Notes I: Symmetric functions

1.1. Fundamental definitions.

We denote by Λ_R the algebra of symmetric functions in infinitely many variables $\mathbf{z} = z_1, z_2, \ldots$, with coefficients in a ring R (typically $R = \mathbb{Z}$ or $\mathbb{Q}$).

In this set-up "symmetric functions" are not actually functions, but formal power series in the variables $\mathbf{z}$. By definition a formal power series f is a function assigning to each (finite) monomial in the $\mathbf{z}$'s a *coefficient* in R. We denote the coefficient of a monomial $\mathbf{z}^\nu$ in f by $f|_{\mathbf{z}^\nu}$, and write f formally as the infinite sum

$$f = \sum c_\nu \mathbf{z}^\nu, \tag{1}$$

where $c_\nu = f|_{\mathbf{z}^\nu}$.

The sum of two formal series is given by coefficient-wise addition. The product is given by the usual rule for multiplication of series, which is well-defined since there are only a finite number of contributions to a given term in the product.

More generally, if we have series $\{f_i\}$ such that each monomial $\mathbf{z}^\nu$ occurs with non-zero coefficient in only finitely many of them, the infinite sum $\sum_i f_i$ is well-defined. If, in addition, all but finitely many of the f_i have constant term 1, the infinite product $\prod_i f_i$ is well-defined.

Note that the formal sum expression for f in (1) is a valid infinite sum.

A formal series f as above is *symmetric* if similar monomials have equal coefficient in f. Monomials $\mathbf{z}^\nu$ and $\mathbf{z}^\mu$ are *similar* if they may be obtained from one another by permuting the variables, that is, if they have the same multiset of exponents.

Abusing terminology, we call f a symmetric *polynomial* if only a finite number of similarity classes of monomials have non-zero coefficient in f. Otherwise we will specifically refer to f as a symmetric *infinite series*. We say f is *homogeneous of degree* n if only terms of degree n occur with non-zero coefficient in f. Clearly a symmetric infinite series is a polynomial in our sense if and only if its (non-zero) terms have bounded degree.

One may regard the symmetric functions as series that are invariant for the action of an infinite symmetric group S_∞ permuting the variables. By S_∞ we might mean the group of all (infinite) permutations of the variables, or its subgroup of permutations fixing all but finitely many variables. Which group we choose makes no difference as far as symmetric functions are concerned.

To recover the classical symmetric polynomials in finitely many variables $z_1, \ldots, z_l$ (which are honest polynomials), we may specialize the remaining variables z_i to zero. The specialization of f is simply obtained by deleting all terms of f involving any of the variables we are zeroing.

1.2. Partitions.

A *partition of* n is a multiset of positive integers, or *parts*, whose sum is n. By convention we always write the parts in decreasing order:

$$\lambda = (\lambda_1 \geq \ldots \geq \lambda_l).$$

It is often convenient to admit into the notation some "parts" of zero at the end, so that a sequence as above of length l can stand for any partition with l or fewer parts.

The *diagram* of a partition is an array of *cells* in $\mathbb{N} \times \mathbb{N}$, with λ_i cells in the i-th row, as for example

$$\lambda = (4, 2, 2); \quad D(\lambda) = \begin{array}{llll} \bullet & \bullet & & \\ \bullet & \bullet & & \\ \bullet & \bullet & \bullet & \bullet \end{array} .$$

The *conjugate* partition λ' is defined by setting λ'_i equal to the number of parts $\geq i$ in λ, so that the diagram of λ' is the transpose of the diagram of λ.

The partitions of n are partially ordered by *dominance*, defined by

$$\lambda \leq \mu \quad \text{iff} \quad \lambda_1 + \cdots + \lambda_k \leq \mu_1 + \cdots + \mu_k \quad \text{for all } k.$$

For $i < j$ let $R_{ij}\lambda = (\lambda_1, \ldots, \lambda_i + 1, \ldots, \lambda_j - 1, \ldots, \lambda_l)$. This sequence may or may not be a partition, but in any case it clearly satisfies the defining condition for $R_{ij}\lambda > \lambda$ in dominance order. When $R_{ij}\lambda$ is a partition, its diagram differs from that of λ by the transfer of a cell from row j down to row i. In class we proved the following characterization of dominance.

Proposition: The dominance order on partitions of n is the transitive closure of the relation $\lambda \to \nu$ if $\nu = R_{ij}\lambda$ for some $i < j$.

Corollary: We have $\lambda \leq \mu$ if and only if $\lambda' \geq \mu'$.

1.3. The monomial basis.

Every similarity class of monomials $\mathbf{z}^\nu$ contains a unique monomial $\mathbf{z}^\lambda$ in which the exponent sequence λ is a partition [of $n =$ (degree of $\mathbf{z}^\lambda$)]. The *monomial symmetric functions* are defined by

$$m_\lambda = \sum \mathbf{z}^\nu, \quad \text{where } \mathbf{z}^\nu \text{ ranges over monomials similar to } \mathbf{z}^\lambda.$$

From the definition it is clear that every symmetric polynomial f is a unique R-linear combination of monomial symmetric functions, which we may also state as follows.

Proposition: The monomial symmetric functions m_λ form a basis of Λ_R as a free R-module. More precisely, $\{m_\lambda : |\lambda| = n\}$ is an R-basis of the submodule $\Lambda_R^{(n)}$ of symmetric functions homogeneous of degree n. In particular $\Lambda_R^{(n)}$ is free of rank equal to the number of partitions of n.

In finitely many variables, it is easy to see that we have the following variant, where $l(\lambda)$ denotes the number of parts of λ.

Proposition: The monomial symmetric functions $\{m_\lambda : l(\lambda) \leq l\}$ form an R-basis of $\Lambda_R(z_1, \ldots, z_l)$, while all m_λ with $l(\lambda) > l$ specialize to zero.

