Course 18.312: Algebraic Combinatorics

Lecture Notes # 18-19 Addendum by Gregg Musiker

March 18th - 20th, 2009

The following material can be found in a number of sources, including Sections $7.3-7.5, 7.7, 7.10-7.11, 7.15-16$ of Stanley's Enumerative Combinatorics Volume 2.

1 Elementary and Homogeneous Symmetric Functions

A polynomial in n variables, $P(x_1, x_2, \ldots, x_n) \in \mathbb{C}[x_1, x_2, \ldots, x_n]$ is known as a **symmetric polynomial** if for any permutation $\sigma \in S_n$, $P(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}) = P(x_1, x_2, \ldots, x_n)$.

An important family of symmetric polynomials is the family of **elementary symmetric functions**.

$$e_k = e_k(x_1, x_2, \ldots, x_n) := \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} x_{i_1} x_{i_2} \cdots x_{i_k}.$$

Notice that $e_0 = 1$, $e_k(x_1, x_2, \ldots, x_n) = 0$ if $k > n$ and the number of terms in $e_k(x_1, x_2, \ldots, x_n)$ is $\binom{n}{k}$. If $\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_\ell]$ is a partition, $e_\lambda := e_{\lambda_1} \cdot e_{\lambda_2} \cdots e_{\lambda_\ell}$.

(Fundamental Theorem of Symmetric Functions) Any symmetric polynomial with coefficients in $\mathbb{C}$ can be written as a $\mathbb{C}$-linear combination of the e_λ's. Furthermore, any symmetric polynomial with coefficients in $\mathbb{Z}$ can be written as a $\mathbb{Z}$-linear combination of the e_λ's.

We will not prove this theorem but will illustrate it for a few important examples of symmetric functions.

Let $E(t) := \sum_{k=0}^{\infty} e_k t^k$. Then $E(t) = \prod_i (1 + x_i t)$. In particular, if we are working with symmetric polynomials in n variables, then i ranges over $\{1, 2, \ldots, n\}$ in this product.

Another important family of symmetric functions is family of **homogeneous** symmetric functions, defined as

$$h_k = h_k(x_1, x_2, \ldots, x_n) := \sum_{1 \le i_1 \le i_2 \le \cdots \le i_k \le n} x_{i_1} x_{i_2} \cdots x_{i_k}.$$

Similarly we let $h_\lambda = h_{\lambda_1} \cdot h_{\lambda_2} \cdots h_{\lambda_\ell}$, $h_0 = 1$, $h_1 = e_1$, and the number of terms in $h_k(x_1, x_2, \ldots, x_n)$ is $\left(\binom{n}{k}\right)$, the number of k-element multisets of $\{1, 2, \ldots, n\}$.

Let $H(t) := \sum_{k=0}^{\infty} h_k t^k$. Then $H(t) = \prod_i \frac{1}{(1-x_i t)}$. As a consequence we get the following result.

Theorem. We have the identity for all $k \ge 1$:

$$\sum_{i=0}^{k} (-1)^k e_i h_{k-i} = 0.$$

Proof. From the above, we see that $H(t)E(-t) = 1$ so the convolution

$$\sum_{i=0}^{k} (-1)^i e_i h_{k-i} = \begin{cases} 1 & \text{if } \ k = 0 \\ 0 & \text{if } \ k \ge 1 \end{cases}.$$

As a corollary, we get that

$$h_k = \sum_{i=1}^{k} (-1)^{i-1} e_i h_{k-i}.$$

Thus by induction, we get explicit expressions for h_k as a polynomial in terms of e_1 through e_k.

Since these identities are true regardless of the number of variables appearing in the polynomials, these are symmetric *function* identities rather than simply identities of polynomials.

2 Power symmetric functions

We define

$$p_k = p_k(x_1, x_2, \ldots, x_n) := x_1^k + x_2^k + \cdots + x_n^k,$$

the **power** symmetric functions, with $p_\lambda = p_{\lambda_1} \cdot p_{\lambda_2} \cdots p_{\lambda_\ell}$

Theorem. These functions satisfy the **Newton-Girard** identities for all $k \geq 1$:

$$\begin{aligned} ke_k &= \sum_{i=1}^{k} (-1)^{i-1} e_{k-i} p_i \\ kh_k &= \sum_{i=1}^{k} h_{k-i} p_i. \end{aligned}$$

