dots$ where the copies of m in μ_i, μ_{i+1} have been replaced by $\{0, 2m\}, \{1, 2m-1\}, \{2, 2m-2\}, \dots, \{m-1, m+1\}$. The coefficient of each q_{ν^i} is 2 and $\ell(\nu^i) \geq \ell(\mu) - 1$. Each of these partitions is larger than μ in the dominance order, so by induction, they are also $\mathbf{Z}$-linear combinations of q_λ with λ strict with the required divisiblity condition. Combining with the extra factor of 2 we picked up, the divisiblity condition is preserved. Hence $\{q_\lambda \mid \lambda \in \mathrm{DPar}(n)\}$ spans Γ_n. (a) shows that Γ_n has rank $|\mathrm{OPar}(n)|$, which is the same as $|\mathrm{DPar}(n)|$ by Lemma 9.1.4. ☐

Proposition 9.1.6. $Q_\lambda \in \Gamma$*. Furthermore, if* $\lambda \in \mathrm{DPar}$*, then*

$$Q_\lambda = \sum_{\mu \in \mathrm{DPar}} a_{\lambda,\mu} q_\mu$$

where $a_{\lambda,\mu} \in \mathbf{Z}$ *and* $a_{\lambda,\mu} \neq 0$ *implies that* $\mu \geq \lambda$*. Furthermore,* $a_{\lambda,\lambda} = 1$ *and* $a_{\lambda,\mu}$ *is divisible by* $2^{\ell(\lambda)-\ell(\mu)}$*. In particular,* $\{Q_\lambda \mid \lambda \in \mathrm{DPar}\}$ *is a* $\mathbf{Z}$*-basis of* Γ.

Proof. First, set $t = -1$ in Theorem 8.1.6 to get

$$Q_\lambda = \sum_\mu a_{\lambda,\mu}(-1) q_\mu$$

where $a_{\lambda,\lambda}(-1) = 1$ and $a_{\lambda,\mu}(-1) \neq 0$ implies $\mu \geq \lambda$. Since each $a_{\lambda,\mu}(t)$ with $\lambda \neq \mu$ is divisible by $1-t$, we conclude that $a_{\lambda,\mu}(-1)$ is even. If λ has all distinct parts, then we can rewrite each q_μ with $\mu > \lambda$ as a $\mathbf{Z}$-linear combination of q_ν with $\nu \geq \mu$ by Proposition 9.1.5. ☐

The bilinear form $\langle , \rangle_t$ has poles at $t = -1$ (e.g., $z_\lambda(t)$ whenever λ has an even part), but this issue disappears if we restrict it to $\Gamma_{\mathbf{Q}}$. In that case, we get:

Proposition 9.1.7. (a) $\langle Q_\lambda, Q_\mu \rangle = 2^{\ell(\lambda)} \delta_{\lambda,\mu}$.

(b) *If* $\lambda, \mu \in \mathrm{OPar}$*, then* $\langle p_\lambda, p_\mu \rangle = z_\lambda \delta_{\lambda,\mu} / 2^{\ell(\lambda)}$.

Proof. From Proposition 8.2.3, we have $\langle Q_\lambda, Q_\mu \rangle_t = b_\lambda(t) \delta_{\lambda,\mu}$. From its definition, $b_\lambda(-1) = 2^{\ell(\lambda)}$. From Proposition 8.2.4, we have $\langle p_\lambda, p_\mu \rangle_t = z_\lambda(t) \delta_{\lambda,\mu}$. Again, from its definition, $z_\lambda(-1) = z_\lambda / 2^{\ell(\lambda)}$. ☐

9.2. **Projective representations.** We consider now the problem of classifying "projective representations" of the symmetric group. We've already discussed (linear) representations. Recall that this can be thought of as either a group homomorphism

$$\rho\colon G \to \mathbf{GL}(V)$$

for some vector space V, or equivalently, as a linear group action of G on V. Let $\mathbf{C}^* \subset \mathbf{GL}(V)$ denote the subgroup of scalar matrices. This is a normal subgroup, and the quotient is the **projective linear group**

$$\mathbf{PGL}(V) = \mathbf{GL}(V)/\mathbf{C}^*.$$

A **projective representation** of a group is a group homomorphism

$$\rho\colon G \to \mathbf{PGL}(V).$$

Note that by composing with the quotient map $\mathbf{GL}(V) \to \mathbf{PGL}(V)$, every linear representation gives a projective representation. Naturally, one can ask if the converse is true: does every projective representation come from a linear representation, i.e., can we linearize it? To give a sense for what this means, suppose ρ is a projective representation of G. Then arbitrarily choose coset representatives $\rho'(g) \in \mathbf{GL}(V)$ for $\rho(g)$ for each g. Then for every $g, h \in G$, there is a scalar $\alpha(g,h) \in \mathbf{C}^*$ such that

$$\rho'(g)\rho'(h) = \alpha(g,h)\rho'(gh),$$

and we are asking if there is a choice of ρ' such that $\alpha(g,h) = 1$ for all g, h.

