D. Laksov, A. Lascoux, P. Pragacz, and A. Thorup

The LLPT Notes

Edited by A. Thorup, 1995

With corrections and additions suggested by D. Grinberg 2017/18

*not part of the 1995 edition of the notes.

The LLPT Notes.

In 1991, a project on cohomology and algebraic geometry was initiated by Dan Laksov (Stockholm), Alain Lascoux (Paris), Piotr Pragacz (Torun) and Anders Thorup (Copenhagen). Part of the project was a presentation of the necessary theory of symmetric polynomials. This set of notes is essentially the first, very preliminary, draft of this part of the project. It was developed in close cooperation between the four writers. The draft in its present form was edited by A. Thorup, and used in a course at the University of Copenhagen in 1995; it has never been approved by the three other writers.

20. December 1995
Anders Thorup

Permutations and symmetric functions

1. The length function.

(1.1) Setup. Fix an *alphabet* A, that is, a finite totally ordered set; the elements of A are called *letters*. The number n will always denote the *size* of the alphabet (the number of letters), $n := |A|$. Every letter a which is not the last letter has a successor, denoted a'. It is often convenient to enumerate the letters, say $A = \{a_1, \dots, a_n\}$. Since A is assumed to be totally ordered, there is a unique enumeration such that the a_i are in increasing order. In this enumeration, the successor of a_i is $a_i' = a_{i+1}$ for $i = 1, \dots, n-1$.

Unless otherwise specified, a *permutation* will mean a permutation of the letters, that is, a bijective map from A onto itself. The group of all permutations will be denoted $\mathfrak{S}(A)$. The multiplication in the group is the usual composition of maps: $\sigma\tau$ maps x to $\sigma(\tau(x))$. The identity map is the unique order preserving permutation of A, denoted 1_A or simply 1. We denote by $\omega = \omega_A$ the unique order reversing permutation of A. Obviously, ω is an involution, that is, $\omega^2 = 1$. Sometimes ω is called the *maximal permutation* of A.

The transposition that interchanges two letters a and b will be denoted $\varepsilon_{a,b}$. Transpositions that interchange two neighbors (with respect to the given order in A) are said to be *simple*. So the simple transpositions are the transpositions $\varepsilon_a := \varepsilon_{a,a'}$ where a is not the last letter. The number of simple transpositions is one less than the number of letters.

Assume that the letters are indexed a_i in the natural order. It is easily seen that the product $\varepsilon_{a_1}\varepsilon_{a_2}\cdots\varepsilon_{a_{n-1}}$ is the n-cycle $(a_1, \dots, a_n)$. It follows, for instance by an inductive argument, that

$$\omega_A = (\varepsilon_{a_1}\cdots\varepsilon_{a_{n-1}})\cdots(\varepsilon_{a_1}\varepsilon_{a_2})\varepsilon_{a_1}.$$

(1.2) Definition. A pair (a, b) of letters will be called an *inversion* for the permutation μ if $a < b$ and $\mu a > \mu b$. The number of inversions of μ is denoted by $\ell(\mu)$ and called the *length* of μ.

(1.3) Lemma. *For any permutation μ, the following six assertions hold:*

(1) *We have equality $\ell(\mu) = 0$ if and only if $\mu = 1$. Moreover, if $\mu \neq 1$, then there is an inversion for μ of the form (a, a').*
(2) *We have equality $\ell(\mu) = 1$ if and only if μ is a simple transposition.*
(3) *For a simple transposition ε_a we have that*

$$\ell(\mu\varepsilon_a) = \begin{cases} \ell(\mu) + 1 & \text{if } \mu a < \mu a', \\ \ell(\mu) - 1 & \text{if } \mu a > \mu a'. \end{cases}$$

More precisely, if $\mu a < \mu a'$*, then the inversions for* $\mu\varepsilon_a$ *are the pairs* $(\varepsilon_a b, \varepsilon_a c)$ *where* (b, c) *is an inversion for* μ *and the pair* (a, a')*, and if* $\mu a > \mu a'$*, then the inversions for* $\mu\varepsilon_a$ *are the pairs* $(\varepsilon_a b, \varepsilon_a c)$ *where* (b, c) *is an inversion for* μ *different from* (a, a')*.*

(4) *We have the inequality* $\ell(\mu) \le n(n-1)/2$*, and equality holds if and only if* $\mu = \omega$*.*
(5) *We have that* $\ell(\mu) = \ell(\mu^{-1})$*.*
(6) *We have that* $\ell(\omega\mu) + \ell(\mu) = \ell(\omega)$*.*

Proof. The five first assertions are easily checked. The sixth follows from the equation $\ell(\omega) = n(n-1)/2$, because each pair (a, b) of letters with $a < b$ is an inversion for exactly one of the permutations μ and $\omega\mu$. □

(1.4) Proposition. *The group* $\mathfrak{S}(A)$ *is generated by the simple transpositions. In fact, any permutation* μ *is a product of* $\ell(\mu)$ *simple transpositions.*

Proof. We prove the second assertion by induction on $\ell(\mu)$. If $\ell(\mu) = 0$, then $\mu = 1$. Hence the assertion holds when $\ell(\mu) = 0$. Assume that $\ell(\mu) = l > 0$ and that the assertion holds for permutations of length $l - 1$. Since $\mu \neq 1$, there is an inversion for μ of the form (a, a'). It follows from (1.3)(3) that $\ell(\mu\varepsilon_a) = l - 1$. Hence $\mu\varepsilon_a$ is a product of $l - 1$ simple transpositions. Therefore $\mu = (\mu\varepsilon_a)\varepsilon_a$ is a product of l simple transpositions. □

(1.5) Note. The proof of Proposition (1.4) is constructive. Let μ be a permutation, and consider the *direct representation* of μ:

$$\mu = (b_1 b_2 \dots b_n). \tag{1.5.1}$$

By definition, if $A = \{a_1, \dots, a_n\}$ with the letters a_i in increasing order, then μ is determined from the sequence in (1.5.1) by $\mu a_i = b_i$. If the b_i are in increasing order, then $b_i = a_i$ and $\mu = 1$. If $\mu \neq 1$, then there is an index $j < n$ such that $b_j > b_{j+1}$. Now interchange in the sequence b_j and b_{j+1}. The new sequence represents the permutation $\mu\varepsilon_{a_j}$. Continue the process until the b_i's appear in increasing order. The sequence obtained at the end represents the identity permutation 1. Hence we have an equation $1 = \mu\varepsilon_{a_{j_1}} \cdots \varepsilon_{a_{j_r}}$. Thus $\mu = \varepsilon_{a_{j_r}} \cdots \varepsilon_{a_{j_1}}$. It follows from the proof of Proposition (1.4) that $r = \ell(\mu)$.

