AN INTRODUCTION TO SCHUR FUNCTIONS

AMRITANSHU PRASAD

1. **Symmetric Functions.** We consider polynomials in n variables $x_1, \ldots, x_n$. Given a multiindex $\alpha = (\alpha_1, \ldots, \alpha_n)$, x^α denotes the monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ A symmetric polynomial in n variables $x_1, \ldots, x_n$ is a polynomial of the form

$$f(x_1, \ldots, x_n) = \sum_\alpha c_\alpha x^\alpha,$$

where, for any permutation $w \in S_n$,

$$c_{(\alpha_1,\ldots,\alpha_n)} = c_{(\alpha_{w(1)},\ldots,\alpha_{w(n)})}.$$

We call the integer partition λ obtained by sorting the coordinates of α the shape of α and write $\lambda = \lambda(\alpha)$. The most obvious example of a symmetric polynomial in n variables is the *monomial symmetric function*, defined for each integer partition λ:

$$m_\lambda = \sum_{\lambda(\alpha)=\lambda} c_\alpha x^\alpha.$$

Note that m_λ is homogeneous of degree $|\lambda|$ (the sum of the parts of λ).

Exercise 1.1. *Take $n = 4$. Compute the monomial symmetric functions $m_{(3)}$, $m_{(2,1)}$, and $m_{(1^3)}$.*

Theorem 1.2. *The polynomials $m_\lambda(x_1, \ldots, x_n)$, as λ runs over all the integer partition of d, form a basis for the space of homogeneous symmetric polynomials of degree d in n variables.*

2. **Complete and Elementary Symmetric Polynomials.** Recall that the coefficients of a polynomial are symmetric polynomials in its roots:

$$(t - x_1)(t - x_2) \cdots (t - x_n) = t^n - e_1(x_1, \ldots, x_n)t^{n-1} + \cdots + (-1)^n e_n(x_1, \ldots, x_n), \tag{1}$$

where coefficient $e_i(x_1, \ldots, x_n)$ of t^{n-i} is given by:

$$e_i(x_1, \ldots, x_n) = \sum_{1 \leq j_1 < \cdots < j_i \leq n} x_{j_1} x_{j_2} \cdots x_{j_i}. \tag{2}$$

The polynomial e_i is called the ith *elementary symmetric polynomial*. By convention, write $e_i(x_1, \ldots, x_n) = 0$, for $i > n$.

The identity (1) can be written more elegantly as:

$$(1 + tx_1) \cdots (1 + tx_n) = \sum_{i=0}^{n} e_i(x_1, \cdots, x_n) t^i.$$

Dually[1], the *complete symmetric polynomials* are defined by the formal identity:

$$\frac{1}{(1 - x_1 t) \cdots (1 - x_n t)} = \sum_{i=0}^{\infty} h_i(x_1, \cdots, x_n) t^i.$$

Example 2.1. *In three variables, we have:*

$$e_2(x_1, x_2, x_3) = x_1 x_2 + x_1 x_3 + x_2 x_3,$$
$$h_2(x_1, x_2, x_3) = x_1^2 + x_1 x_2 + x_1 x_3 + x_2^2 + x_2 x_3 + x_2^3.$$

Exercise 2.2. *Show that*

$$h_i(x_1, \cdots x_n) = \sum_{1 \leq j_1 \leq \cdots \leq j_i \leq n} x_{j_1} \cdots x_{j_i}.$$

More generally, for any integer partition $\lambda = (\lambda_1, \ldots, \lambda_l)$, define:

$$h_\lambda = h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_l},$$
$$e_\lambda = e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_l}.$$

Theorem 2.3. *Given integer partitions $\lambda = (\lambda_1, \ldots, \lambda_l)$ and $\mu = (\mu_1, \cdots, \mu_m)$ of and integer d, let $M_{\lambda\mu}$ denote the number of integer matrices (a_{ij}) with non-negative entries whose ith row sums to λ_i for each i, and whose jth column sums to μ_j for each j. Then*

$$h_\lambda = \sum_\mu M_{\lambda\mu} m_\mu.$$

Dually, let $= N_{\lambda\mu}$ denote the number of integer matrices (a_{ij}) with entries 0 or 1, whose ith row sums to λ_i for each i, and whose jth column sums to μ_j for each j.

$$e_\lambda = \sum_\mu N_{\lambda\mu} m_\mu.$$

[1]We will refer to the replacing of $(1 + u)$ by $(1 - u)^{-1}$ in a formal identity as *dualization*.