Remark 9.2.1. Since $\rho'(f)(\rho'(g)\rho'(h)) = (\rho'(f)\rho'(g))\rho'(h)$ for any $f, g, h \in G$, we have

$$\alpha(f, gh)\alpha(g,h) = \alpha(f,g)\alpha(fg,h).$$

Also, $\alpha(1,g) = 1 = \alpha(g,1)$ for all $g \in G$. This means that α is a 2-cocycle on G in the sense of group cohomology. Note that for any function $\delta\colon G \to \mathbf{C}^*$, we could have modified the choice of $\rho'(g)$ by replacing it with $\delta(g)\rho'(g)$ and this would replace our 2-cocycle by another one defined by

$$\beta(g,h) = \delta(g)\delta(h)\delta(gh)^{-1}\alpha(g,h).$$

These cocycles are cohomologous, and in fact, ρ having a linearization is equivalent to α being cohomologous to 0, i.e., defining a trivial cohomology class (in $\mathrm{H}^2(G; \mathbf{C}^*)$). $\square$

In general, it is not possible to find such a choice of ρ', and the previous remark makes the obstruction precise. However, Schur proved that there exists a central extension

$$1 \to \mathrm{H}^2(G; \mathbf{C}^*) \to \widetilde{G} \to G \to 1$$

such that any projective representation of G can be lifted to a linear representation of $\widetilde{G}$, i.e., the dashed arrow in the following diagram can always be filled in:

$$\begin{tikzcd} \widetilde{G} \arrow[r, dashed] \arrow[d] & \mathbf{GL}(V) \arrow[d] \\ G \arrow[r, "\rho"] & \mathbf{PGL}(V) \end{tikzcd}.$$

The group $\widetilde{G}$ is called the **representation group** of G. We will not discuss its construction, however, we note that $\mathrm{H}^2(G; \mathbf{C}^*)$ is always a finite group.

We are concerned with the case $G = \Sigma_n$. Then

$$\mathrm{H}^2(\Sigma_n; \mathbf{C}^*) \cong \begin{cases} 1 & \text{if } n \le 3 \\ \mathbf{Z}/2 & \text{if } n \ge 4 \end{cases}.$$

The group $\widetilde{\Sigma}_n$ can be given the following presentation. It has generators $z, t_1, \dots, t_{n-1}$ subject to the relations:

$$\begin{aligned} z^2 &= 1 & & \\ zt_j &= t_j z & & 1 \le j \le n-1 \\ t_j^2 &= z & & 1 \le j \le n-1 \\ (t_j t_{j+1})^3 &= z & & 1 \le j \le n-2 \\ t_j t_k &= z t_k t_j & & \text{if } |j-k| > 1. \end{aligned}$$

The third relation implies that z is redundant as a generator, but it will be useful to have a name for this element. If we set $z = 1$ in the above relations, then one gets the usual presentation for Σ_n (where t_j is the transposition that exchanges j and $j+1$). In particular, $Z = \{1, z\}$ is a normal subgroup of $\widetilde{\Sigma}_n$ and we have a central extension

$$1 \to Z \to \widetilde{\Sigma}_n \to \Sigma_n \to 1.$$

So this is a finite group of order $2n!$ and is a representation group for Σ_n for $n \ge 4$. We let $\pi \colon \widetilde{\Sigma}_n \to \Sigma_n$ be the quotient map.

Lemma 9.2.2. *Let λ be a partition of n and let c_λ be the set of permutations with cycle type λ.*

- *If λ has no even parts (and hence c_λ contains even permutations), then $\pi^{-1}(c_\lambda)$ is a union of two conjugacy classes.*
- *If c_λ contains odd permutations (i.e., $n - \ell(\lambda)$ is odd) and every part of λ is different (i.e., $m_i(\lambda) \le 1$ for all i), then $\pi^{-1}(c_\lambda)$ is a union of two conjugacy classes.*
- *In all other cases, $\pi^{-1}(c_\lambda)$ is a single conjugacy class.*

We will omit the proof; see [HH, Theorem 3.8] for a proof.

Recall that $\mathrm{OPar}(n)$ is the set of partitions of n such that all parts are odd numbers and that $\mathrm{DPar}(n)$ is the set of partitions of n such that all parts are distinct. Furthermore, let $\mathrm{DPar}^+(n)$ denote the subset of $\mathrm{DPar}(n)$ consisting of partitions such that $n - \ell(\lambda)$ is even, and similarly, let $\mathrm{DPar}^-(n)$ be the subset of partitions such that $n - \ell(\lambda)$ is odd.

We need a way to label the conjugacy classes in $\widetilde{\Sigma}_n$. Given a partition λ of n, define

$$\begin{aligned} T_j &= t_{a+1} t_{a+2} \cdots t_{a+\lambda_j - 1} \qquad (a = \lambda_1 + \cdots + \lambda_{j-1}), \\ \sigma^\lambda &= T_1 T_2 \cdots T_{\ell(\lambda)}. \end{aligned}$$

Then $\pi(\sigma^\lambda) \in c_\lambda$. In the case that $\pi^{-1}(c_\lambda)$ is two conjugacy classes, σ^λ and $z\sigma^\lambda$ are representatives for the two conjugacy classes. We define

$$\begin{aligned} c_\lambda^+ &= \{\tau \sigma^\lambda \tau^{-1} \mid \tau \in \widetilde{\Sigma}_n\}, \\ c_\lambda^- &= \{\tau z \sigma^\lambda \tau^{-1} \mid \tau \in \widetilde{\Sigma}_n\}. \end{aligned}$$

The $\pm$ convention has been chosen so that the trace of c_λ^+ in a certain representation, to be constructed in the next section, is positive.

At this point, we could forget all about projective representations and focus on the problem of classifying irreducible representations of $\widetilde{\Sigma}_n$. If V is an irreducible representation of $\widetilde{\Sigma}_n$, then since z is a central element, it must act by a scalar on V by Schur's lemma. But it has order 2, so it either acts as 1 or -1. In the first case, V can be thought of as a representation of Σ_n, and we have already studied this case. Hence, we'll just be interested in the case when z acts by -1, which we will call the **negative representations**. Alternatively, we study modules over the **twisted group algebra**

$$\mathbf{C}[\widetilde{\Sigma}_n]/(z+1).$$

Proposition 9.2.3. *The number of irreducible negative representations of* $\widetilde{\Sigma}_n$ *is* $|\mathrm{OPar}(n)| + |\mathrm{DPar}^-(n)|$.