(1.6) Proposition. *For any presentation* $\mu = \tau_1 \cdots \tau_r$ *of* μ *as a product of* r *simple transpositions, we have that* $\ell(\mu) \le r$ *and* $\ell(\mu) \equiv r \pmod 2$*.*

Proof. We prove the Proposition by induction on the number r of factors in the presentation of μ. If $r = 0$, then $\mu = 1$ and the two assertions hold. Assume that $r > 0$ and that the two assertions hold for all presentations with $r - 1$ factors. Consider the presentation $\nu := \tau_1 \cdots \tau_{r-1}$. Since $\mu = \nu\tau_r$, it follows from Lemma (1.3)(3) that $\ell(\mu) = \ell(\nu) \pm 1$. In particular, we have that $\ell(\mu) \le \ell(\nu) + 1$ and $\ell(\mu) \equiv \ell(\nu) + 1 \pmod 2$. By the induction hypothesis, we have that $\ell(\nu) \le r - 1$ and $\ell(\nu) \equiv r - 1 \pmod 2$. Therefore, the two assertions hold for the presentation of μ. □

(1.7) Corollary. *For any two permutations μ and ν in $\mathfrak{S}(A)$, we have the inequality $\ell(\mu) + \ell(\nu) \geq \ell(\mu\nu)$.*

Proof. The Corollary follows immediately from the two previous Propositions. ∎

(1.8) Definition. The *signature* of a permutation μ, denoted $\operatorname{sign}\mu$, is the number

$$\operatorname{sign}\mu := (-1)^{\ell(\mu)}.$$

It follows from Proposition (1.6) that the signature is a homomorphism of groups,

$$\mathfrak{S}(A) \to \{\pm 1\}.$$

By (1.3), the signature of a simple transposition is equal to -1. It is easy to see that the signature of any transposition is equal to -1. It follows that the map sign is independent of the given order of the letters of A.

2. The Coxeter–Moore relations.

(2.1) Definition. A sequence $(\tau_1, \dots, \tau_r)$ of simple transpositions will be called a *presentation* of the permutation μ in $\mathfrak{S}(A)$ if $\mu = \tau_1 \cdots \tau_r$. By Proposition (1.6), for any presentation $(\tau_1, \dots, \tau_r)$ of μ, we have that $r \geq \ell(\mu)$. The presentation is said to be *minimal* if $r = \ell(\mu)$. Minimal presentations of a given permutation μ exist by Proposition (1.4).

(2.2) Lemma. *Let $(\tau_1, \dots, \tau_r)$ be a minimal presentation of a permutation μ. For $s = 1, \dots, r$, let b_s be the letter for which $\tau_s = \varepsilon_{b_s}$. Then the inversions for μ are the pairs $\big(\tau_r \cdots \tau_{s+1}(b_s), \tau_r \cdots \tau_{s+1}(b'_s)\big)$ for $s = 1, \dots, r$.*

In particular (when $r \geq 1$), the pair (b_r, b'_r) is an inversion for μ.

Proof. We shall prove the Lemma by induction on $r = \ell(\mu)$. Clearly, the assertion holds for $r = 0$. Assume that $r \geq 1$ and that the assertion holds for all permutations of length $r - 1$. It follows from Lemma (1.3)(3) that $\nu := \mu\tau_r = \tau_1 \cdots \tau_{r-1}$ has length $r - 1$. By the induction hypothesis, we have that the inversions of ν are the pairs $\big(\tau_{r-1} \cdots \tau_{s+1}(b_s), \tau_{r-1} \cdots \tau_{s+1}(b'_s)\big)$ for $s = 1, \dots, r-1$. Since $\ell(\nu\tau_{b_r}) = \ell(\mu) = \ell(\nu) + 1$, it follows from Lemma (1.3)(3) that the assertion holds for μ. □

(2.3) Lemma (The exchange property). *Let $(\tau_1, \dots, \tau_r)$ and $(\sigma_1, \dots, \sigma_r)$, for $r \geq 1$, be two minimal presentations of the same permutation μ. Then, for some $q = 1, \dots, r$, there is a presentation of μ of the form $(\sigma_1, \tau_1, \dots, \widehat{\tau_q}, \dots, \tau_r)$, where the hat indicates an omitted transposition.*

Proof. Assume $\sigma_1 = \varepsilon_a$ and $\tau_i = \varepsilon_{b_i}$ for $i = 1, \dots, r$. By Lemma (2.2), we have that (a, a') is an inversion for $\mu^{-1} = \sigma_r \cdots \sigma_1$. By the same Lemma, since $\mu^{-1} = \tau_r \cdots \tau_1$, there is a q such that $a = \tau_1 \cdots \tau_{q-1}(b_q)$ and $a' = \tau_1 \cdots \tau_{q-1}(b'_q)$. It follows that the permutation $\tau := (\tau_1 \cdots \tau_{q-1})\tau_q(\tau_1 \cdots \tau_{q-1})^{-1}$ interchanges a and a'. However, τ is conjugate to the transposition τ_q and hence τ is a transposition. Since τ interchanges a and a', it follows that $\tau = \varepsilon_a = \sigma_1$. As a consequence, we have the equation,

$$\sigma_1\tau_1 \cdots \tau_{q-1} = \tau_1 \cdots \tau_{q-1}\tau_q.$$

Clearly, the assertion of the Lemma is obtained after multiplication by $\tau_{q+1} \cdots \tau_r$. □

(2.4) Remark. Every transposition τ is an involution. In particular, for every simple transposition ε_a we have that $\varepsilon_a^2 = 1$. Consider a second simple transposition ε_b with $b \neq a$. Then we have the relations,

$$\begin{aligned} \varepsilon_a\varepsilon_b\varepsilon_a &= \varepsilon_b\varepsilon_a\varepsilon_b \quad \text{if } a \text{ and } b \text{ are neighbors,} \\ \varepsilon_a\varepsilon_b &= \varepsilon_b\varepsilon_a \quad \text{if } a \text{ and } b \text{ are not neighbors.} \end{aligned}$$

Indeed, if a and b are neighbors, we may assume that $b = a'$, and then $\varepsilon_a\varepsilon_b\varepsilon_a = \varepsilon_{a,b'} = \varepsilon_b\varepsilon_a\varepsilon_b$. If a and b are not neighbors, then the permutations ε_a and ε_b are disjoint, and hence they commute.