Proof. We first prove the second identity involving elementary symmetric functions. A monomial in the expansion:

$$e_\lambda = \prod_{i=1}^{l} \sum_{j_1 < \cdots < j_{\lambda_j}} x_{j_1} \cdots x_{j_{\lambda_i}}$$

is a product of summands, one chosen from each of the l factors. Construct an $l \times m$ matrix (a_{ij}) corresponding to such a choice as follows: if the summand $x_{j_1} \cdots x_{j_{\lambda_i}}$ is chosen from the ith factor, then set the entries $a_{i,j_1}, \dots, a_{i,j_{\lambda_j}}$ to be 1 (the remaining entries of the ith row are 0). Clearly the ith row of such a matrix sums to λ_i. The monomial corresponding to this choice is x^μ if, for each j, the the number of i for which x_j appears in $a_{i,j_1}, \dots, a_{i,j_{\lambda_j}}$, which is the sum of the jth column of the matrix (a_{ij}). It follows that the coefficient of x^μ, and hence the coefficient of m_μ in the expansion of e_λ in the basis of monomial symmetric functions of degree n, is $N_{\lambda\mu}$.

A similar proof can be given for the first identity involving complete symmetric functions. The only difference is that a variables are repeated in the monomials that appear in h_i. Counting the number of repetitions (instead of just recording 0 or 1) gives non-negative integer matrices. $\square$

3. **Alternating Polynomials.** An *alternating polynomial* in $x_1, \dots, x_n$ is of the form:

$$f(x_1, \dots, x_n) = \sum_\alpha c_\alpha x_\alpha, \tag{3}$$

where, $c_{w(\alpha)} = \epsilon(w) c_\alpha$ for every multiindex α as in Section 1. Here $\epsilon : S_n \to \{\pm 1\}$ denotes the sign function. Equivalently, an alternating polynomial is one whose sign is reversed upon the interchange of any two variables.

Exercise 3.1. *If α is a multiindex where $\alpha_i = \alpha_j$ for some $i \neq j$, then $c_\alpha = 0$.*

In particular, every monomial in an alternating polynomial must be composed of distinct powers. Moreover, the polynomial is completely determined by the coefficients c_α of strictly decreasing multiindices, namely, multiindices of the form $\alpha = (\alpha_1, \dots, \alpha_n)$ with $\alpha_1 > \cdots > \alpha_n$.

Exercise 3.2. *Let δ denote the strictly increasing multiindex $(n-1, n-2, \dots, 1, 0)$ of lowest degree. Given an integer partition with at most n parts, we will pad it with 0's so that it can be regarded as a weakly decreasing multiindex. Then $\lambda \mapsto \lambda + \delta$ is a bijection from the set* *of integer partitions with at most n onto the set of strictly decreasing multiindices.*

Example 3.3. *Let $\lambda = (\lambda_1, \dots, \lambda_n)$ be a weakly decreasing multiindex. The polynomial:*

$$a_{\lambda+\delta} = \det(x_i^{\lambda_j+n-j})$$

is alternating, with unique strictly decreasing monomial $x^{\lambda+\delta}$.

Exercise 3.4. *The alternating polynomial of the form* (3) *is equal to*

$$\sum_\lambda c_\lambda a_{\lambda+\delta},$$

the sum being over all weakly decreasing multiindices.