Proof. We prove the second identity, involving the power symmetric functions and the homogeneous symmetric functions. Let

$$P(t) = \sum_{k=1}^{\infty} p_k t^k.$$

Notice that

$$\frac{d}{dt}\Big(H(t)\Big) = H'(t) = \sum_{k=0}^{\infty} kh_k t^{k-1},$$

and the logarithmic derivative

$$\begin{aligned} \frac{H'(t)}{H(t)} = \frac{d}{dt}\Big(\log\ H(t)\Big) &= \frac{d}{dt}\Big(\log\ \prod_i (1 - x_i t)^{-1}\Big) \\ &= \frac{d}{dt}\Big(\sum_i - \log(1 - x_i t)\Big) \\ &= \frac{d}{dt}\Big(\sum_i \sum_{j=1}^{\infty} \frac{(x_i\ t)^j}{j}\Big) \\ &= \sum_{j=1}^{\infty}\Big(\sum_i x_i^j\ t^{j-1}\Big) \\ &= \sum_{k=1}^{\infty} p_k t^{k-1} = \frac{P(t)}{t}. \end{aligned}$$

Thus $P(t)H(t) = tH'(t)$ and each coefficient of t^k in the convolution on the LHS, $\sum_{i=1}^{k} h_{k-i} p_i$, eqauls the coefficient of t^{k-1} in $H'(t)$, namely kh_k.

The proof of the first identity is analogous. We leave it to the reader.

As above, we can use these identities like these to rewrite p_k's in terms of e_λ's or h_λ's, respectively, or vice-versa. First we introduce some notation.

For $i \geq 1$, let $m_i = m_i(\lambda)$ copies of the number i in λ. (Note that $m_i = 0$ for $i > |\lambda|$.) $z_\lambda = \prod_{i=1}^{\infty} i^{m_i} \cdot (m_i)!$. Let $\epsilon_\lambda = (-1)^{m_2 + m_4 + m_6 + \cdots}$.

Lemma. If $\lambda \vdash n$ and has ℓ nonzero parts, then $\epsilon_\lambda = (-1)^{n-\ell}$. In particular, ϵ_λ is the sign of the permutation that contains $m_i(\lambda)$ i-cycles (for $i \geq 1$).

Proof. Left to the reader.

Using this notation we obtain the following result.

Theorem.

$$\begin{aligned} h_k &= \sum_{\lambda \vdash k} \frac{p_\lambda}{z_\lambda} \\ e_k &= \sum_{\lambda \vdash k} \epsilon_\lambda \frac{p_\lambda}{z_\lambda} \end{aligned}$$

Proof. We saw in the last proof that

$$\frac{d}{dt}\left(\log H(t) \right) = \frac{P(t)}{t}.$$

As a consequence,

$$\sum_{k=0}^{\infty} h_k t^k = \exp\left(\sum_{k=1}^{\infty} \frac{p_k}{k} t^k \right).$$

The exponential of a series, $\exp(\sum_{k=1}^{\infty} a_k t^k) = \exp(A(t))$ equals the sum $\sum_{i=0}^{\infty} \frac{A(t)^i}{i!}$, which can be rewritten as the double sum

$$\sum_{i=0}^{\infty} \sum_{\substack{\text{unordered compostion } r_1+r_2+r_3+\cdots+r_i=i \\ \text{each } r_j \text{ is a nonnegative integer}}} \binom{i}{r_1, r_2, \ldots, r_i} \frac{(a_1 t)^{r_1} (a_2 t^2)^{r_2} \cdots (a_i t^i)^{r_i}}{i!}$$

after expanding each term by the multinomial theorem.

Since the order of the composition does not matter, and only nonzero parts contribute to the summands, we can think of these r_j's as the number of j's in a partition $\lambda \vdash i$, i.e. each such composition gives rise to a λ so that $r_j = m_j(\lambda)$. We then use the above notation to rephrase this sum as

$$\exp(A(t)) = \sum_{i=0}^{\infty} \sum_{\lambda \vdash i} \binom{i}{m_1, m_2, \ldots, m_i} \frac{(a_1^{m_1} a_2^{m_2} \cdots a_i^{m_i})\ t^i}{i!}.$$

We leave as an exercise that the coefficient of t^k in $\exp\left(\sum_{k=1}^{\infty} \frac{p_k}{k} t^k \right)$ is $\sum_{\lambda \vdash k} \frac{p_\lambda}{z_\lambda}$.