Proof. From Lemma 9.2.2, the number of conjugacy classes of $\widetilde{\Sigma}_n$ minus the number of conjugacy classes of Σ_n is $|\mathrm{OPar}(n) \cup \mathrm{DPar}^-(n)|$, so this is the number of irreducible negative representations. To finish, we observe that $\mathrm{OPar}(n) \cap \mathrm{DPar}^-(n) = \varnothing$: if $\lambda \in \mathrm{OPar}(n)$, then $n - \ell(\lambda) = \sum_i (\lambda_i - 1)$ is even, so $\lambda \notin \mathrm{DPar}^-(n)$. □

9.3. **Clifford algebras and the basic spin representation.** The trivial representation provides a basic example of a linear representation of the symmetric group and was used to essentially get the characters of all of the others in our approach using the Frobenius characteristic map. This is not a negative representation of $\widetilde{\Sigma}_n$, and hence we need a replacement for it. This will be done using Clifford algebras.

Given a positive integer n, let $\mathbf{Cl}_n$ be the associative unital $\mathbf{C}$-algebra with generators $\xi_1, \dots, \xi_n$ and relations

$$\xi_i^2 = 1, \qquad \xi_i \xi_j = -\xi_j \xi_i \quad (i \neq j).$$

Given an increasing sequence $A = a_1 < \cdots < a_k$ of integers between 1 and n, define

$$\xi_A = \xi_{a_1} \xi_{a_2} \cdots \xi_{a_k}$$

(we allow $k = 0$, in which case $\xi_A = 1$).

Let $M_2(\mathbf{C})$ be the algebra of 2×2 complex matrices. Let $i = \sqrt{-1}$, and define the following elements in $M_2(\mathbf{C})$:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \varepsilon = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad y = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}.$$

They form a linear basis for $M_2(\mathbf{C})$.

Below, we will consider the k-fold tensor product $M_2(\mathbf{C})^{\otimes k}$. Elements are linear combinations of $a_1 \otimes \cdots \otimes a_k$ with $a_i \in M_2(\mathbf{C})$ and multiplication is defined componentwise: $(a_1 \otimes \cdots \otimes a_k)(b_1 \otimes \cdots \otimes b_k) = a_1 b_1 \otimes \cdots \otimes a_k b_k$.

Proposition 9.3.1. *If $n = 2k$ is even, then there is an algebra isomorphism*

$$\rho \colon \mathbf{Cl}_{2k} \to M_2(\mathbf{C})^{\otimes k}$$

defined by

$$\rho(\xi_{2j-1}) = \varepsilon^{\otimes (j-1)} \otimes x \otimes 1^{(k-j)}, \qquad \rho(\xi_{2j}) = \varepsilon^{\otimes (j-1)} \otimes y \otimes 1^{(k-j)}.$$

Furthermore, $\{\rho(\xi_A)\}$ forms a basis for $M_2(\mathbf{C})^{\otimes k}$ and $\{\xi_A\}$ forms a basis for $\mathbf{Cl}_{2k}$ as A ranges over all increasing sequences.

Proof. First, we need to show that ρ is well-defined: this means showing that $\rho(\xi_i)^2 = 1$ and $\rho(\xi_i)\rho(\xi_j) = -\rho(\xi_j)\rho(\xi_i)$ for $i \neq j$. The first follows from the fact that $\varepsilon^2 = x^2 = y^2 = I_2$. The second follows from the fact that $x\varepsilon = -\varepsilon x$, $y\varepsilon = -\varepsilon y$, and $xy = -yx$.

We leave the claim that $\rho(\xi_A)$ forms a basis as A ranges over all possible increasing sequences as an exercise.

Finally, it is clear from the relations that ξ_A span $\mathbf{Cl}_{2k}$. Since their images under ρ are linearly independent, they themselves must be linearly independent. ☐

We now use this to understand $\mathbf{Cl}_{2k+1}$. Let ρ' denote the map $\mathbf{Cl}_{2k} \to M_2(\mathbf{C})^{\otimes k}$ just defined and we think of $\mathbf{Cl}_{2k}$ as a subalgebra of $\mathbf{Cl}_{2k+1}$.

Proposition 9.3.2. *If $n = 2k+1$ is odd, then there is an algebra isomorphism*

$$\rho\colon \mathbf{Cl}_{2k+1} \to M_2(\mathbf{C})^{\otimes k} \oplus M_2(\mathbf{C})^{\otimes k}$$

defined by

$$\rho(\xi_j) = (\rho'(\xi_j), \rho'(\xi_j)) \quad (1 \le j \le 2k),$$
$$\rho(\xi_{2k+1}) = (i^k \rho'(\xi_{2k}) \cdots \rho'(\xi_2)\rho'(\xi_1), -i^k \rho'(\xi_{2k}) \cdots \rho'(\xi_2)\rho'(\xi_1)).$$

Furthermore, $\{\rho(\xi_A)\}$ forms a basis for $M_2(\mathbf{C})^{\otimes k} \oplus M_2(\mathbf{C})^{\otimes k}$ and $\{\xi_A\}$ forms a basis for $\mathbf{Cl}_{2k+1}$ as A ranges over all increasing sequences.

Proof. First, we check this is well-defined: $(\rho'(\xi_{2k}) \cdots \rho'(\xi_1))^2 = (-1)^s$ where $s = (2k-1) + (2k-2) + \cdots + 1 = k(2k-1)$, which is the same as $(-1)^k$. Hence $\rho(\xi_{2k+1})^2 = 1$. Also, we have $\rho(\xi_j)\rho(\xi_{2k+1}) = (-1)^{2k-1}\rho(\xi_{2k+1})\rho(\xi_j)$ for $j \le 2k$. Next, for each $B \subseteq \{1,\dots,2k\}$, $\rho(\xi_B \xi_{2k+1}) = \pm(\rho(\xi_{B^c}), -\rho(\xi_{B^c}))$ where $B^c = \{1,\dots,2k\} \setminus B$, which proves the last claim. ☐

It will be convenient to define $\zeta = \xi_1\xi_2\cdots\xi_{2k+1}$ and use the basis $\{\xi_A, \zeta\xi_A \mid A \subseteq \{1,\dots,2k\}\}$ for $\mathbf{Cl}_{2k+1}$.