4. **Cauchy's bialternant form of a Schur function.** The simplest polynomial of the form $a_{\lambda+\delta}$ arises when $\lambda = 0$; a_δ is the Vandermonde determinant:

$$a_\delta = \prod_{1\leq i<j\leq n} (x_i - x_j).$$

Exercise 4.1. *Show that, for every weakly decreasing multiindex λ, $a_{\lambda+\delta}$ is divisible by a_δ in the ring of polynomials in $x_1, \dots, x_n$.*

Exercise 4.2. *Show that $f \mapsto fa_\delta$ is an isomorphism of the space of symmetric polynomials in $x_1, \dots, x_n$ of degree d onto the space of alternating polynomials of degree $d + \binom{n}{2}$.*

The above exercise allows us to give the historically oldest definition of Schur functions—*Cauchy's bialternant formula*:

$$s_\lambda(x_1, \dots, x_n) = a_{\lambda+\delta}/a_\delta, \tag{4}$$

for any partition λ with at most n parts. If λ has more than n parts, set $s_\lambda(x_1, \dots, x_n) = 0$. This is clearly a symmetric function of degree $|\lambda|$. When λ has more than n parts, we shall write $s_\lambda(x_1, \dots, x_n) = 0$

Theorem 4.3. *As λ runs over all integer partitions of d with at most n parts, the Schur functions $s_\lambda(x_1, \dots, x_n)$ form a basis of the space of all homogeneous symmetric functions in $x_1, \dots, x_n$ of degree d.*

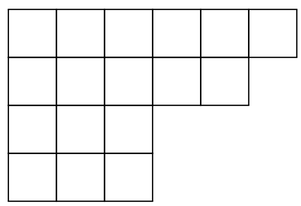
Proof. This follows from Exercises 3.4 and 4.2. □

Exercise 4.4 (Stability of Schur functions)**.** *Show that substituting $x_n = 0$ in the Schur function $s_\lambda(x_1, \dots, x_n)$ with n variables gives the corresponding Schur function $s_\lambda(x_1, \dots, x_{n-1})$ with $n - 1$ variables.*

5. **Pieri's rule.** The set of integer partitions is endowed with the *containment order*. We say that a partition $\lambda = (\lambda_1, \dots, \lambda_l)$ *contains* a partition $\mu = (\mu_1, \dots, \mu_m)$ if $l \geq m$, and $\lambda_i \geq \mu_i$ for every $i = 1, \cdots, m$. We write $\lambda \supset \mu$ or $\mu \subset \lambda$. Recall that the Young diagram of the partition λ is the set of points

$$\{(i,j) \mid 1 \leq i \leq l,\ 1 \leq j \leq \lambda_i\}.$$

Visually, each node (i,j) of the Young diagram is replaced by a box, and the box corresponding to (i,j) is placed in the ith row and jth column (matrix notation). Thus, the Young diagram of $\lambda = (6,5,3,3)$ is depicetd by:



Note that containment of partitions is nothing but the containment relation on their Young diagrams. By abuse of notation, we will also use λ to denote the Young diagram of λ.

By a skew-shape, we mean a difference of Young diagrams $\lambda \setminus \mu$, where $\lambda \supset \mu$. We write λ/μ for this skew-shape. A skew-shape is called a *horizontal strip* (respectively, a *vertical strip*) if it has at most one element in each vertical column (respectively, horizontal row).

Theorem 5.1. *For every partition λ, and every positive integer k,*

$$s_\lambda h_k = \sum_{\mu} s_\mu,$$

where the sum runs over all partitions $\mu \supset \lambda$ such that μ/λ is a horizontal strip of size k. Dually,

$$s_\lambda e_k = \sum_{\mu} s_\mu,$$

where the sum runs over all partitions $\mu \supset \lambda$ such that μ/λ is a vertical strip of size k.

Proof. The first identity is equivalent to showing that:

$$a_{\lambda+\delta} \sum_{\alpha_1+\cdots+\alpha_n=k} x_1^{\alpha_1} \cdots x_n^{\alpha_n} = \sum_{\mu} a_{\mu+\delta},$$

the sum on the right being over all partitions $\mu \supset \lambda$ such that μ/λ is a horizontal strip.