In general, an ordered set $g_1, \dots, g_{n-1}$ of elements in a semi-group G are said to satisfy the *Coxeter–Moore relations* if

(1) $g_j g_k = g_k g_j \quad \text{if} \quad |k - j| > 1$,
(2) $g_j g_k g_j = g_k g_j g_k \quad \text{if} \quad |k - j| = 1$.

In particular, the simple transpositions ε_{a_j}, with the letters a_j in increasing order, satisfy the Coxeter–Moore relations.

(2.5) Definition. Two presentations are said to be *Coxeter–Moore equivalent* if one can be obtained from the other by a finite number (possibly none) of the following two *allowable replacements*: Given a presentation $(\tau_1, \dots, \tau_r)$. If two consecutive transpositions τ_i and τ_{i+1} are disjoint, then it is allowed to replace τ_i, τ_{i+1} by τ_{i+1}, τ_i. If for three consecutive transpositions τ_i, τ_{i+1}, and τ_{i+2}, we have that $\tau_i = \tau_{i+2} = \varepsilon_a$ and $\tau_{i+1} = \varepsilon_b$, where a and b are neighboring letters, then it is allowed to replace $\tau_i, \tau_{i+1}, \tau_{i+2}$ by $\tau_{i+1}, \tau_i, \tau_{i+1}$.

Clearly, two Coxeter–Moore equivalent presentations contain the same number of simple transpositions. As the simple transpositions satisfy the Coxeter–Moore relations, it follows that two Coxeter–Moore equivalent presentations are presentations of the same permutation.

(2.6) Proposition. *Any two minimal presentations of the same permutation are Coxeter–Moore equivalent.*

Any presentation which is not a minimal presentation is Coxeter–Moore equivalent to a presentation in which two consecutive transpositions are equal.

Proof. To prove the first assertion, consider two minimal presentations, labelled α and β, of a permutation μ,

$$\alpha = (\tau_1, \dots, \tau_r), \qquad \beta = (\sigma_1, \dots, \sigma_r).$$

Then $r = \ell(\mu)$. We have to prove that α and β are (Coxeter–Moore) equivalent.

Clearly, the equivalence holds when $r = 1$. Proceed by induction on $r = \ell(\mu)$. Assume that $r \geq 2$ and that the assertion holds for minimal presentations of permutations of length $r - 1$.

The following observation will be useful: The equivalence holds if $\sigma_r = \tau_r$. This follows by applying the induction hypothesis to the two minimal presentations $(\tau_1, \dots, \tau_{r-1})$ and $(\sigma_1, \dots, \sigma_{r-1})$ of $\mu\sigma_r$. Similarly, the equivalence holds if $\sigma_1 = \tau_1$.

Now, by the exchange property, there is, for some $q = 1, \dots, r$, a presentation of μ of the form $(\sigma_1, \tau_1, \dots, \widehat{\tau_q}, \dots, \tau_r)$. By the observation, the new presentation is equivalent to β. Hence we may assume that β is the new presentation, $\beta = (\sigma_1, \tau_1, \dots, \widehat{\tau_q}, \dots, \tau_r)$. Again, by the observation, the equivalence of α and β holds if $q < r$. Hence we may assume that $q = r$, that is,

$$\alpha = (\tau_1, \dots, \tau_r), \qquad \beta = (\sigma_1, \tau_1, \dots, \tau_{r-1}).$$

Again, by the exchange property, there is a presentation of μ of the form $(\tau_1, \dots)$ where the dots indicate the transpositions of β with one omitted. If it was the first transposition, σ_1, that was omitted, then the new presentation would have the form $(\tau_1, \tau_1, \dots)$, contradicting the minimality. Therefore, the new presentation is of the form $(\tau_1, \sigma_1, \tau_1, \dots, \widehat{\tau_s}, \dots, \tau_{r-1})$ for some $s = 1, \dots, r-1$. Again, by the observation, this new presentation is equivalent to α, and hence we may assume that α is this new presentation, that is, for some $s = 1, \dots, r-1$,

$$\alpha = (\tau_1, \sigma_1, \tau_1, \dots, \widehat{\tau_s}, \dots, \tau_{r-1}), \quad \beta = (\sigma_1, \tau_1, \dots, \tau_{r-1}).$$

The equivalence holds if $r = 2$. Indeed, if $r = 2$ then $\alpha = (\tau_1, \sigma_1)$ and $\beta = (\sigma_1, \tau_1)$ and we have the equation $\mu = \tau_1\sigma_1 = \sigma_1\tau_1$. Since $\ell(\mu) = 2$, we have that $\tau_1 \neq \sigma_1$. Clearly, then the equation $\tau_1\sigma_1 = \sigma_1\tau_1$ for transpositions implies that τ_1 and σ_1 are disjoint. Thus β is obtained from α by an allowable replacement, and hence α and β are equivalent.

Thus we may assume that $r \geq 3$. By the observation, if $s < r-1$, then the equivalence holds. So assume that $s = r-1$. Then $s \geq 2$, and so the presentation α has the form $\alpha = (\tau_1, \sigma_1, \tau_1, \dots)$. Since α is minimal, we have that $\tau_1 \neq \sigma_1$ and $\tau_1\sigma_1 \neq \sigma_1\tau_1$. Therefore, the simple transpositions σ_1 and τ_1 are associated to neighboring letters. Thus with an allowable replacement we may obtain from α a presentation of the form $(\sigma_1, \tau_1, \sigma_1, \dots)$. By the observation, the replaced presentation is equivalent to β. Therefore α and β are equivalent and the first assertion of the Proposition has been proved.