Next, we have an isomorphism

$$M_2(\mathbf{C})^{\otimes k} \cong M_{2^k}(\mathbf{C}),$$

which can be seen by picking a total ordering on the set of vectors $\{e_{i_1} \otimes \cdots \otimes e_{i_k} \mid i_j \in \{0,1\}\}$. Hence, when $n = 2k$ is even, $\mathbf{Cl}_{2k}$ has a 2^k-dimensional module that we call Δ, and when $n = 2k+1$ is even, $\mathbf{Cl}_{2k+1}$ has two non-isomorphic 2^k-dimensional modules that we call Δ_+ and Δ_-. We'll use $\mathrm{Tr}_\Delta(v)$ to denote the trace of $v \in \mathbf{Cl}_{2k}$ on Δ, and similarly for $\mathrm{Tr}_{\Delta_\pm}(v)$.

Proposition 9.3.3. (1) *For $n = 2k$, we have*

$$\mathrm{Tr}_\Delta\left(\sum_{A \subseteq \{1,\dots,2k\}} c_A \xi_A\right) = 2^k c_\varnothing.$$

(2) *For $n = 2k+1$, we have*

$$\mathrm{Tr}_{\Delta_\pm}\left(\sum_{A \subseteq \{1,\dots,2k\}} (c_A \xi_A + d_A \zeta \xi_A)\right) = 2^k c_\varnothing \pm (2i)^k d_\varnothing.$$

Proof. (1) If $A \neq \varnothing$, then $\rho(\xi_A)$ is a tensor product of 2×2 matrices, some of which are a non-empty product of some subset of $\{\varepsilon, x, y\}$. All such products have zero trace, and the trace of a tensor product is the product of the traces. On the other hand, the trace of 1 is just the dimension of the module, which is 2^k.

(2) The trace of $\zeta\xi_A$ on Δ_+ is the trace of the first component of $\rho(\zeta\xi_A)$, which is $i^k\rho'(\xi_A)$, and hence has trace 0 unless $A = \varnothing$. Otherwise, the trace is i^k. Likewise, the trace on Δ_- is the trace of the second component of $\rho(\zeta\xi_A)$. □

Let $\mathbf{Cl}_m^\times$ denote the group of invertible elements in $\mathbf{Cl}_m$ under multiplication.

Proposition 9.3.4. *Pick complex numbers $a_1, \dots, a_{n-1}, b_1, \dots, b_{n-1}$ such that $a_j^2 + b_j^2 = -1$ and $a_{j+1}b_j = \frac{1}{2}$. Then we have a group homomorphism $r_n \colon \widetilde{\Sigma}_n \to \mathbf{Cl}_n^\times$ defined by*

$$r_n(t_i) = a_i\xi_i + b_i\xi_{i+1}, \qquad r_n(z) = -1.$$

In particular, if $b_{n-1} = 0$, this gives a group homomorphism $\widetilde{\Sigma}_n \to \mathbf{Cl}_{n-1}^\times$.

Proof. It suffices to show that r_n preserves all of the relations among the generators $t_1, \dots, t_{n-1}$. Namely, $r_n(t_i)^2 = -1$, $(r_n(t_i)r_n(t_{i+1}))^3 = -1$, and $r_n(t_i)r_n(t_j) = -r_n(t_j)r_n(t_i)$ if $j > i+1$.

For the first one, the definition of the Clifford algebra gives $r_n(t_i)^2 = a_i^2 + b_i^2$, so the relation follows from our assumption. For the second, we note that $r_n(t_i)r_n(t_{i+1}) + r_n(t_{i+1})r_n(t_i) = 2a_ia_{i+1} = 1$. So we can replace $r_n(t_{i+1})r_n(t_i)$ by $-r_n(t_i)r_n(t_{i+1}) + 1$ which we do below (underlining the part that we replace):

$$\begin{aligned} &r_n(t_i)\underline{r_n(t_{i+1})r_n(t_i)}r_n(t_{i+1})r_n(t_i)r_n(t_{i+1}) \\ &= -r_n(t_i)^2r_n(t_{i+1})^2r_n(t_i)r_n(t_{i+1}) + r_n(t_i)\underline{r_n(t_{i+1})r_n(t_i)}r_n(t_{i+1}) \\ &= -r_n(t_i)r_n(t_{i+1}) - r_n(t_i)^2r_n(t_{i+1})^2 + r_n(t_i)r_n(t_{i+1}) = -1. \end{aligned}$$

Finally, the last relation follows directly from the Clifford algebra, and does not use the relations on the a_i and b_i, since there are no repeating ξ_i in the relation. □

We can find several choices of a_i, b_i satisfying the above relations and $b_{n-1} = 0$, pick any such one and denote the resulting homomorphism

$$\psi \colon \widetilde{\Sigma}_n \to \mathbf{Cl}_{n-1}^\times.$$

When $n = 2k+1$ is odd, we can compose with the map $\mathbf{Cl}_{2k}^\times \to \mathbf{GL}_{2^k}(\mathbf{C})$ coming from the module Δ. This is a negative representation of $\widetilde{\Sigma}_{2k+1}$, and we let φ^{2k+1} be the corresponding character.

Likewise, when $n = 2k$ is even, we can compose with either of the maps $\mathbf{Cl}_{2k-1}^\times \to \mathbf{GL}_{2^{k-1}}(\mathbf{C})$ coming from the modules $\Delta_\pm$. These are negative representations of $\widetilde{\Sigma}_{2k}$ and we let $\varphi_\pm^{2k}$ be the corresponding characters and let $\varphi^{2k} = \varphi_+^{2k} + \varphi_-^{2k}$.