Writing $\alpha = (\alpha_1, \ldots, \alpha_n)$, the sum on the left hand side can be regarded as a sum of determinants:

$$a_{\lambda+\delta} \sum_{|\alpha|=k} x_1^{\alpha_1} \cdots x_n^{\alpha_n} = \sum_{|\alpha|=k} a_{\lambda+\alpha+\delta}. \tag{5}$$

Suppose there exists an integer α_i such that $\alpha_i > \lambda_i - \lambda_{i+1}$ (in other words, $(\lambda+\alpha)/\lambda$ is not a horizontal strip), then define $\beta = (\beta_1, \ldots, \beta_n)$ by $\beta_i = \alpha_{i+1} - (\lambda_i - \lambda_{i+1} + 1)$, $\beta_{i+1} = \alpha_i + (\lambda_i - \lambda_{i+1} + 1)$, and $\beta_j = \alpha_j$ for all $j \notin \{i, i+1\}$. Then $a_{\lambda+\alpha+\delta} = -a_{\lambda+\beta+\delta}$. So the only terms that survive on the right hand side of (5) are of the form $a_{\mu+\delta}$, where μ/λ is a horizontal strip.

The proof of the second identity in the theorem is similar (in fact, a little simpler) and is left to the reader as an exercise. $\square$

As a special case of Pieri's rule, we have:

Corollary 5.2. *For every positive integer k,*

$$s_{(k)} = h_k, \text{ and } s_{(1^k)} = e_k.$$

6. Schur to Complete and Elementary via Tableaux.

Pieri's rule allows us to compute the complete and elementary symmetric functions h_λ and e_λ in terms of Schur functions.

Example 6.1. *Repeated application of Pieri's rule gives,:*

$e_{(2,2,1)} =$

$$\begin{tikzcd}[column sep=small]
 & & & & e_2e_2e_1 \arrow[d] & & \\
 & & & & s_{(1^2)}e_2e_1 \arrow[dll] \arrow[d] \arrow[dr] & & \\
 & & s_{(2^2)}e_1 \arrow[dll] \arrow[dl] \arrow[d] & & s_{(2,1^2)}e_1 \arrow[dl] \arrow[d] \arrow[dr] & s_{(1^4)}e_1 \arrow[dr] & \\
s_{(3,2)} & s_{(2^2,1)} & s_{(3,1^2)} & s_{(2^2,1)} & s_{(2,1^3)}] & s_{(2,1^3)} & s_{(1^5)}
\end{tikzcd}$$

The steps going from the first line of the above calculation to each term of the last line can be recorded by putting numbers into Young diagrams:

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

7. The Lindström-Gessel-Viennot Lemma.

Let R be a commutative ring. Let S be any set of points, and $v : S \times S \to R$ be any function (we think of w as a *weight function.* Given $s, t \in S$, a path in S from s to t is is a sequence $\omega = (s = s_0, s_1, \ldots, s_k = t)$ of distinct points in S. We denote this by $\omega : s \to t$. The weight of the path ω is defined to be:

$$v(\omega) = v(s_0, s_1)v(s_1, s_2)\cdots v(s_{k-1}, s_k).$$

Definition 7.1 (non-crossing paths)**.** Two paths $\omega = (s_0, \dots, s_k)$ and $\eta = (t_0, \dots, t_l)$ are said to be non-crossing if $s_i \neq t_j$ for all $0 \leq i \leq k$ and $0 \leq t \leq l$.

Fix points $A_1, \dots, A_n$ and $B_1, \dots, B_n$ in S, and define an $n \times n$ matrix (a_{ij}) by:

$$a_{ij} = \sum_{\omega : A_i \to B_j} v(\omega).$$

Theorem 7.2 (Lindström-Gessel-Viennot Lemma)**.** *The determinant of the matrix* (a_{ij}) *defined above is given by:*

$$\det(a_{ij}) = \sum_{\omega_i : A_i \to B_i} v(\omega_1)\cdots v(\omega_n),$$

where the sum is over all n-tuples $(\omega_1, \dots, \omega_n)$ *of pairwise non-crossing paths.*