The second assertion is proved by induction on the number r of factors in the presentation. Clearly, the number of factors of a non-minimal presentation is at least 2. Moreover, a presentation (τ_1, τ_2) is minimal unless $\tau_1 = \tau_2$. Hence the assertion holds when $r = 2$. Assume that $r > 2$ and that the assertion holds for presentations with $r-1$ factors. Let $\alpha = (\tau_1, \dots, \tau_r)$ be a non-minimal presentation of μ. Then $r > \ell(\mu)$. Now $\beta := (\tau_1, \dots, \tau_{r-1})$ is a presentation of $\nu := \mu\tau_r$. Clearly, if β is non-minimal then, by the induction hypothesis, the assertion holds for α. So assume that the presentation β is minimal. Then $\ell(\nu) = r-1$. Since $\mu = \nu\tau_r$ is not of length r, it follows from Lemma (1.3)(3) that $\ell(\mu) = r-2$. Hence there is a minimal presentation $(\sigma_1, \dots, \sigma_{r-2})$ of μ. Since $\nu = \mu\tau_r$, it follows that $\delta := (\sigma_1, \dots, \sigma_{r-2}, \tau_r)$ is a presentation of ν. Moreover, the presentation δ is minimal because $\ell(\nu) = r-1$. Therefore, by the first part of the Proposition, the presentation β is Coxeter–Moore equivalent to δ. It follows that α is Coxeter–Moore equivalent to the presentation $(\sigma_1, \dots, \sigma_{r-2}, \tau_r, \tau_r)$. Hence the assertion holds for α.

Thus both assertions of the Proposition have been proved. $\square$

3. The Bruhat–Ehresman order.

4. Young subgroups.

5. Symmetric polynomials.

(5.1) Setup. Fix a commutative ring R. Consider the ring $R[A]$ of polynomials with coefficients in R in the letters of A. With the letters of A in increasing order $a_1, \dots, a_n$, the *monomials* in $R[A]$ are the products,

$$a_1^{j_1} \cdots a_n^{j_n},$$

where $(j_1, \dots, j_n)$ is a sequence of n nonnegative integers. The monomials form an R-basis of $R[A]$ since, by definition, every polynomial f is an R-linear combination,

$$f = \sum f_{j_1,\dots,j_n} a_1^{j_1} \cdots a_n^{j_n},$$

with uniquely determined coefficients $f_{j_1,\dots,j_n}$ in R.

The notation is simplified through the use of multi indices. A *multi index* J of *size* n is a sequence $J = (j_1, \dots, j_n)$ of n nonnegative integers. Associate with J the monomial,

$$a^J := a_1^{j_1} \cdots a_n^{j_n}.$$

Then the coefficients of a polynomial f are the elements f_J of R, for all multi indices J. If $f_J \neq 0$, then the monomial a^J is said to *appear* in f. The *degree* of a multi index J is the sum of the entries, $\|J\| = j_1 + \cdots + j_n$, and the *degree* of the monomial a^J is the sum of the exponents, $\|J\|$. If $f \neq 0$ is a polynomial, then the *degree* and the *order* of f are, respectively, the maximal and the minimal degree of the monomials appearing in f. The polynomial f is said to be *homogeneous of degree d* if all monomials appearing in f are of degree d. According to this definition, the zero polynomial is homogeneous of every degree d.

Multi indices are ordered (graded lexicographically) as follows: we write $I < J$ if either $\|I\| < \|J\|$ or $\|I\| = \|J\|$ and there is a $p = 1, \dots, n$ such that $i_q = j_q$ for $q = 1, \dots, p-1$ and $i_p < j_p$. Clearly, the order is a total order, and there is only a finite number of multi indices less than a given. The smallest multi index is $(0, \dots, 0)$. With respect to addition of multi indices, we have that if $I < J$ then $I + K < J + K$.

According to the order on the multi indices, there is a total order on the monomials: $a^I < a^J$ if $I < J$. (Note that the order on the a_i's as monomials of degree 1 is the reverse of the given order on the a_i's as letters.) The *leading monomial* of a non-zero polynomial f is the biggest monomial a^J appearing in f; the corresponding coefficient f_J is called the *leading coefficient* and $f_J a^J$ is called the *leading term*. Addition of multi indices corresponds to multiplication of monomials. Hence, if $a^I \leq a^J$ and $a^L \leq a^K$ and one of the inequalities is strict, then $a^I a^L < a^J a^K$. It follows that if f and g are nonzero polynomials with leading terms $f_J a^J$ and $g_K a^K$, then every monomial appearing in the product fg is at most equal to a^{J+K}; moreover, if $f_J g_K \neq 0$, then $f_J g_K a^{J+K}$ is the leading term of fg.

A second (partial) order on multi indices of size n is defined as follows: We write $I \prec J$ (or $J \succ I$) if $i_p \leq j_p$ for $p = 1, \dots, n$ with at least one strict inequality. Clearly, for the corresponding monomials we have $I \preceq J$ iff a^I divides a^J.

(5.2) Note. The number of monomials of degree d in $n > 0$ variables is equal to the binomial coefficient,

$$\binom{d+n-1}{n-1}.$$

Indeed, the number is equal to the number of multi indices $J = (j_1, \dots, j_n)$ such that $\|J\| = j_1 + \cdots + j_n = d$. Now, there is a bijective map $J \mapsto J'$ from the set of all multi indices onto the set of all strictly increasing multi indices, defined by

$$j'_p := j_1 + \cdots + j_p + (p-1) \quad \text{for } p = 1, \dots, n.$$

Under this map, we have that $\|J\| = d$ if and only if $j'_n = d + n - 1$. Hence, the number of multi indices J such that $\|J\| = d$ is equal to the number of strictly increasing multi indices $(k_1, \dots, k_n)$ such that $k_n = d + n - 1$. Clearly, the latter number is equal to the number of subsets with $n - 1$ elements of the $d + n - 1$ integers $0, 1, \dots, d + n - 2$.

(5.3) Remark. The algebra $R[A]$ of polynomials has the following universal property: Given a homomorphism $R \to S$ of commutative rings and a set $\alpha = (\alpha_1, \dots, \alpha_n)$ of n elements of S. Then there is a unique homomorphism of R-algebras $R[A] \to S$ such that $a_i \mapsto \alpha_i$. The map is called *evaluation* of polynomials at α. The value of f under this map is denoted

$$f(\alpha) = f(\alpha_1, \dots, \alpha_n);$$

it is said to be obtained by the *substitution* $a_i \mapsto \alpha_i$ (or $a_i := \alpha_i$) for $i = 1, \dots, n$.

(5.4) Definition. The symmetric group $\mathfrak{S}(A)$ acts on the algebra $R[A]$ of polynomials. Indeed, any permutation σ can be viewed as a permutation of the variables of $R[A]$ and as such it extends uniquely to an R-algebra automorphism of $R[A]$, denoted $f \mapsto \sigma f$. It is given by the substitution $\sigma f = f(\sigma a_1, \dots, \sigma a_n)$. Obviously, we have the equation $(\sigma\tau)f = \sigma(\tau f)$ for permutations σ and τ.