1.4. Elementary symmetric functions.

We define the k-th *elementary symmetric function* to be

$$e_k = m_{(1^k)}.$$

This is the sum of all products of k distinct variables z_i.

For a partition λ we set

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

Proposition: We have

$$e_\lambda = m_{\lambda'} + \sum_{\nu<\lambda'} c_{\lambda\nu} m_\nu$$

for some integers $c_{\lambda\nu}$.

Proof: The coefficient of m_ν in the expansion by monomials of e_λ is the same as the coefficient $e_\lambda|_{\mathbf{z}^\nu}$ (this principle holds for expansion of any symmetric function by monomials). This is the number of ways to realize $\mathbf{z}^\nu$ as a product of products $\mathbf{z}_{S_i}$, where for a set of indices $S = \{s_1, \ldots, s_k\}$, $\mathbf{z}_S$ denotes $z_{s_1} \cdots z_{s_k}$, and we require $|S_i| = \lambda_i$. Representing the sets S_i by rows of a zero-one matrix, we see that $c_{\lambda\nu}$ is the number of such matrices M with row-sums λ and column-sums ν.

Now for M having fixed row-sums, denoting the column-sums by $\tilde{\nu}_i$, we can simultaneously maximize $\tilde{\nu}_1 + \cdots + \tilde{\nu}_k$ for all k by the "rolling-ball" process: tilt the matrix to the left and roll the ones into the left-most positions in each row. This done, we will have $\tilde{\nu} = \lambda'$. Hence if $c_{\lambda\nu} \neq 0$ we have $\nu \leq \lambda'$, with equality for a unique matrix M, so that $c_{\lambda\lambda'} = 1$.

Corollary: The elementary symmetric functions e_λ form an R-basis of Λ_R. Equivalently, Λ_R is isomorphic to the polynomial ring $R[e_1, e_2, \ldots]$, *i.e.*, the elementary symmetric functions e_k are algebraically independent.

Since λ' has at most l parts iff the parts of λ are at most l, we have the following variant in l variables, known as the fundamental theorem of symmetric functions.

Corollary: $\Lambda_R(z_1, \ldots, z_l)$ is isomorphic to $R[e_1, \ldots, e_l]$.

1.5. Complete homogeneous symmetric functions.

We define the k-th *complete homogeneous symmetric function* to be

$$h_k = \sum_{|\lambda|=k} m_\lambda.$$

This is the sum of all monomials of degree k in the variables $\mathbf{z}$. As with the elementary symmetric functions we define

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

We will show presently that these form a basis of Λ_R; this result is not as easy as for the elementary symmetric functions. For now we note that by reasoning similar to that for e_λ we can obtain the following.

Proposition: Let $h_\lambda = \sum_\nu c_{\lambda\nu} m_\nu$. Then $c_{\lambda\nu}$ is the number of non-negative integer matrices with row sums λ and column-sums ν. In particular, $c_{\lambda\nu} = c_{\nu\lambda}$.

1.6. Schur functions.

For the moment, we work with a fixed finite number of variables $z_1, \ldots, z_l$, and take R to be $\mathbb{Z}$. Let α be a partition with *distinct parts*

$$\alpha = (\alpha_1 > \cdots > \alpha_l).$$

(Here we allow $\alpha_l = 0$ so α could have l or $l-1$ parts). The smallest such partition is

$$\delta = (l-1, l-2, \ldots, 1, 0),$$

and we may clearly put

$$\alpha = \lambda + \delta,$$

where λ is an arbitrary partition with at most l parts.

We define

$$a_\alpha(z_1, \dots, z_l) = \det \begin{bmatrix} z_1^{\alpha_1} & \dots & z_1^{\alpha_l} \\ \vdots & & \vdots \\ z_l^{\alpha_1} & \dots & z_l^{\alpha_l} \end{bmatrix} = \sum_{w \in S_l} \epsilon(w) w(\mathbf{z}^\alpha),$$

where $\epsilon(w)$ denotes the sign of the permutation w. We require α to have distinct parts so that a_α will not be zero. Note that a_α is *alternating*, *i.e.*, it changes sign when we exchange two variables z_i and z_j, since this exchanges two rows of the determinant.

Since a_α vanishes upon setting $z_i = z_j$, it must be divisible by

$$\prod_{i<j} (z_i - z_j). \tag{2}$$

In particular, this is true of a_δ. But both a_δ and the above product have degree $\binom{n}{2}$, so a_δ must be a constant multiple of (2). By considering the coefficient of $\mathbf{z}^\delta$ we see that the constant is 1, so we have the *Vandermonde identity*

$$a_\delta = \prod_{i<j} (z_i - z_j).$$

Now for any λ with $l(\lambda) \leq l$, both $a_{\lambda+\delta}$ and a_δ are alternating, and hence the "bialternant"

$$s_\lambda(\mathbf{z}) = a_{\lambda+\delta}/a_\delta$$

is a symmetric polynomial in $z_1, \dots, z_l$. These symmetric polynomials are the *Schur functions*. (The definition, which predates Schur, is Jacobi's).

Lemma: For $l(\lambda) \leq l$ we have

$$a_{\lambda+\delta_{l+1}}(z_1, \dots, z_l, 0) = (z_1 \cdots z_l) a_{\lambda+\delta_l}(z_1, \dots, z_l).$$

Proof: Write out the two determinants and check that the result follows by elementary properties of determinants.

Corollary: For $l(\lambda) \leq l$ we have $s_\lambda(z_1, \dots, z_l, 0) = s_\lambda(z_1, \dots, z_l)$, where the first Schur function is as defined in $l+1$ variables and the second is as defined in l variables.

Corollary: Let $s_\lambda(\mathbf{z}) = \sum_\mu K_{\lambda\mu} m_\mu(\mathbf{z})$. Then the integers $K_{\lambda\mu}$ do not depend on the number of variables l, provided only that we take it large enough for s_λ to be defined.