8. **The Jacobi-Trudi Identities.**

$$s_\lambda = \det(h_{\lambda_i - i + j}), \quad s_{\lambda'} = \det(e_{\lambda_i - i + j}).$$

For the first Jacobi-Trudi identity take S to be the positive cone in the the integer lattice:

$$S = \{(i, j) \mid i \geq 0,\ j > 0 \text{ are integers}\}.$$

Set the weight $v((i, j), (i + 1, j))$ of each rightward horizontal edge to be x_j for $j = 1, \dots, n$, the weight of each downward vertical edge $v((i, j), (i, j - 1))$ to be 1 for all $j = 2, \dots, n$. The remaining weights are all zero.

Lemma 8.1. *For all integers $i > 0$ and $j \geq 0$, we have:*

$$\sum_{\omega : (i,n) \to (i+j,1)} v(\omega) = h_j(x_1, \dots, x_n).$$

Proof. Only rightward or downward steps have non-zero weights. So every path with non-zero weight is composed of unit downward and rightward steps. A path with non-zero weight from (i, n) to $(i + j, 1)$ must have exactly j rightward steps, say in rows $n \geq i_1 \geq i_2 \cdots \geq i_j \geq 1$. The weight of such a path is $x_{i_1} \cdots x_{i_j}$, and hence, the sum of the weights of all such paths is $h_j(x_1, \dots, x_n)$. For an example, see Fig. 1. □

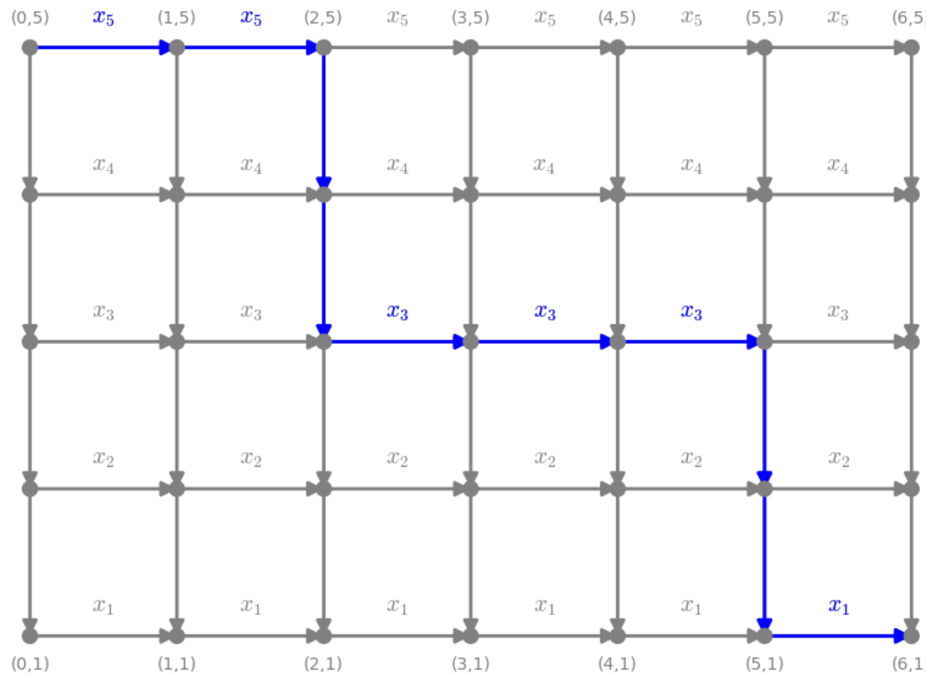


FIGURE 1. A path from $(0, 5)$ to $(6, 1)$ representing the monomial $x_1x_3^3x_5^2$ in h_6.

Let $A_i = (n - i, n)$ and $B_i = (\lambda_i + n - i, 1)$ for $i = 1, \dots, n$. Then by Lemma 8.1,

$$\sum_{\omega:A_i \to B_j} v(\omega) = h_{\lambda_j+i-j}.$$