A polynomial f in $R[A]$ is called *symmetric* if it is invariant under the action of $\mathfrak{S}(A)$, that is, if

$$\sigma(f) = f \text{ for all } \sigma \in \mathfrak{S}(A). \tag{5.4.1}$$

The symmetric polynomials in $R[A]$ form an R-subalgebra, denoted $\mathrm{Sym}_R[A]$.

A polynomial f is called *anti-symmetric* if it is semi-invariant under the action of $\mathfrak{S}(A)$ in the sense that

$$\sigma(f) = \mathrm{sign}(\sigma) f \text{ for all } \sigma \in \mathfrak{S}(A). \tag{5.4.2}$$

Clearly, the anti-symmetric polynomials in $R[A]$ form a module over the ring $\mathrm{Sym}_R[A]$ of symmetric polynomials. In particular, a product of a symmetric polynomial and an anti-symmetric polynomial is an anti-symmetric polynomial. Similarly, a product of two anti-symmetric polynomials is a symmetric polynomial.

The symmetric group is generated by the simple transpositions. It follows that a polynomial is symmetric if it is unchanged whenever two neighbor variables are interchanged. Similarly, a polynomial is anti-symmetric if it changes sign whenever two neighbor variables are interchanged.

(5.5). If σ is a permutation in $\mathfrak{S}(A)$, then $\sigma(a^J)$ is the monomial $a^{\sigma J}$, where σJ is the multi index obtained from J by permuting the entries as follows: With the given enumeration of A, we can identify $\mathfrak{S}(A)$ with the symmetric group $\mathfrak{S}_n$ of permutations of the numbers $1, 2, \dots, n$. Then σJ is obtained from J by moving, for $p = 1, \dots, n$, the entry j_p from its position p to the position $\sigma(p)$. In other words,

$$\sigma(j_1, \dots, j_n) = (j_{\sigma^{-1}1}, \dots, j_{\sigma^{-1}n}).$$

Symmetry is detected on the coefficients: f is symmetric if and only if $f_{\sigma J} = f_J$ for all J and all $\sigma \in \mathfrak{S}(A)$; it suffices that f_J is unchanged whenever two neighbor entries in J are interchanged. Similarly, f is anti-symmetric if and only if $f_{\sigma J} = \operatorname{sign}(\sigma) f_J$; it suffices that f_J changes sign whenever two neighbor entries in J are interchanged.

(5.6) Example. Consider the following polynomial of $R[A]$:

$$\Delta(a_1, \dots, a_n) := \prod_{a<b} (b - a) = \prod_{p<q} (a_q - a_p).$$

The factor $a_q - a_p$ is homogeneous of degree 1. It may be viewed as a monic polynomial in the single variable a_q over the polynomial ring generated by the variables different from a_q, and as such it is regular in $R[A]$. Consequently, Δ is regular, and homogeneous of degree equal to the number, $n(n-1)/2$, of factors. The leading monomial in Δ is $a_1^{n-1} a_2^{n-2} \cdots a_n^0$ and the leading coefficient is $(-1)^{n(n-1)/2}$.

Let σ be a permutation. Then $\sigma\Delta$ is the product of the factors $\sigma b - \sigma a$ for $a < b$. Clearly, if (a, b) is not an inversion for σ, then $\sigma b - \sigma a$ is one of the factors of Δ and if (a, b) is an inversion for σ, then $\sigma b - \sigma a = -(\sigma a - \sigma b)$ is equal to -1 times a factor of Δ. It follows that

$$\sigma\Delta = (-1)^{\ell(\sigma)} \Delta.$$

Hence Δ is an anti-symmetric polynomial. As a consequence, the square Δ^2 is a symmetric polynomial. The square Δ^2 is called the *discriminant*.

(5.7) Definition. For each multi index J, define the *monomial symmetric polynomial* $m^J = m^J(A)$ as the following sum of monomials,

$$m^J = \sum{}' \sigma(a^J),$$

where the sum is over all *different* monomials of the form $\sigma(a^J)$ for $\sigma \in \mathfrak{S}(A)$. In other words, m^J is the sum of all monomials a^I where I can be obtained from J by a permutation of the entries. The polynomial m^J is obviously a symmetric polynomial, and homogeneous of degree $\|J\|$.

Clearly, if the multi index I is a permutation of J, then $m^I = m^J$. Among the multi indices that are permutations of J, the largest, K say, is characterized as being decreasing, that is, by the property that $k_1 \geq k_2 \geq \cdots \geq k_n$. Hence the monomial symmetric polynomials are naturally parametrized by the decreasing multi indices K. Note that if K is a decreasing multi index, then the leading term in m^K is the monomial a^K.

In the notation m^K for a decreasing multi index K, the trailing zeros in K are often omitted. We may also omit the commas separating the entries of K, when the separation is clear from the context. (However, for the smallest multi index $K = (0, \dots, 0)$ we write $m^0 = m^{0\dots0} = 1$.) For instance, for the alphabet with $n = 3$ letters a, b, c, we have that

$$m^1 = a + b + c,$$
$$m^2 = a^2 + b^2 + c^2, \quad m^{11} = ab + ac + bc,$$
$$m^3 = a^3 + b^3 + c^3, \quad m^{21} = a^2b + a^2c + ab^2 + ac^2 + b^2c + bc^2, \quad m^{111} = abc.$$

(5.8) Lemma. *The monomial symmetric polynomials m^K, for all decreasing multi indices K, form an R-basis for* $\mathrm{Sym}_R[A]$.

Proof. The terms of the polynomial m^K are the monomials a^J where J is a permutation of K. Hence a polynomial f is an R-linear combination of the m^K if and only if, for all decreasing multi indices K we have that $f_J = f_K$ when J is a permutation of K, that is, if and only if f is a symmetric polynomial. □

(5.9) Definition. The d'th *elementary symmetric polynomial* $e_d = e_d(A)$ is the sum of all products of d different letters. Thus $e_0 = 1$ (the empty product is equal to 1) and $e_d = 0$ when $d > n$. Clearly, $e_1 = a_1 + \cdots + a_n$, and $e_n = a_1 \cdots a_n$. In general,

$$e_d = \sum_{1 \le i_1 < \cdots < i_d \le n} a_{i_1} \cdots a_{i_d}.$$

Obviously, e_d is a symmetric polynomial, and homogeneous of degree d. Note that e_d for $d \le n$ is the special monomial symmetric polynomial,

$$e_d = m^{1\dots10\dots0} = m^{1\dots1},$$

with d occurrences of 1.