Theorem 9.3.5. *We have*

$$\begin{aligned} \varphi^{2k+1}(\sigma^\lambda) &= 2^{(\ell(\lambda)-1)/2} \quad &&\text{if } \lambda \in \mathrm{OPar}(2k+1), \\ \varphi^{2k}(\sigma^\lambda) &= 2^{\ell(\lambda)/2} &&\text{if } \lambda \in \mathrm{OPar}(2k). \end{aligned}$$

In all other cases, $\varphi^n(\sigma^\lambda) = 0$.

See [Ste, Theorem 3.3] for a more refined calculation of the values of $\varphi_\pm^{2k}$, though we won't need it.

Proof. Recall the definition of σ^λ. It is a product $T_1T_2\cdots T_{\ell(\lambda)}$ where $T_j = t_{r+1}t_{r+2}\cdots t_{r+\lambda_j-1}$ where $r = \lambda_1 + \cdots + \lambda_{j-1}$. Applying ψ to T_j, we get

$$(a_{r+1}\xi_{r+1} + b_{r+1}\xi_{r+2})(a_{r+2}\xi_{r+2} + b_{r+2}\xi_{r+3})\cdots(a_{r+\lambda_j-1}\xi_{r+\lambda_j-1} + b_{r+\lambda_j-1}\xi_{r+\lambda_j}).$$

Each ξ_i appears in at most one $\psi(T_j)$, so the constant term (the coefficient of $\xi_\varnothing$) of $\psi(\sigma^\lambda)$ is the product of the constant terms of the $\psi(T_j)$. Furthermore, there is no constant term of $\psi(T_j)$ if there is an odd number of factors, so there is only a constant term if each λ_j is odd, i.e., $\lambda \in \mathrm{OPar}$. In that case, the constant term of $\psi(T_j)$ is

$$(b_{r+1}a_{r+2})(b_{r+3}a_{r+4})\cdots(b_{r+\lambda_j-2}a_{r+\lambda_j-1}) = (1/2)^{(\lambda_j-1)/2} = 2^{(1-\lambda_j)/2}.$$

and hence the constant term of $\psi(\sigma^\lambda)$ is $2^{(\ell(\lambda)-n)/2}$.

Now we finish with Proposition 9.3.3. If $n = 2k+1$, then $\varphi^{2k+1}(\sigma^\lambda)$ is 2^k times the constant term of $\psi(\sigma^\lambda)$, so $\varphi^{2k+1}(\sigma^\lambda) = 2^{(\ell(\lambda)-1)/2}$. If $n = 2k$, then $\varphi^{2k}_+(\sigma^\lambda) + \varphi^{2k}_-(\sigma^\lambda)$ is $2 \cdot 2^{k-1}$ times the constant term of $\psi(\sigma^\lambda)$, so we get $2^{\ell(\lambda)/2}$. $\square$

9.4. **Character ring.** Recall that for symmetric groups, we identified $\Sigma_n \times \Sigma_m$ as a subgroup of Σ_{n+m}. This allowed us to build an induction product: given representations V, W of Σ_n, Σ_m, respectively, we constructed $\mathrm{Ind}_{\Sigma_n \times \Sigma_m}^{\Sigma_{n+m}} V \otimes W$.

We want an analogue for negative representations. However, the problem is that the tensor product of two negative representations is no longer negative. To get around this, we introduce a formalism.

We let $\mathscr{G}$ be the collection of triples (G, z, σ) such that G is a finite group, $z \in G$ is a central element of order 2, and $\sigma \colon G \to \mathbf{Z}/2$ is a homomorphism. Our primary example is $(\widetilde{\Sigma}_n, z, \mathrm{sgn})$. Given two such triples (G_1, z_1, σ_1) and (G_2, z_2, σ_2), first define $G_1 \widetilde{\times} G_2$ to be the product $G_1 \times G_2$ with the multiplication

$$(x_1, x_2)(y_1, y_2) = (z_1^{\sigma_2(x_2)\sigma_1(y_1)} x_1 y_1, x_2 y_2).$$

This has a central subgroup $\{(1,1), (z_1, 1), (1, z_2), (z_1, z_2)\}$ and we define $G_1 \widehat{\times} G_2$ to be the group $(G_1 \widetilde{\times} G_2)/\langle (z_1, z_2)\rangle$. Then in the quotient, $(z_1, 1) = (1, z_2)$ and we define σ by $\sigma(x, y) = \sigma_1(x) + \sigma_2(y)$. This gives another triple in $\mathscr{G}$.

Of interest to us is that $\widetilde{\Sigma}_n \widehat{\times} \widetilde{\Sigma}_m \subseteq \widetilde{\Sigma}_{n+m}$.

There are several approaches to starting with negative representations of G_1 and G_2 and building one for $G_1 \widehat{\times} G_2$. However, they are either very intricate or require one to work with $\mathbf{Z}/2$-graded representations. We will skip this and state one result at the end of this section that we will use later. In the case of $G_1 = \widetilde{\Sigma}_m$ and $G_2 = \widetilde{\Sigma}_n$, the construction in [J] gives us a form of induction which is easy to describe on character rings.

Via π, we get a sign function $\widetilde{\Sigma}_n \to \Sigma_n \to \{1, -1\}$, which is a 1-dimensional representation (not a negative one). Given a representation V of $\widetilde{\Sigma}_n$, we can twist it by the sign character by defining a new action of g on v to be $\mathrm{sgn}(g)gv$. This twist is called the **associate representation**, and a representation is **self-associate** if it is isomorphic to its associate. If χ is the character of a representation, then it is self-associate if and only if $\chi(g) = 0$ whenever $\pi(g)$ is odd.

We let Θ_n be the $\mathbf{Z}$-span of the negative characters χ on $\widetilde{\Sigma}_n$ such that $\chi(g) = 0$ whenever the cycle type of $\pi(g)$ has an even part.