In infinitely many variables, and over any coefficient ring R, we now define the Schur functions s_λ by the formula

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

where the $K_{\lambda\mu}$ are as in the corollary. Then we automatically have that the specialization of s_λ to l variables, for $l \geq l(\lambda)$, agrees with the bialternant formula. The specialization of s_λ to fewer than $l(\lambda)$ variables, let us note, is zero. To see this, observe that in the bialternant formula for s_λ in $l = l(\lambda)$ variables, we have $\alpha_i > 0$ for all i, so if we specialize any further variable to zero, the determinant in the the numerator will have a zero row and will vanish.

There is a simple procedure for expressing any symmetric polynomial f in terms of Schur functions. Namely, specialize f to l variables, where l is large enough (l at least the degree of f will do), multiply f by a_δ, and extract the coefficient of $\mathbf{z}^{\lambda+\delta}$, which is also the coefficient of $a_{\lambda+\delta}$ in $a_\delta f$ and hence the coefficient of s_λ in f. Note that, at least over $\mathbb{Z}$ or any ring R in which 2 is not a zero-divisor, the $a_{\lambda+\delta}$'s form a "monomial" basis of the alternating polynomials just as the m_λ's do for the symmetric polynomials.

Over $\mathbb{Z}$, this shows in particular that every m_λ is a $\mathbb{Z}$-linear combination of Schur functions (of the same degree), and therefore, for each n, the square integer matrix $[K_{\lambda\mu}]$ with rows and columns indexed by partitions λ, μ of n is invertible. This implies the following result, for any coefficient ring R.

Proposition: The Schur functions s_λ form an R-basis of Λ_R.

1.7. Power-sums.

We define the k-th *power-sum symmetric function* to be

$$p_k = m_{(k)} = \sum_i z_i^k.$$

As with the elementary and complete homogeneous symmetric functions, we define

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

Proposition: We have

$$p_\lambda = \sum_{\nu \geq \lambda} u_{\lambda\nu} m_\lambda,$$

with $u_{\lambda\lambda} = \prod_i r_i!$, where $\lambda = (1^{r_1}, 2^{r_2}, \dots)$.

Proof: By considering the monomials that could appear in the product $p_{\lambda_1} \cdots p_{\lambda_k}$ we see immediately that $u_{\lambda\nu} = 0$ unless λ *refines* ν, that is, the parts of λ can be obtained by further partitioning the parts of ν. It is not hard to see that if λ refines ν then also $\lambda \leq \nu$ in the dominance order.

The formula for $u_{\lambda\lambda}$, the coefficient of $\mathbf{z}^\lambda$ in p_λ, is straightforward.

We now assume that our ground ring R contains $\mathbb{Q}$, so that we have division by non-zero integers, and the triangular matrix $[u_{\lambda\nu}]$ is invertible.

Corollary: Over R containing $\mathbb{Q}$, the power-sums p_λ form a basis of Λ_R, and we have $\Lambda_R \cong R[p_1, p_2, \dots]$.

1.8. Plethystic substitution.

Assume $\mathbb{Q} \subseteq R$. Then by Corollary 1.7, if $\mathbb{F}$ is an R-algebra, and $f_1, f_2, \dots \in \mathbb{F}$ are arbitrary, we may define a unique R-algebra homomorphism $\Lambda_R \to \mathbb{F}$ mapping the k-th power-sum p_k to f_k, for all k.

Specifically, we take $\mathbb{F}$ to be the algebra of formal Laurent series with coefficients in R, in an alphabet of variables, or *letters*, $a_1, a_2, \dots$. (In practice the letters will be denoted x_i, y_i, q, t, etc. rather than a_i.) Now given $A \in \mathbb{F}$, we define

$$\mathrm{ev}_A \colon \Lambda_{\mathbb{F}} \to \mathbb{F}$$

by the rule

$$\mathrm{ev}_A(p_k) = A|_{a_i \mapsto a_i^k},$$

that is, the image of p_k is the result of replacing each letter by its k-th power in A. Note that we can perfectly well define this on $\Lambda_{\mathbb{F}}$, not just on Λ_R, which will be useful later on. Here the symmetric function variables $\mathbf{z}$ are necessarily *not* part of the alphabet of letters, otherwise it would make no sense to speak of $\Lambda_{\mathbb{F}}$.

Definition: The *plethystic substitution* of A into f is $\mathrm{ev}_A(f)$, henceforth denoted $f[A]$.

Here is the simplest, but most important, example of this construction. Let $X = x_1 + x_2 + \cdots$, where $x_1, x_2, \ldots$ are letters. Then by definition, $p_k[X] = X|_{x_i \mapsto x_i^k} = p_k(x_1, x_2, \ldots) = p_k(\mathbf{x})$. Since plethystic substitution is an algebra homomorphism it immediately follows that for all $f \in \Lambda$ we have

$$f[X] = f(\mathbf{x}), \quad \text{when } X = x_1 + x_2 + \cdots.$$

Here X may be a finite or infinite sum of letters. This means we are now free to forget about our original variables $\mathbf{z}$: we may view Λ_R purely formally, and recover symmetric functions in actual variables $\mathbf{x}$ through the plethystic substitution of their formal sum X.

For clarity, let me repeat the set-up here. We first fix a ground ring R and an alphabet of letters, denoting by $\mathbb{F}$ the algebra of formal Laurent series in the letters. We form the algebra of symmetric functions $\Lambda_{\mathbb{F}}$ in formal variables $\mathbf{z}$ which are not part of the alphabet of letters. Immediately and forever, however, we forget about the variables $\mathbf{z}$, viewing $\Lambda_{\mathbb{F}}$ purely formally as $\mathbb{F}[p_1, p_2, \ldots]$. To recover an actual symmetric function $f(\mathbf{x})$ in *letters* $\mathbf{x}$, we take the plethystic evaluation $f[X]$, where here and throughout, X will always stand for the formal sum $x_1 + x_2 + \cdots$.