Equivalently, the e_d may be defined by the following expansion in the polynomial ring $R[A][T]$ in one variable T over $R[A]$:

$$\prod_{a \in A} (T - a) = T^n - e_1 T^{n-1} + \cdots + (-1)^n e_n.$$

(5.10) Theorem. *The products $e^I := e_1^{i_1} \cdots e_n^{i_n}$, for all multi indices I, form an R-basis for* $\mathrm{Sym}_R[A]$.

Proof. The leading term in e_d is the monomial $a^{1\dots10\dots0}$. Hence the leading term of e^I is the monomial a^K, where

$$K = (i_1 + \cdots + i_n, i_2 + \cdots + i_n, \dots, i_{n-1} + i_n, i_n). \tag{5.10.1}$$

To prove the Theorem, we have to prove that any symmetric polynomial f has an expansion $f = \sum_I r_I e^I$ with uniquely determined coefficients $r_I \in R$. To prove the existence, assume that $f \neq 0$, and consider the leading monomial, a^K say, of f. Since f is symmetric, it follows, for instance from (5.8), that K is a decreasing multi index. Hence there is a unique multi index I such that (5.10.1) holds. Then the polynomial f and the polynomial $f_K e^I$ have the same leading term, namely $f_K a^K$. Consequently, the difference $f - f_K e^I$ is either the zero polynomial or its leading monomial is strictly less than the leading monomial of f. If the difference $f - f_K e^I$ is non-zero, repeat the argument. By induction, in a finite number of steps, we obtain the required expansion of f.

By almost the same argument, unicity holds. Indeed, it follows for a sum $\sum_I r_I e^I$ that the non-zero terms have different leading monomials, and so a non-trivial sum is never the zero polynomial. □

(5.11) Remark. An important application of the theorem is the following: Consider a monic polynomial $P \in R[T]$, say,

$$P = T^n + r_1 T^{n-1} + \cdots + r_{n-1} T + r_n. \tag{5.11.1}$$

Let S be a commutative ring containing R as a subring, and assume that P as a polynomial in $S[T]$ has an expansion as a product,

$$P = (T - \alpha_1) \cdots (T - \alpha_n). \tag{5.11.2}$$

Evaluation at $\alpha = (\alpha_1, \dots, \alpha_n)$ is the homomorphism of R-algebras $R[A] \to S$,

$$f \mapsto f(\alpha_1, \dots, \alpha_n).$$

By expanding the product (5.11.2) it follows that

$$e_d(\alpha_1, \dots, \alpha_n) = (-1)^d r_d.$$

So, up to a sign, evaluation of the elementary symmetric polynomials yield the coefficients of P. It follows from the Theorem that if f is a symmetric polynomial in $R[A]$, then the value $f(\alpha_1, \dots, \alpha_n)$ is an R-linear combination of products $r_1^{i_1} \cdots r_n^{i_n}$. In other words, without knowing the "roots" α_i of P it is possible to express, for a symmetric polynomial $f \in R[A]$, the value $f(\alpha_1, \dots, \alpha_n)$ as a polynomial in the coefficients of P.

(5.12) Example. The polynomial e^I is obviously a sum of monomials a^J. Hence, in the notation of the proof of theorem (5.10), there is an equation,

$$e^I = m^K + \sum_{L<K} \alpha_{I,L} m^L,$$

where the sum is over decreasing multi indices L less that K and the coefficients $\alpha_{I,L}$ are non-negative integers. The equation expresses the basis e^I in terms of the basis m^K, and it follows that the base change matrix is an upper triangular matrix with 1 in the diagonal. So it is easy to invert the matrix and express the m^K in terms of the basis e^I.

In degree 0, we have that $e^{0\ldots0} = 1 = m^0$, and in degrees 1, 2, and 3,

$$e_1 = m^1,$$
$$e_2 = m^{11}, \quad e_1^2 = m^2 + 2m^{11},$$
$$e_3 = m^{111}, \quad e_1e_2 = m^{21} + 3m^{111}, \quad e_1^3 = m^3 + 3m^{21} + 6m^{111}$$

By solving the equations, it follows that

$$m^1 = e_1,$$
$$m^{11} = e_2, \quad m^2 = e_1^2 - 2e_2,$$
$$m^{111} = e_3, \quad m^{21} = e_1e_2 - 3e_3, \quad m^3 = e_1^3 - 3e_1e_2 + 3e_3.$$

(5.13) Definition. The d'th *power sum* $p_d = p_d(A)$, for $d \geq 1$, is the sum of the d'th powers of the variables, that is, $p_d = \sum_a a^d$. Equivalently,

$$p_d = m^d.$$

The *complete symmetric* polynomial $s_d = s_d(A)$ is the sum of all monomials of degree d, that is, $s_d = \sum_{\|J\|=d} a^J$. Equivalently,

$$s_d = \sum_{\|K\|=d}' m^K,$$

where the sum is over decreasing multi indices.

(5.14). The polynomials e_d, s_d, and p_d appear naturally as coefficients of power series. Indeed, in the power series ring $R[A][[T]]$ in one variable, we have the equations,

$$e(T) = e(A)(T) := \prod_{a \in A}(1 + aT) = \sum_{d=0}^{\infty} e_d T^d, \tag{5.14.1}$$

$$s(T) = s(A)(T) := \prod_{a \in A} \frac{1}{1 - aT} = \sum_{d=0}^{\infty} s_d T^d, \tag{5.14.2}$$

$$p(T) = p(A)(T) := \sum_{a \in A} \frac{a}{1 - aT} = \sum_{d=0}^{\infty} p_{d+1} T^d. \tag{5.14.3}$$