3 Monomial Symmetric Functions

An even simpler family of symmetric functions is the family of **monomial** symmetric functions.

$$m_\lambda = m_\lambda(x_1, x_2, \ldots, x_n) := \sum_{[\alpha_1,\alpha_2,\ldots,\alpha_n] \text{ is a rearrangement of } [\lambda_1,\lambda_2,\ldots,\lambda_\ell,0,0,\ldots,0]} x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$$

if $n > \ell$, the number of nonzero parts in λ, and we set $m_\lambda(x_1, x_2, \ldots, x_n)$ to be zero otherwise.

(Note that when we think of m_λ as a formal symmetric function, i.e. in an infinite number of variables, this second case never occurs.)

Remark. Note that unlike the e_λ's, h_λ's and p_λ's, $m_\lambda \neq m_{\lambda_1} \cdot m_{\lambda_2} \cdots m_{\lambda_n}$.

Observation. $e_n = m_{[1^n]}$, $p_n = m_{[n]}$, and $h_n = \sum_{\lambda \vdash n} m_\lambda$.

4 Schur Functions

We define a fifth family of symmetric functions by using determinants. Let $\Delta(x_1, x_2, \ldots, x_n)$ denote the determinant of the matrix

$$a_\delta = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ x_1 & x_2 & x_3 & \ldots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \ldots & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \ldots & x_n^{n-1} \end{bmatrix}.$$

Theorem.

$$\Delta(x_1, x_2, \ldots, x_n) = \prod_{1 \le i < j \le n} (x_j - x_i).$$

Furthermore, $\Delta(x_1, x_2, \ldots, x_n)$ is the nonzero polynomial with smallest degree and the property that

$$\Delta(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}) = \operatorname{sgn}(\sigma)\Delta(x_1, x_2, \ldots, x_n)$$

for any permutation $\sigma \in S_n$. In particular, if σ is a **transposition** that just switches x_i and x_j, we get $-\Delta(x_1, x_2, \ldots, x_n)$ on the RHS.

Such a polynomial is called an **alternating** polynomial, and it follows from above that all alternating polynomials must be divisible by $\Delta(x_1, x_2, \ldots, x_n)$. We can build other alternating polynomials by taking the determinant of

$$a_{\lambda+\delta} = \begin{bmatrix} x_1^{\lambda_n} & x_2^{\lambda_n} & x_3^{\lambda_n} & \ldots & x_n^{\lambda_n} \\ x_1^{\lambda_{n-1}+1} & x_2^{\lambda_{n-1}+1} & x_3^{\lambda_{n-1}+1} & \ldots & x_n^{\lambda_{n-1}+1} \\ x_1^{\lambda_{n-2}+2} & x_2^{\lambda_{n-2}+2} & x_3^{\lambda_{n-2}+2} & \ldots & x_n^{\lambda_{n-2}+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^{\lambda_1+(n-1)} & x_2^{\lambda_1+(n-1)} & x_3^{\lambda_1+(n-1)} & \ldots & x_n^{\lambda_1+(n-1)} \end{bmatrix},$$

for any partition $\lambda = [\lambda_1, \lambda_2, \ldots, \lambda_n]$ with at most n parts, written in weakly decreasing order.

Consequently, the quotient

$$s_\lambda = s_\lambda(x_1, x_2, \ldots, x_n) = \frac{\det(a_{\lambda+\delta})}{\det(a_\delta)}$$

is the quotient of two alternating polynomials, and is in fact a symmetric polynomial (function). We call these s_λ's **Schur functions**.

Remark. Note that like the m_λ's, $s_\lambda \neq s_{\lambda_1} \cdot s_{\lambda_2} \cdots s_{\lambda_n}$.

The Schur functions are very important in the theory of representation theory of S_n and GL_n. We will not discuss such connections further in the course, although there are many possible final projects on this topic.

There is a beautiful formula for writing the s_λ's in terms of the h_μ's (equivalently the e_μ's). The following two formulas are known as the **Jacobi-Trudi Identity**.