More generally, if $A = t_1 + t_2 + \cdots$ is a sum of monomials in the letters, each with coefficient 1, then $f[A] = f(t_1, t_2, \ldots)$. Note that if A is a series with positive integer coefficients, we can treat it as a sum of monomials with coefficient 1 by repeating terms as many times as their coefficients indicate, *e.g.*, $f[(x+y)^2] = f[x^2 + 2xy + y^2] = f(x^2, xy, xy, y^2)$.

One caution must be observed with plethystic substitution: since the letters have a distinguished role, it is generally not permissible to substitute for them inside the plethystic bracket. Thus if f is homogeneous of degree n it is true that $f[tX] = t^n f[X]$, where t is a letter, but NOT true that $f[-X] = (-1)^n f[X]$. This happens because the evaluation map ev_A is defined in terms of replacing letters by their k-th powers, but it makes no sense to "replace -1 by its k-th power." However the substitution of a monomial in the letters for a letter in a plethystic identity IS permissible, since this operation is compatible with the definition of ev_A.

Because we defined plethystic substitution using power-sums, we have assumed to this point that our coefficient ring R contains $\mathbb{Q}$. For series A with *integer coefficients* , however, possible to define the plethystic substitution $f[A]$ for $f \in \Lambda_{\mathbb{F}}$ over an arbitrary ground ring R, as we shall see below.

1.9. Plethystic identities.

If we take $X = x_1 + x_2 + \cdots$, where $\mathbf{x}$ is infinite, then the $p_k[X] = p_k(\mathbf{x})$ are algebraically independent. It follows that any plethystic equation in X which holds for $X = x_1 + x_2 + \cdots$, holds identically. The same principle may be applied with more than one formal alphabet $X, Y, \dots$.

Example: From the definition of the elementary symmetric functions it is easy to see that

$$e_n(x_1, x_2, \dots, y_1, y_2, \dots) = \sum_{k+l=n} e_k(x_1, x_2, \dots)e_l(y_1, y_2, \dots),$$

where here and elsewhere we adopt the convention that $e_0 = 1$. This proves the plethystic identity

$$e_n[X+Y] = \sum_{k+l=n} e_k[X]e_l[Y], \tag{3}$$

for arbitrary X, Y. By similar reasoning we have the identity

$$h_n[X+Y] = \sum_{k+l=n} h_k[X]h_l[Y]. \tag{4}$$

1.10. The "Cauchy kernel" Ω.

It is useful for many purposes to define the symmetric infinite series

$$\Omega = \sum_{n\geq 0} h_n,$$

with the convention that $h_0 = 1$.

The identity (4) implies the following identity.

Proposition: $\Omega[X+Y] = \Omega[X]\Omega[Y]$.

This immediately leads to some interesting formulas. To begin with, for a single letter x, we have $h_n[x] = h_n(x) = x^n$, so

$$\Omega[x] = \frac{1}{1-x}.$$

Then by the multiplicative property, with $X = x_1 + x_2 + \cdots$ as usual, we have

$$\Omega[X] = \sum_n h_n(\mathbf{x}) = \prod_i \frac{1}{1-x_i}.$$

Now using $\Omega[X]\Omega[-X] = 1$, we have

$$\Omega[-X] = \prod_i (1-x_i) = \sum_n (-1)^n e_n(\mathbf{x}).$$

Here the last sum is easily seen to be equal to the product, by inspection. Comparing homogeneous components in each degree n we obtain the plethystic identity

$$h_n[-X] = (-1)^n e_n[X], \tag{5}$$

and hence also

$$e_n[-X] = (-1)^n h_n[X]. \tag{6}$$

Note that X is not a letter—it is *arbitrary*, so it is correct to substitute $-X$ for X in (5) to get (6).

Combining (6) with (3), we have

$$e_n[X-Y] = \sum_{k+l=n} (-1)^l e_k[X] h_l[Y]. \tag{7}$$

This formula allows us to evaluate $e_n[A]$ for any A with integer coefficients, without reference to the definition of plethysm in terms of power sums. We simply write A as $X-Y$ where X and Y have positive integer coefficients, repeat terms as necessary to obtain sums of monomials with coefficient 1, and evaluate using (7).

Since $\Lambda_R \cong R[e_1, e_2, \ldots]$, we can define $f[A]$ for A with integer coefficients, over any coefficient ring R.

Another application of Ω is to express the complete homogeneous symmetric functions h_n in terms of the power-sums. For this, take $X = x_1 + x_2 + \cdots$ as always, so

$$\begin{aligned}
\Omega[X] &= \prod_i \frac{1}{1-x_i} \\
&= \exp(\sum_i \log \frac{1}{1-x_i}) \\
&= \exp(\sum_i \sum_{k>0} x_i^k/k) \\
&= \exp(\sum_{k>0} p_k[X]/k) \\
&= \prod_{k>0} \exp(p_k[X]/k) \\
&= \prod_{k>0} \sum_{r\geq 0} p_k[X]^r/r!k^r \\
&= \sum_\lambda p_\lambda[X]/z_\lambda,
\end{aligned}$$

where if $\lambda = (1^{r_1}, 2^{r_2}, \ldots)$, we define $z_\lambda = \prod_k r_k! k^{r_k}$. Hence in particular we have

$$h_n = \sum_{|\lambda|=n} p_\lambda/z_\lambda.$$

The quantity z_λ has an interesting combinatorial interpretation: it is the number of elements in the symmetric group S_n commuting with any given permutation w whose cycle type (the partition given by the lengths of its disjoint cyles) is λ. The significance of this point will appear later on.