It follows from the equations that the power series $s(T)$ is the inverse of the power series $e(-T)$, that is, we have the equation $e(-T)s(T) = 1$. Hence, since $e_0 = 1$, the coefficients

s_d of $s(T)$ are determined recursively from the coefficients $(-1)^d e_d$ of $e(-T)$. For instance, in low degrees we obtain the formulas,

$$\begin{aligned} e_0 s_0 = 1;&\quad s_0 = 1,\\ s_1 - e_1 = 0;&\quad s_1 = e_1,\\ s_2 - e_1 s_1 + e_2 = 0;&\quad s_2 = e_1^2 - e_2,\\ s_3 - e_1 s_2 + e_2 s_1 - e_3 = 0;&\quad s_3 = e_1^3 - 2e_1e_2 + e_3,\\ s_4 - e_1 s_3 + e_2 s_2 - e_3 s_1 + e_4 = 0;&\quad s_4 = e_1^4 - 3e_1^2 e_2 + 2e_1e_3 + e_2^2 - e_4. \end{aligned}$$

Similarly, it follows from the equations that the power series $p(T)$ is equal to the logarithmic derivative of $s(T)$ and equal to -1 times the logarithmic derivative of $e(-T)$, that is, we have the equations $p(T) = s(T)'/s(T) = -(e(-T))'/e(-T)$ or, equivalently,

$$s(T)p(T) = s'(T), \qquad e(-T)p(T) = e'(-T). \tag{5.14.4}$$

Hence, the coefficients p_{d+1} of $p(T)$ are determined recursively from the coefficients e_d of $e(T)$. For instance, in low degrees we obtain the formulas,

$$\begin{aligned} e_0 p_1 = e_1;&\quad p_1 = e_1,\\ p_2 - e_1 p_1 = -2e_2;&\quad p_2 = e_1^2 - 2e_2,\\ p_3 - e_1 p_2 + e_2 p_1 = 3e_3;&\quad p_3 = e_1^3 - 3e_1e_2 + 3e_3,\\ p_4 - e_1 p_3 + e_2 p_2 - e_3 p_1 = -4e_4;&\quad p_4 = e_1^4 - 4e_1^2 e_2 + 4e_1e_3 + 2e_2^2 - 4e_4. \end{aligned}$$

6. Alternating polynomials.

(6.1) Lemma. *The following three conditions on a polynomial $f = \sum_J f_J a^J$ in $R[A]$ are equivalent:*

(i) *The coefficients f_J are alternating in the multi index J, that is, f_J changes sign when two entries at different positions in J are interchanged and f_J vanishes when two entries of different position in J are equal.*

(ii) *The polynomial f is anti-symmetric and divisible by the product $\prod_{p<q}(a_q - a_p)$.*

(iii) *The polynomial f is anti-symmetric and, for all $q > p$, the substitution $a_q := a_p$ in f yields the zero polynomial.*

Proof. As noted in (5.5), the polynomial f is anti-symmetric if and only if f_J changes sign when two entries of J are interchanged. Therefore, to prove the equivalence of the three conditions, we may assume that f is anti-symmetric.

Let $p < q$ be arbitrary integers between 1 and n, and let τ be the transposition in $\mathfrak{S}(A)$ that interchanges a_p and a_q. Then, to prove the equivalence of (i) and (iii), it suffices to prove the following assertion: the substitution $a_q := a_p$ in f yields zero if and only if $f_J = 0$ for all multi indices J such that $j_p = j_q$. To prove the latter assertion, decompose f into two sums of monomials,

$$f = \sum_{j_p = j_q} f_J a^J + \sum_{j_q < j_p} \bigl(f_J a^J + f_{\tau J} a^{\tau J}\bigr). \tag{6.1.1}$$

Clearly, the substitution $a_q := a_p$ in a^J and in $a^{\tau J}$ yield the same result. Moreover, since f is anti-symmetric, we have that $f_{\tau J} = -f_J$. Consequently, the substitution $a_q := a_p$ in the second sum of (6.1.1) yields zero. Therefore, the substitution $a_q := a_p$ in f yields zero, if and only if the substitution $a_q := a_p$ in the first sum yields zero. Obviously, the substitution $a_q := a_p$ in the first sum yields zero if and only if all the coefficients f_J in the first sum are equal to zero. Hence the equivalence of (i) and (iii) holds.

To prove the equivalence of (ii) and (iii), note that the substitution $a_q := a_p$ in f yields zero if and only if f is divisible by the difference $a_q - a_p$. Hence the equivalence is a consequence of the following general observation (applied to the differences $d_i = a_q - a_p$ for $p < q$): Assume for finitely many elements d_i in a commutative ring S that the class $d_i \pmod{d_j}$ for all $i \neq j$ is regular in the residue ring $S/d_j S$. Then an element $f \in S$ is divisible by every d_i iff f is divisible by the product of all the d_i. ∎

(6.2) Definition. A polynomial f is said to be *alternating* if it satisfies the equivalent conditions in Lemma (6.1). Clearly, the alternating polynomials form an R-submodule $\mathrm{Alt}_R[A]$ of $R[A]$. In fact, it follows from any of the characterizations (ii) or (iii) that $\mathrm{Alt}_R[A]$ is a $\mathrm{Sym}_R[A]$-submodule, that is, a product of a symmetric polynomial and an alternating polynomial is alternating.

If 2 is a regular element in R then the alternating polynomials are simply the anti-symmetric polynomials. Indeed, if f is anti-symmetric and J is a multi-index with two equal entries, then $f_J = -f_J$ and hence $2f_J = 0$.

(6.3) Definition. Consider the $\infty \times n$ matrix,

$$V = V(a_1, \dots, a_n) := \begin{pmatrix} 1 & \dots & 1 \\ a_1 & \dots & a_n \\ \vdots & & \vdots \\ a_1^i & \dots & a_n^i \\ \vdots & & \vdots \end{pmatrix}, \tag{6.3.1}$$

whose rows are naturally indexed $0, 1, 2, \dots$. For any multi-index $J = (j_1, \dots, j_n)$, consider the $n \times n$ matrix $V^J = V^J(a_1, \dots, a_n)$ obtained by selecting from the matrix V the n rows with indices $j_1, \dots, j_n$, and denote by $\Delta^J = \det V^J$ its determinant, that is,

$$\Delta^J = \Delta^{j_1,\dots,j_n}(a_1, \dots, a_n) := \begin{vmatrix} a_1^{j_1} & \dots & a_n^{j_1} \\ \vdots & & \vdots \\ a_1^{j_n} & \dots & a_n^{j_n} \end{vmatrix}. \tag{6.3.2}$$

The special determinant obtained when J is the sequence $0, 1, \dots, n-1$ will be called the *Vandermonde determinant* and denoted $\Delta(a_1, \dots, a_n)$, that is,

$$\Delta(a_1, \dots, a_n) := \begin{vmatrix} 1 & \dots & 1 \\ a_1 & \dots & a_n \\ \vdots & & \vdots \\ a_1^{n-1} & \dots & a_n^{n-1} \end{vmatrix}. \tag{6.3.3}$$

The determinants Δ^J are polynomials in $R[A]$. Clearly, they are homogeneous of degree $\|J\|$. It follows from usual properties of determinants, as functions of the columns, that the condition (6.1)(iii) holds for Δ^J. Hence Δ^J is an alternating polynomial. Moreover, the determinants Δ^J are alternating in the entries of J, that is, if σ is a permutation in $\mathfrak{S}(A)$ then $\Delta^{\sigma J} = (\operatorname{sign} \sigma)\Delta^J$ and $\Delta^J = 0$ if the multi index J has two equal entries.