Lemma 9.4.1. *Θ_n is the space spanned by the characters of the self-associate negative representations of $\widetilde{\Sigma}_n$.*

Proof. Let χ be the character of a negative representation of $\widetilde{\Sigma}_n$. For any $g \in \widetilde{\Sigma}_n$, we have $\chi(g) = -\chi(zg)$. As mentioned above, the preimage of every conjugacy class $c_\lambda \subset \Sigma_n$ either splits into two in $\widetilde{\Sigma}_n$, or remains a single conjugacy class. In the second case, we have $\chi(g) = \chi(zg)$, so $\chi(g) = 0$ for any g in such a case.

Now suppose χ is the character of a self-associate negative representation and pick $g \in \widetilde{\Sigma}_n$ such that $\pi(g)$ has an even part. If $\pi(g)$ is odd, then by self-association, $\chi(g) = 0$. Otherwise, $\pi(g)$ is even, and so by Lemma 9.2.2, the preimage of the conjugacy class of $\pi(g)$ is a single conjugacy class, so by the above paragraph, $\chi(g) = 0$. Hence $\chi \in \Theta_n$.

Conversely, suppose $\chi \in \Theta_n$ and pick g such that $\pi(g)$ is odd. Then $\pi(g)$ has at least one even cycle, so $\chi(g) = 0$ by definition of Θ_n. $\square$

Now set

$$\Theta = \bigoplus_{n \geq 0} \Theta_n.$$

The product is defined like in the case for symmetric groups. Given $f \in \Theta_m$ and $g \in \Theta_n$, we have $f \cdot g \in \Theta_{m+n}$ defined by

$$(f \cdot g)(c^+_\gamma) = \sum_{\gamma = \alpha \cup \beta} \frac{z_\gamma}{z_\alpha z_\beta} f(c^+_\alpha) g(c^+_\beta).$$

For the following facts (which we will not explain here), see [J, §4A and Lemma 4.29]:

Lemma 9.4.2. *(1) If f and g are characters of representations, then so is $f \cdot g$.*
(2) If n and m are even, then $\frac{1}{2}\varphi^m \cdot \varphi^n$ is the character of a representation.

9.5. **Odd characteristic map.** We define the **odd (Frobenius) characteristic map** by

$$\begin{aligned} \mathrm{och}\colon \Theta_n &\to \Gamma_n \otimes \mathbf{C} \\ f &\mapsto \sum_{\lambda \in \mathrm{OPar}(n)} z_\lambda^{-1} 2^{\ell(\lambda)/2} f(c^+_\lambda) p_\lambda \end{aligned}$$

The class functions on $\widetilde{\Sigma}_n$ have an inner product defined as before:

$$\langle f, g \rangle_{\widetilde{\Sigma}_n} = \frac{1}{2 \cdot n!} \sum_{\sigma \in \widetilde{\Sigma}_n} f(\sigma) \overline{g(\sigma)}.$$

This restricts to an inner product on Θ_n which we will also denote by $\langle , \rangle_{\widetilde{\Sigma}_n}$. Recall that we have a bilinear form on $\Gamma_n \otimes \mathbf{Q}$ by specializing the bilinear form $\langle , \rangle_t$ to $t = -1$. We extend it to a Hermitian bilinear form on $\Gamma_n \otimes \mathbf{C}$ by $\langle \alpha f, \beta g \rangle = \alpha \overline{\beta} \langle f, g \rangle$ for $\alpha, \beta \in \mathbf{C}$.

Theorem 9.5.1. och *is an isometry, i.e., for all $f, g, \in \Theta_n$, we have $\langle f, g \rangle_{\widetilde{\Sigma}_n} = \langle \mathrm{och}(f), \mathrm{och}(g) \rangle$.*

Proof. By Proposition 9.1.7, we have

$$\langle \mathrm{och}(f), \mathrm{och}(g) \rangle = \sum_{\lambda \in \mathrm{OPar}(n)} z_\lambda^{-2} 2^{\ell(\lambda)} f(c^+_\lambda) \overline{g(c^+_\lambda)} \frac{z_\lambda}{2^{\ell(\lambda)}} = \sum_{\lambda \in \mathrm{OPar}(n)} z_\lambda^{-1} f(c^+_\lambda) \overline{g(c^+_\lambda)}.$$

On the other hand, since $f, g \in \Theta_n$, they take the value 0 on σ if $\pi(\sigma)$ has an even part in its cycle type. Also, $|c^+_\lambda| = |c^-_\lambda| = n!/z_\lambda$ by Lemma 9.2.2, so

$$\langle f, g \rangle_{\widetilde{\Sigma}_n} = \frac{1}{2 \cdot n!} \sum_{\lambda \in \mathrm{OPar}(n)} \frac{n!}{z_\lambda} (f(c^+_\lambda) \overline{g(c^+_\lambda)} + f(c^-_\lambda) \overline{g(c^-_\lambda)})$$

Finally, use that $f(c^+_\lambda) = -f(c^-_\lambda)$ and $g(c^+_\lambda) = -g(c^-_\lambda)$ (since the conjugacy classes differ by multiplication by z, which acts as the scalar -1 on a negative representation). $\square$

Proposition 9.5.2. och *is a ring homomorphism.*

Proof. This follows from the definitions. For $f \in \Theta_m$ and $g \in \Theta_n$, we have:

$$\begin{aligned}
\mathrm{och}(f)\mathrm{och}(g) &= \left(\sum_{\alpha \in \mathrm{OPar}(m)} z_\alpha^{-1} 2^{\ell(\alpha)/2} f(c_\alpha^+) p_\alpha\right)\left(\sum_{\beta \in \mathrm{OPar}(n)} z_\beta^{-1} 2^{\ell(\beta)/2} f(c_\beta^+) p_\beta\right)\\
&= \sum_{\gamma \in \mathrm{OPar}(m+n)} \sum_{\gamma = \alpha\cup\beta} z_\alpha^{-1} z_\beta^{-1} 2^{\ell(\gamma)/2} f(c_\alpha^+) g(c_\beta^+) p_\gamma\\
&= \sum_{\gamma \in \mathrm{OPar}(m+n)} z_\gamma^{-1} 2^{\ell(\gamma)/2} (f\cdot g)(c_\gamma^+) p_\gamma\\
&= \mathrm{och}(f\cdot g).
\end{aligned}$$