1.11. The involution ω.

Since $\Lambda_R \cong R[e_1, e_2, \ldots]$, there is a unique endomorphism

$$\omega\colon \Lambda_R \to \Lambda_R \quad \text{defined by} \quad \omega e_k = h_k.$$

(Caution: ω is defined on the algebra of symmetric functions in infinitely many variables but is not compatible with specialization to finitely many variables, since the latter makes

$e_k = 0$ for $k > l$, but does not make $h_k = 0$.) By virtue of (6), for symmetric functions homogeneous of degree n, we have

$$\omega f[X] = (-1)^n f[-X].$$

Thus ω is essentially plethystic substitution of $-X$ for X, up to a sign depending on the degree. In particular we immediately see that

$$\omega^2 = 1,$$

i.e., ω is an involution. In particular, ω is an isomorphism, which yields the following result.

Proposition: The complete homogeneous symmetric functions h_λ form an R-basis of Λ_R.

1.12. First Cauchy formula and Hall inner product.

As usual, let $X = x_1 + x_2 + \cdots$, $Y = y_1 + y_2 + \cdots$. Then we have the *first Cauchy identity*

$$\begin{aligned} \Omega[XY] &= \prod_{i,j} \frac{1}{1 - x_i y_j} \\ &= \prod_i \Omega[x_i Y] \\ &= \prod_i \sum_n x_i^n h_n[Y] \\ &= \sum_{\nu_1, \nu_2, \ldots} \mathbf{x}^\nu h_\nu[Y] \\ &= \sum_\lambda m_\lambda[X] h_\lambda[Y]. \end{aligned}$$

By symmetry we also have

$$\Omega[XY] = \sum_\lambda h_\lambda[X] m_\lambda[Y],$$

which is another version of the fact we observed earlier, that the matrix giving the expansion of the h_λ by the μ_ν is symmetric.

Definition: The *Hall inner product* $\langle -, - \rangle$ is the R-bilinear form on Λ_R uniquely defined by requiring

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

i.e., we make the monomial symmetric functions and the complete homogeneous symmetric functions into dual bases.

According to this definition, the coefficient of m_λ in the monomial expansion of any f is given by $\langle h_\lambda, f \rangle$. In particular,

$$\langle h_\lambda, h_\mu \rangle = c_{\lambda\mu}, \quad \text{where} \quad h_\lambda = \sum_\mu c_{\lambda\mu} m_\mu.$$

Since $c_{\lambda\mu} = c_{\mu\lambda}$, the inner product $\langle -, - \rangle$ is *symmetric.*

Proposition: Taking $X = x_1 + x_2 + \cdots$, and considering the Hall inner product with respect to symmetric functions in $\mathbf{x}$, we have the identity

$$\langle \Omega[AX], f(\mathbf{x})\rangle = f[A].$$

Proof. Both sides are linear in f so it suffices to verify the identity for $f = m_\lambda$. Then the left hand side is the coefficient of $h_\lambda[X]$ in $\Omega[AX]$, which is $m_\lambda[A]$ by the Cauchy identity. Indeed, the Proposition can be viewed as a basis-free formulation of the Cauchy identity, as is made more precise by the following important corollary.

Corollary: Let $\{u_\lambda\}$ be a homogeneous R-basis of Λ_R. Then $\{v_\lambda\}$ is the Hall-dual basis if and only if we have the Cauchy identity

$$\Omega[XY] = \sum_\lambda u_\lambda[X] v_\lambda[Y].$$

Proof: The R-basis $\{u_\lambda\}$ has a unique dual basis $\{v_\lambda\}$, since the Hall inner product is non-degenerate on each homogeneous component $\Lambda_R^{(n)}$, which is a finitely generated free R-module. There is also a unique expansion of $\Omega[XY]$ in terms of the u_λ with some coefficients $w_\lambda[Y]$. Thus the "if" and "only if" are both equivalent to the identity $v_\lambda = w_\lambda$. By the previous Proposition, $v_\lambda[Y] = \langle \Omega[XY], v_\lambda[X]\rangle_{\mathbf{x}} = w_\lambda[Y]$.

Corollary: We have the identity

$$\langle g[X]\Omega[AX], f[X]\rangle_{\mathbf{x}} = \langle g[X], f[A+X]\rangle_{\mathbf{x}}.$$

In other words the operator of multiplication by $\Omega[AX]$ is Hall-adjoint to the plethystic shift operator sending $f[X]$ to $f[A+X]$.

Proof. Here we use another handy plethystic trick. Since both sides are linear in $g[X]$, it is sufficient to verify the identity with $g[X] = \Omega[BX]$, where B is arbitrary, since by the Cauchy formula, $\Omega[BX]$ is a linear combination of the symmetric function basis elements $m_\lambda[X]$ with coefficients $h_\lambda[B]$ that are themselves linearly independent for general B (say, $B = b_1 + b_2 + \cdots$).

Now by Proposition 1.12 we have

$$\langle \Omega[BX]\Omega[AX], f[X]\rangle = \langle \Omega[(A+B)X], f[X]\rangle = f[A+B] = \langle \Omega[BX], f[A+X]\rangle.$$

1.13. Raising operators for Schur functions.

In l variables $x_1, \dots, x_l$, using the Vandermonde identity, we may write the bialternant formula for the Schur function s_λ as

$$s_\lambda(\mathbf{x}) = \frac{\sum_{w\in S_l} \epsilon(w) w(\mathbf{x}^{\lambda+\delta})}{\prod_{i<j}(x_i - x_j)}.$$

The denominator is alternating, so we can account for the sign $\epsilon(w)$ by bringing it inside the sum to get

$$s_\lambda(\mathbf{x}) = \sum_{w\in S_l} w\left(\frac{\mathbf{x}^{\lambda+\delta}}{\prod_{i<j}(x_i - x_j)}\right) = \sum_{w\in S_l} w\left(\frac{\mathbf{x}^{\lambda}}{\prod_{i<j}(1 - x_j/x_i)}\right). \tag{8}$$

Here we have divided numerator and denominator by $\mathbf{x}^\delta$ to obtain the last formula.