Theorem. If λ has ℓ nonzero parts, let JT_ℓ be the ℓ-by-ℓ matrix whose (i, j)th entry is $h_{\lambda_i - i + j}$, where we set $h_0 = 1$ and $h_{-k} = 0$ for $k < 0$. Then

$$s_\lambda = \det JT_\ell.$$

Recall that λ^T is the conjugate (or transpose) of λ. Let JT_ℓ' be the matrix whose (i, j)th entry is $e_{\lambda_i - i + j}$. Then we also obtain

$$s_{\lambda^T} = \det JT_\ell'.$$

Example.

$$s_{4,1}(x_1, x_2, x_3) = \det \begin{bmatrix} h_{4-1+1} & h_{4-1+2} & h_{4-1+3} \\ h_{1-2+1} & h_{1-2+2} & h_{1-2+3} \\ h_{0-3+1} & h_{0-3+2} & h_{0-3+3} \end{bmatrix} = \det \begin{bmatrix} h_4 & h_5 & h_6 \\ 1 & h_1 & h_2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Proof. We let $e_j^{(\ell)}$ denote the jth elementary symmetry function on the alphabet $\{x_1, x_2, \ldots, x_{\ell-1}, x_{\ell+1}, \ldots, x_n\}$.

$$\begin{aligned}\left(\sum_{i\geq 0} h_i t^i\right)\left(\sum_{j=0}^{n-1} e_j^{(\ell)}(-t)^j\right) &= \prod_{i=1}^{n} \frac{1}{1-x_i t} \prod_{\substack{m=1\\ m\neq \ell}}^{n} (1-x_m t) \\ &= \frac{1}{1-x_\ell t} = 1 + x_\ell t + x_\ell^2 t^2 + \ldots\end{aligned}$$

As a special application, we take the coefficient of t^{α_i} on both sides and obtain

$$\sum_{j=0}^{n-1} h_{\alpha_i - j} e_j^{(\ell)} (-1)^j = \sum_{j=1}^{n} h_{\alpha_i - n + j} e_{n-j}^{(\ell)} (-1)^{n-j} = x_\ell^{\alpha_i}.$$

This identity implies the matrix equation

$$H_\alpha E = A_\alpha,$$

where we let the entries of A_α be $x_j^{\alpha_i}$'s, the entries of H_α be h_{α_i-n+j}'s and the entries of E be $(-1)^{n-i} e_{n-i}^{(j)}$'s.

If we let $\alpha = [n-1, n-2, \ldots, 2, 1, 0]$ (resp. $\lambda + [n-1, n-2, \ldots, 2, 1, 0]$), the right-hand-side gives precisely the entries of the matrix appearing in the denominator (resp. numerator) of the Schur function.

It suffices to show that $\det E = \det A_{[n-1,n-2,\ldots,2,1,0]} = \Delta(x_1, x_2, \ldots, x_n)$, and thus we obtain

$$\det H_{\lambda+[n-1,n-2,\ldots,2,1,0]} = \frac{\det A_{\lambda+[n-1,n-2,\ldots,2,1,0]}}{\det A_{[n-1,n-2,\ldots,2,1,0]}}.$$

The formula $\det E = \det A_{[n-1,n-2,\ldots,2,1,0]}$ follows from the fact that $A_{[n-1,n-2,\ldots,2,1,0]} = H_{[n-1,n-2,\ldots,2,1,0]} E$ and $H_{[n-1,n-2,\ldots,2,1,0]}$ is an upper triangular matrix with ones on the diagonal. We saw $\det A_{[n-1,n-2,\ldots,2,1,0]} = \Delta(x_1, x_2, \ldots, x_n)$ above.

We close these notes with an alternative, more combinatorial definition, of Schur functions.

We begin by generalizing the definition of Standard Young Tableaux (SYT). Recall that a SYT of shape λ, $\lambda \vdash n$, is a filling of a Young diagram of shape λ using exactly the numbers $\{1, 2, \ldots, n\}$ such that the numbers in each row increase as we proceed to the right, and the numbers in each column increase as we proceed downwards.

A **Semi-standard Young Tableaux** (SSYT) of shape λ using no number smaller than 1 or larger than n is a filling of the Young diagram so that the numbers in each row **weakly increase** and the numbers in each column **strictly decrease**.

We define the weight x_T of a SSYT T to be the product $\prod_{i=1}^{n} x_i^{\#i'\text{s appearing in } T}$.

Theorem.

$$s_\lambda(x_1, x_2, \ldots, x_n) = \sum_{\text{SSYT } T \text{ of shape } \lambda \text{ using no number outside } 1 \leq i \leq n} x_T.$$

Proof. Omitted.

The proof of this theorem along with associated results or applications of SSYT is a possible final project.