□

Define the parity of a partition λ to the parity of $|\lambda| - \ell(\lambda)$ and set

$$\varepsilon(\lambda) = \begin{cases} 1 & \text{if } \lambda \text{ is odd} \\ 0 & \text{if } \lambda \text{ is even} \end{cases}$$
$$d(\lambda) = 2^{(\varepsilon(\lambda)-\ell(\lambda))/2}.$$

Proposition 9.5.3. $d(\lambda) q_\lambda$ *is in the image of* och.

Proof. We do this by induction on $\ell(\lambda)$. When $\ell(\lambda) = 1$, we have $\lambda = (n)$. If $n = 2k+1$ is odd, then $d(n) = 2^{-1/2}$ and by Theorem 9.3.5, we have

$$d(n)^{-1}\mathrm{och}(\varphi^{2k+1}) = 2^{1/2} \sum_{\lambda \in \mathrm{OPar}(2k+1)} z_\lambda^{-1} 2^{\ell(\lambda)/2} 2^{(\ell(\lambda)-1)/2} p_\lambda = \sum_{\lambda \in \mathrm{OPar}(2k+1)} z_\lambda^{-1} 2^{\ell(\lambda)} p_\lambda.$$

By Lemma 9.1.3, this is the same as q_{2k+1}, so we conclude that $\mathrm{och}(\varphi^{2k+1}) = d(2k+1) q_{2k+1}$ in this case.

If $n = 2k$ is even, then $d(n) = 1$ and again by Theorem 9.3.5, we have

$$\mathrm{och}(\varphi_+^{2k} + \varphi_-^{2k}) = \sum_{\lambda \in \mathrm{OPar}(2k)} z_\lambda^{-1} 2^{\ell(\lambda)/2} 2^{\ell(\lambda)/2} p_\lambda = \sum_{\lambda \in \mathrm{OPar}(2k)} z_\lambda^{-1} 2^{\ell(\lambda)} p_\lambda.$$

Again, by Lemma 9.1.3, this is the same as q_{2k}, so $\mathrm{och}(\varphi^{2k}) = q_{2k}$.

In general, suppose λ has a even parts and b odd parts, so that $\ell(\lambda) = a+b$. If a is even, then $\ell(\lambda) = |\lambda| \pmod 2$, so λ is even. Also, by Lemma 9.4.2,

$$2^{-a/2}\varphi^{\lambda_1}\cdots\varphi^{\lambda_{a+b}}$$

is the character of a representation. From above, $\mathrm{och}(\varphi^n) = 2^{-1/2} q_n$ if n is odd, and $\mathrm{och}(\varphi^n) = q_n$ if n is even, so

$$\mathrm{och}(2^{-a/2}\varphi^{\lambda_1}\cdots\varphi^{\lambda_{a+b}}) = 2^{-\ell(\lambda)/2} q_\lambda = d(\lambda) q_\lambda.$$

The case when a is even is similar. □

Corollary 9.5.4. $d(\lambda) Q_\lambda$ *is in the image of* och.

Proof. Write $Q_\lambda = q_\lambda + \sum_{\mu > \lambda} a_{\lambda,\mu} q_\mu$. By Proposition 9.1.6, $a_{\lambda,\mu}$ an integer divisible by $2^{\ell(\lambda)-\ell(\mu)}$. So it suffices to know that $d(\lambda) a_{\lambda,\mu}/d(\mu) \in \mathbf{Z}$, and for that, we need that $(\varepsilon(\lambda) - \varepsilon(\mu) + \ell(\lambda) - \ell(\mu))/2$ is a non-negative integer, but this is clear from the definition of ε. □

Now set

$$\widetilde{\chi}^\lambda = \operatorname{och}^{-1}(d(\lambda)Q_\lambda).$$

Theorem 9.5.5. *If λ is even, then $\widetilde{\chi}^\lambda$ is an irreducible negative self-associate character. If λ is odd, then $\widetilde{\chi}^\lambda = \widetilde{\chi}^{\lambda,+} + \widetilde{\chi}^{\lambda,-}$ is a sum of two irreducible, non-isomorphic, associate characters.*

Proof. First, we have

$$\langle d(\lambda)Q_\lambda, d(\mu)Q_\mu\rangle = 2^{\varepsilon(\lambda)}\delta_{\lambda,\mu}.$$

By Theorem 9.5.1, och is an isometry, so the $\widetilde{\chi}^\lambda$ are orthogonal. In particular, since they are $\mathbf{Z}$-linear combinations of characters, we conclude that $\widetilde{\chi}^\lambda = \pm[\lambda]$ where $[\lambda]$ is an irreducible negative self-associate character when λ is even, and that $\widetilde{\chi}^\lambda = \pm[\lambda,+] \pm [\lambda,-]$ where $[\lambda,\pm]$ are irreducible non-isomorphic, associate, negative characters when λ is odd.