Now let S_{l-1} be the subgroup of S_l consisting of permutations that fix 1. We can isolate the role of the first part λ_1 of λ by organizing the terms in (8) according to cosets of S_{l-1}

containing w. Note that the elements v of a given right coset are those with a given value of $v(1)=k$. The sum with w ranging only over S_{l-1} is equal to

$$\frac{x_1^{\lambda_1}}{\prod_{j\neq 1}(1-x_j/x_1)} s_\nu(x_2,\ldots,x_l),$$

where $\nu=(\lambda_2,\lambda_3,\ldots)$. Operating on this with a coset representative v such that $v(1)=k$, and summing over cosets, that is, over k, gives the total sum in the form

$$s_\lambda(\mathbf{x})=\sum_k \frac{x_k^{\lambda_1}}{\prod_{j\neq k}(1-x_j/x_k)} s_\nu[X-x_k]. \tag{9}$$

Here we have conveniently used the plethystic notation in terms of $X=x_1+\cdots+x_l$ for the Schur function s_ν in the variables $\mathbf{x}$ with x_k omitted.

The interesting aspect of the above formula appears when we compare it with the following partial fraction expansion for $\Omega[zX]$, where $X=x_1+\cdots+x_l$ as above, and we treat $\Omega[zX]$ as a function of z:

$$\Omega[zX]=\prod_{i=1}^{l}\frac{1}{1-zx_i}=\sum_{k=1}^{l}\frac{1}{1-zx_k}\prod_{j\neq k}\frac{1}{1-x_j/x_k}. \tag{10}$$

To fully exploit the resemblance between the terms above with those of (9), consider the following simple identity, valid for any polynomial or formal power series $f(z)$ in z whose terms involve only non-negative powers z^n.

Lemma:

$$\frac{1}{1-zu}f(z^{-1})|_{z^0}=\sum_{n\geq 0} z^n u^n f(z^{-1})|_{z^0}=\sum_{n\geq 0} u^n f(z^{-1})|_{z^{-n}}=f(u).$$

Here the vertical bar denotes the taking of a coefficient, and the statement of the Lemma contains its proof, since the last equation is essentially the *definition* of $f(u)$. More generally, we have for $r\geq 0$

$$\frac{1}{1-zu}f(z^{-1})|_{z^r}=u^r f(u),$$

as can easily be seen by replacing $f(z)$ with $z^r f(z)$ in the Lemma. Now summing this over $u=x_1,\ldots,x_l$ and using (10) yields the plethystic identity

$$f[X-z^{-1}]\Omega[zX]|_{z^r}=\sum_{k=1}^{l}\frac{x_k^r}{\prod_{j\neq k}(1-x_j/x_k)} f[X-x_k],$$

valid when $X=x_1+\cdots+x_l$. Note that the substitution of x_k for z^{-1} inside the plethystic bracket is permissible, since both are letters.

Definition: The *raising operator* B_r is the coefficient of z^r in the plethystic operator $B(z)$ defined by

$$B(z)f=f[X-z^{-1}]\Omega[zX].$$

Later we will see q-analogs of these raising operators in connection with Hall-Littlewood polynomials—the q versions are known as Jing's vertex operators. For now, our calculations above prove the raising operator formulas

$$s_\lambda[X] = B_{\lambda_1} s_{(\lambda_2,\lambda_3,\dots)}[X],$$

and hence

$$s_\lambda[X] = B_{\lambda_1} B_{\lambda_2} \cdots B_{\lambda_l}(1).$$

Since we derived the formulas for a finite but arbirtrarily number of variables $X = x_1 + \cdots + x_l$, they are actually valid as plethystic identities; in particular they are valid in infinitely many variables.

The operators B_r do not commute, but there is a commutation formula, which is of interest to derive. Introducing separate dummy variables u, v we have

$$\begin{aligned} B(u)B(v)f &= B(u)f[X - v^{-1}]\Omega[vX] \\ &= f[X - u^{-1} - v^{-1}]\Omega[v(X - u^{-1})]\Omega[uX] \\ &= \Omega[-v/u]f[X - u^{-1} - v^{-1}]\Omega[(u+v)X]. \end{aligned}$$

Since $\Omega[-v/u] = (1 - v/u)$, as u and v are letters, we have

$$uB(u)B(v)f = (u - v)f[X - u^{-1} - v^{-1}]\Omega[(u+v)X].$$

Now the right-hand side changes sign if we exchange u and v, giving the identity

$$uB(u)B(v) = -vB(v)B(u).$$

To extract information about the operators B_r, we may compare coeffecients on each side. Doing this with the coefficient of $u^{r+1}v^{s+1}$ yields the commutation relation

$$B_r B_{s+1} = -B_s B_{r+1}. \tag{11}$$

In particular, taking $r = s$ we have

$$B_r B_{r+1} = 0. \tag{12}$$

The commutation relations may be used to express any product of the operators B_r as plus-or-minus such a product with the indices decreasing, or zero. In particular this gives a definition for s_ν when ν is an integer sequence that is not a partition, as plus-or-minus a Schur function, or zero. You can easily verify that this definition agrees with what the bialternant formula would suggest for s_ν, as it should, since we did not actually use the fact that λ was a partition in deriving the raising operator formula from the bialternant formula.

1.14. Cauchy formula.

Let $Y = y_1 + \cdots + y_l$. We may write $s_\lambda[X] = B_{\lambda_1} \cdots B_{\lambda_l}(1)$ as

$$s_\lambda[X] = B(y_1) \cdots B(y_l) 1|_{\mathbf{y}^\lambda}.$$

Now it is an easy exercise to evaluate the above operator product, much as we did for $B(u)B(v)$ in 1.13, to obtain

$$\begin{aligned} s_\lambda[X] &= \Omega[XY]\Omega[-\sum_{i<j} y_j/y_i]|_{\mathbf{y}^\lambda} \\ &= \Omega[XY]\prod_{i<j}(1-y_j/y_i)|_{\mathbf{y}^\lambda} \\ &= \Omega[XY]\prod_{i<j}(y_i-y_j)|_{\mathbf{y}^{\lambda+\delta}}. \end{aligned}$$

Now recall that the coefficient of $s_\lambda[Y]$ in any symmetric function $f[Y]$ is found by multiplying $f[Y]$ by $a_\delta(\mathbf{y})$ and extracting the coefficient of $\mathbf{y}^{\lambda+\delta}$. In view of this the above identity reads

$$\Omega[XY] = \sum_\lambda s_\lambda[X]s_\lambda[Y].$$

Of course if we obtain this with $Y = y_1 + \cdots + y_l$, for a fixed l, the sum only ranges over terms with $l(\lambda) \leq l$, all others having $s_\lambda[Y] = 0$. However, l was arbitrary, so the identity is valid as a plethystic idenitity, with the sum ranging over all λ. This is the *Cauchy formula* for Schur functions. By 1.12, it has the following consequence.

Proposition: The Schur functions form an orthonormal basis of Λ_R with respect to the Hall inner product $\langle -, - \rangle$.

1.15. Pieri formulas.

We may derive expressions for elementary and complete homogeneous symmetric functions in terms of Schur functions from formulas known as *Pieri formulas* for the multiplication of a Schur function by e_k or h_k. Viewing $\Omega[\pm uX]$ as a generating function for elementary or homogeneous symmetric functions, we can obtain the Pieri formulas from commutation formulas for the raising operator generating function $B(z)$ with the operator of multiplication by $\Omega[\pm uX]$. We begin with $\Omega[-uX]$, which is the easier case.

We have

$$\begin{aligned} B(z)\Omega[-uX]f &= \Omega[-u(X-z^{-1})]f[X-z^{-1}]\Omega[zX] \\ &= \Omega[u/z]\Omega[-uX]B(z)f, \end{aligned}$$

or

$$\Omega[-uX]B(z) = (1-u/z)B(z)\Omega[-uX]$$

as an operator identity, where $\Omega[-uX]$ stands for the operator of multiplication by itself. Taking the coefficient of $(-u)^k z^r$ yields the operator identity

$$e_k B_r = (B_r e_k + B_{r+1}e_{k-1}). \tag{13}$$

We'll need an auxilliary result to properly handle B_0:

$$B_0\Omega[-uX] = \Omega[u/z]\Omega[-uX]\Omega[zX]|_{z^0} = \frac{1}{1-uz^{-1}}\Omega[uX]\Omega[-zX]|_{z^0}.$$

Now by Lemma 1.13, the last expression is equal to $\Omega[uX]\Omega[-uX]=1$, so we have

$$B_0(e_k)=\begin{cases}1 & k=0\\ 0 & \text{otherwise.}\end{cases}$$

Combining this with (13) for $r=0$, $f=1$ we see by induction on k that

$$B_1^k(1)=e_k.$$

Now let us use (13) to compute $e_k s_\lambda[X]$. Applying (13) repeatedly to move the e_k to the right we obtain

$$e_k B_{\lambda_1}\cdots B_{\lambda_l}(1)=\sum_\nu B_{\nu_1}\cdots B_{\nu_l}(e_{k-j})=\sum_\nu B_{\nu_1}\cdots B_{\nu_l}B_1^{k-j}(1),$$

where ν ranges over all integer sequences obtained from λ by increasing some (or none, or all) of the parts by 1, and j is the number of parts so increased.

By (12), any terms in the above sum for which $(\nu,1^{k-j})$ is not a partition will vanish. This leaves one term for each partition $\mu=(\nu,1^{k-j})$ obtainable from λ by increasing some parts by 1 and adding some parts of 1 at the end (effectively increasing extraneous parts of zero to 1). Such a partition is said to differ from λ by the addition of a *vertical k-strip* (k being the size), a terminology whose significance in terms of partition diagrams is apparent. Denoting the condition that μ is λ plus a vertical k-strip by $\mu/\lambda\in V_k$, we may express our first Pieri formula as follows.

Proposition:

$$e_k s_\lambda=\sum_{\mu/\lambda\in V_k} s_\mu.$$

For the second Pieri formula we rearrange our previous operator identity to read

$$\Omega[uX]B(z)=\frac{1}{1-u/z}B(z)\Omega[uX],$$

obtaining

$$h_k B_r=\sum_{j=0}^{k} B_{r+j}h_j.$$

Note also that $h_k=B_k(1)$, directly from the definition of B_k. Hence

$$h_k s_\lambda[X]=h_k B_{\lambda_1}\cdots B_{\lambda_l}(1)=\sum_{j_1+\cdots+j_{l+1}=k} B_{\lambda_1+j_1}\cdots B_{\lambda_l+j_l}B_{j_{l+1}}(1).$$

As we have seen, each term on the right is plus-or-minus a Schur function, or zero. I claim there is perfect cancellation of all terms with some index i such that $\lambda_i+j_i>\lambda_{i-1}$. Consider such a term and let i be the greatest such index. Compare

$$B_{\lambda_1+j_1}\cdots B_{\lambda_{i-1}+j_{i-1}}B_{\lambda_i+j_i}\cdots 1;$$
$$B_{\lambda_1+j_1}\cdots B_{\lambda_i+j_i-1}B_{\lambda_{i-1}+j_{i-1}+1}\cdots 1.$$

The second term has $\lambda_i+j_i-1 \geq \lambda_{i-1}$ and $\lambda_{i-1}+j_{i-1}+1 > \lambda_i$ so $\lambda_i+j_i-1 = \lambda_{i-1}+j'_{i-1}$, say, and $\lambda_{i-1}+j_{i-1}+1 = \lambda_i + j'_i$. Written this way the second term also satisfies $\lambda_i + j'_i > \lambda_{i-1}$, with i the greatest such index, and the same construction applied to it returns us to the first term. Thus terms of this type pair off with opposite sign, except when $\lambda_i + j_i = \lambda_{i-1}+1$, in which case the term is zero.