Recall $\xi^\lambda = \operatorname{och}^{-1}(d(\lambda)q_\lambda)$ is the character of a negative representation. Then

$$\xi^\lambda = \widetilde{\chi}^\lambda + \sum_{\mu>\lambda} b_{\lambda,\mu}\widetilde{\chi}^\mu$$

for some integers $b_{\lambda,\mu}$. First suppose λ is even. If $\widetilde{\chi}^\lambda = -[\lambda]$, then $\langle \xi^\lambda, \widetilde{\chi}^\lambda\rangle_{\widetilde{\Sigma}_n} \le 0$, but we know that

$$\langle \xi^\lambda, \widetilde{\chi}^\lambda\rangle_{\widetilde{\Sigma}_n} = \langle \widetilde{\chi}^\lambda, \widetilde{\chi}^\lambda\rangle_{\widetilde{\Sigma}_n} = 2^{\varepsilon(\lambda)}$$

by orthogonality, so we conclude that $\widetilde{\chi}^\lambda$ is an irreducible character when λ is even. Similarly, we can conclude that $\widetilde{\chi}^\lambda = [\lambda,+] + [\lambda,-]$ when λ is odd. ◻

In particular, we have

$$Q_\lambda = \sum_{\mu\in\mathrm{OPar}(n)} z_\mu^{-1} 2^{(\ell(\lambda)+\ell(\mu)-\varepsilon(\lambda))/2}\widetilde{\chi}^\lambda(c_\mu^+)p_\mu.$$

This is enough when λ is even since $\widetilde{\chi}^\lambda$ vanishes on any conjugacy class which has an even part. This approach will not give us the values of $\widetilde{\chi}^{\lambda,\pm}$ on other conjugacy classes (though we know they sum to 0). See [Ste, §7] for the values. Once they are given, we can verify that they are correct by using the orthogonality relations.

Appendix A. Change of bases

In the table below, we summarize the change of bases we have discussed. For instance, the column for h_λ and row m_μ describes the coefficient of m_μ in the expansion of h_λ.

	m_λ	e_λ	h_λ	p_λ	s_λ
m_μ	$\delta_{\lambda,\mu}$	$M_{\lambda,\mu}$ (Lemma 2.2.1)	$N_{\lambda,\mu}$ (§2.4)	$R_{\lambda,\mu}$ (§2.5)	Kostka numbers $K_{\lambda,\mu}$
e_μ		$\delta_{\lambda,\mu}$			Jacobi–Trudi identity
h_μ			$\delta_{\lambda,\mu}$		Jacobi–Trudi identity
p_μ		Theorem 2.5.5	Theorem 2.5.5	$\delta_{\lambda,\mu}$	$\chi^\lambda(\mu)/z_\mu$
s_μ		Pieri rule	Pieri rule	$\chi^\mu(\lambda)$	$\delta_{\lambda,\mu}$

References

[A] Kaan Akin, On complexes relating the Jacobi-Trudi identity with the Bernstein-Gel'fand-Gel'fand resolution *J. Algebra* **117** (1988), no. 2, 494–503.

[CW] Shun-Jen Cheng, Weiqiang Wang, *Dualities and representations of Lie superalgebras*, Graduate Studies in Mathematics **144**, American Mathematical Society, Providence, RI, 2012.

[Fu1] William Fulton, *Young Tableaux, with applications to representation theory and geometry*, London Mathematical Society Student Texts **35**, Cambridge University Press, Cambridge, 1997.

[Fu2] William Fulton, Eigenvalues, invariant factors, highest weights, and Schubert calculus, *Bull. Amer. Math. Soc. (N.S.)* **37** (2000), no. 3, 209–249, arXiv:math/9908012v3.

[FH] William Fulton, Joe Harris, *Representation Theory*, Graduate Texts in Math. **129**, Springer-Verlag, 2004.

[GNW] Curtis Greene, Albert Nijenhuis, Herbert S. Wilf, A probabilistic proof of a formula for the number of Young tableaux of a given shape, *Adv. in Math.* **31** (1979), no. 1, 104–109.

[HH] P. N. Hoffman, J. F. Humphreys, *Projective representations of the symmetric groups. Q-functions and shifted tableaux*, Oxford Mathematical Monographs, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1992.

[Ho] Roger Howe, Perspectives on invariant theory: Schur duality, multiplicity-free actions, and beyond, *Israel Mathematical Conference Proceedings* **8**, 1995.

[Hu] James E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Graduate Texts in Math. **9**, Springer-Verlag, 1978.

[J] Tadeusz Józefiak, Characters of projective representations of symmetric groups, *Expo. Math.* **7** (1989), 193–247.

[Mac] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, second edition, Oxford Mathematical Monographs, Oxford, 1995.

[Mal] F. Miller Maley, The Hall polynomial revisited, *J. Algebra* **184** (1996), 363–371.

[Ma] Laurent Manivel, *Symmetric functions, Schubert polynomials and degeneracy loci*, SMF/AMS Texts and Monographs **6**, American Mathematical Society, Providence, RI.

[O] Andrei Okounkov, Symmetric functions and random partitions, arXiv:math/0309074v1.

[MTV] E. Mukhin, V. Tarasov, A. Varchenko, Schubert calculus and representations of the general linear group, *J. Amer. Math. Soc.* **22** (2009), no. 4, 909–940, arXiv:0711.4079v1.

[Sc] Olivier Schiffmann, Lectures on Hall algebras, arXiv:math/0611617v2.

[Se] Jean-Pierre Serre, *Linear Representations of Finite Groups*, Graduate Texts in Math. **42**, Springer-Verlag, 1977.

[Sta] Richard P. Stanley, *Enumerative combinatorics. Vol. 2.* With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin, Cambridge Studies in Advanced Mathematics **62**, Cambridge University Press, Cambridge, 1999.

[Ste] John R. Stembridge, Shifted tableaux and the projective representations of symmetric groups, *Adv. Math* **74** (1989), 87–134.

[W] Jerzy Weyman, *Cohomology of Vector Bundles and Syzygies*, Cambridge University Press, Cambridge, 2003.

[Z] A. V. Zelevinskii, Resolutions, dual pairs and character formulas, *Funktsional. Anal. i Prilozhen.* **21** (1987), no. 2, 74–75.

Department of Mathematics, University of Wisconsin, Madison
URL: http://math.wisc.edu/~svs/
E-mail address: svs@math.wisc.edu