The surviving terms have $\lambda + \mathbf{j} = \mu$, where μ is λ plus a *horizontal* k-strip (conjugate concept to vertical k-strip). This gives the second Pieri formula.

Proposition:

$$h_k s_\lambda = \sum_{\mu/\lambda \in H_k} s_\mu.$$

Note that either Pieri formula completely determines every s_λ, by providing an inductive rule for expressing the e_λ's (respectively the h_λ's) in terms of Schur functions. From the conjugate symmetry between the two formulas we obtain the following corollary.

Corollary: We have $\omega s_\lambda = s_{\lambda'}$, or $s_\lambda[-X] = (-1)^{|\lambda|} s_{\lambda'}[X]$.

Another corollary, this one to the second Pieri rule, is the combinatorial formula for the *Kostka coefficients* $K_{\lambda\mu}$ giving Schur functions in terms of monomials. Recall the adjunction formula $\langle f[X+A], g\rangle = \langle f, \Omega[AX]g\rangle$ from 1.10. In particular, taking $A = u$ a single letter, and comparing coefficients of u^k, we have

$$\langle f[X+u], g\rangle|_{u^k} = \langle f, h_k g\rangle.$$

(This is a plethystically disguised version of the dual-basis relationship between the m_λ's and h_λ's.) By the second Pieri formula,

$$\langle s_\nu[X+u], s_\lambda\rangle|_{u^k} = \langle s_\nu, h_k s_\lambda\rangle$$

is 1, if ν/λ is a horizontal k-strip, and zero otherwise. In other words,

$$s_\nu[X+u] = \sum_k u^k \sum_{\lambda\colon \nu/\lambda \in H_k} s_\lambda.$$

Iterating this, we see that the coefficient of $\mathbf{u}^\kappa$ in $s_\nu[u_1 + \cdots + u_l]$ is equal to the number of sequences

$$\emptyset = \lambda^{(0)} \subseteq \lambda^{(1)} \subseteq \cdots \subseteq \lambda^{(l)} = \nu \quad \text{such that } \lambda^{(i)}/\lambda^{(i-1)} \in H_{\kappa_i} \text{ for all } i = 1, \ldots, l.$$

Filling in the difference diagram between $\lambda^{(i-1)}$ and $\lambda^{(i)}$ with i's, we get a filling of the diagram of ν with κ_1 1's, κ_2 2's, and so on, with non-decreasing rows and columns, in which the horizontal strip condition on the positions of each i means that *columns* increase *strictly*. Such a filling is a *column strict Young tableau of shape* ν and *content* κ.

Corollary: The Kostka coefficient $K_{\lambda\mu}$ given by the expansion

$$s_\lambda = \sum_\mu K_{\lambda\mu} m_\mu$$

is equal to the number of column-strict Young tableaux of shape λ and content μ.

It also follows that the number of column-strict Young tableaux is symmetric with respect to permutations of the content vector κ. *Exercise:* prove this directly and combinatorially by considering what happens for adjacent transpositions in κ.

1.16. Jacobi-Trudi formula.

Consider the problem of expanding Schur functions in terms of the complete homogeneous symmetric functions h_μ. By Hall duality, the coefficient of h_μ in s_λ is given by

$$\langle s_\lambda, m_\mu \rangle,$$

which is also the coefficient of s_λ in m_μ. We find the latter by the usual device for Schur function expansions to be

$$\begin{aligned} m_\mu a_\delta|_{\mathbf{x}^{\lambda+\delta}} &= \sum_{w \in S_l} \epsilon(w) m_\mu|_{\mathbf{x}^{\lambda+\delta-w(\delta)}} \\ &= \sum_{w \in S_l} \epsilon(w) \begin{cases} 1 & \text{if } \mathbf{x}^{\lambda+\delta-w(\delta)} \approx \mathbf{x}^\mu, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

This implies the *Jacobi-Trudi identity*, given as follows.

Proposition:

$$s_\lambda = \sum_{w \in S_l} \epsilon(w) h_{\lambda+\delta-w(\delta)} = \det \begin{bmatrix} h_{\lambda_1} & h_{\lambda_1+1} & h_{\lambda_1+2} & \cdots \\ h_{\lambda_2-1} & h_{\lambda_2} & h_{\lambda_2+1} & \cdots \\ & & \vdots & \\ \cdots & & h_{\lambda_l-1} & h_{\lambda_l} \end{bmatrix},$$

with the convention that $h_0 = 1$ and $h_{-k} = 0$ for $-k$ negative.

The format of the Jacobi-Trudi matrix on the right is: $h_{\lambda_1}, \ldots, h_{\lambda_l}$ on the diagonal, indices changing by 1 from column to column within each row. Here l can be any integer greater than or equal to the number of parts of λ. Applying ω gives the analog for elementary symmetric functions.

Corollary:

$$s_{\lambda'} = \det \begin{bmatrix} e_{\lambda_1} & e_{\lambda_1+1} & e_{\lambda_1+2} & \cdots \\ e_{\lambda_2-1} & e_{\lambda_2} & e_{\lambda_2+1} & \cdots \\ & & \vdots & \\ \cdots & & e_{\lambda_l-1} & e_{\lambda_l} \end{bmatrix}.$$