Note that, since Δ^J is an alternating polynomial, we have that Δ^J is divisible by the product $\prod_{p<q}(a_q - a_p)$. We prove in Corollary (6.7) that the notation Δ used for the latter product in (SYM.5.6) is not in conflict with the notation used here.

(6.4) Example. For $n = 3$ and the alphabet with the letters a, b, c, we have that

$$\Delta^{ijk}(a, b, c) = \begin{vmatrix} a^i & b^i & c^i \\ a^j & b^j & c^j \\ a^k & b^k & c^k \end{vmatrix} = a^i b^j c^k + a^k b^i c^j + a^j b^k c^i - a^i b^k c^j - a^j b^i c^k - a^k b^j c^i.$$

In particular,

$$\Delta(a, b, c) = -a^2b + a^2c + ab^2 - ac^2 - b^2c + bc^2 = (b-a)(c-a)(c-b).$$

(6.5) Note. The Vandermonde determinant defined in (6.3) is the determinant used by Jacobi [1]. Up to a sign, the determinant is independent of the given ordering of the letters of A. Our choice of sign differs from that used by Macdonald [2] and others. Indeed, the Vandermonde determinant defined in [2] differs by the sign $(-1)^{n(n-1)/2}$ from ours.

(6.6) Proposition. *The determinant Δ^J is given by the following formula,*

$$\Delta^J = \sum_{\sigma \in \mathfrak{S}(A)} \operatorname{sign}(\sigma)\sigma(a^J). \tag{6.6.1}$$

Moreover, the determinants Δ^J, for all strictly increasing multi indices J, form an R-basis for the module $\mathrm{Alt}_R[A]$ *of alternating polynomials in $R[A]$.*

Proof. The formula (6.6.1) is just the usual expression for the determinant (6.3.2).

Let f be an alternating polynomial. By (6.1)(i), the only monomials appearing in f are of the form a^I where all entries in the multi index I are different. If all the entries in a multi index I are different, then they may be arranged into strictly increasing order by a unique permutation, that is, we have that $I = \sigma J$ where σ is a permutation and J is a strictly increasing multi index, both uniquely determined by I. Moreover, since f is anti-symmetric, we have that $f_I = (\operatorname{sign}\sigma) f_J$, that is, the term $f_I a^I$ is equal to $f_J(\operatorname{sign}\sigma)a^{\sigma J}$. It follows that polynomials on the right side of (6.6.1), for all strictly increasing multi indices J, form an R-basis for $\mathrm{Alt}_R[A]$. ∎

(6.7) Corollary. *For the Vandermonde determinant we have the equation,*

$$\Delta = \Delta(a_1, \dots, a_n) = \prod_{p<q} (a_q - a_p). \tag{6.7.1}$$

Moreover, multiplication by Δ is an isomorphism from the R-submodule $\mathrm{Sym}_R[A]$ *of symmetric polynomials onto the R-submodule of* $\mathrm{Alt}_R[A]$ *of alternating polynomials. Finally, the symmetric polynomials,*

$$s^J(A) := \Delta^J/\Delta,$$

for all strictly increasing multi indices J, form an R-basis for the module $\mathrm{Sym}_R[A]$ *of symmetric polynomials.*

Proof. In (6.7.1) the Vandermonde determinant is alternating and hence divisible by the right side. Moreover, the two polynomials have the same degree, namely $n(n-1)/2$, and the right side is homogeneous and regular. Hence the left side is a constant times the right side. It is easy to see that the coefficient to $a_1^0 a_2^1 \cdots a_n^{n-1}$ in both polynomials is equal to 1. Therefore, the two polynomials are equal.

By Lemma (6.1), every alternating polynomial is divisible by Δ. Clearly, if a polynomial f is divisible by Δ, say $f = g\Delta$ where g is unique since Δ is regular, then f is alternating if and only if g is symmetric. Hence the second assertion of the Corollary holds. The final assertion follows from the second and the description of the R-basis for $\mathrm{Alt}_R[A]$ in Proposition (6.6). ∎

(6.8) Definition. The symmetric polynomials $s^J = \Delta^J/\Delta$ of (6.7) are called *Schur polynomials*. Since Δ^J is homogeneous of degree $\|J\|$ and Δ is homogeneous of degree $n(n-1)/2$, it follows that s^J is homogeneous of degree $\|J\| - n(n-1)/2$.

The polynomials s^J for strictly increasing multi indices J will be called *proper Schur polynomials*. They form an R-basis for $\mathrm{Sym}_R[A]$. For the smallest strictly increasing multi index $(0, 1, \dots, n-1)$, we have that

$$s^{0,1,\dots,n-1} = 1.$$

The Schur polynomials s^J are alternating in the multi index J, that is, $s^J = 0$ if J has two equal entries and s^J changes sign when two entries at different position in J are interchanged. In particular, any Schur polynomial s^J is either equal to 0 or, up to sign, equal to a proper Schur polynomial.

In general, it is hard to compute the Schur polynomials directly as the determinants divided by Δ, and later we will prove other relations involving the Schur polynomials. As an example, let us prove here the following formulas, for $p = 0, \dots, n$:

$$s^{0,1,\dots,\hat{p},\dots,n} = e_{n-p}\,. \tag{6.8.1}$$

Consider the polynomial $D(T)$ in $R[A][T]$ defined as the Vandermonde determinant

$$D(T) := \Delta(a_1, \dots, a_n, T) = \begin{vmatrix} 1 & \dots & 1 & 1 \\ a_1 & \dots & a_n & T \\ \vdots & & \vdots & \vdots \\ a_1^n & \dots & a_n^n & T^n \end{vmatrix}.$$

By (6.7.1), applied to the alphabet $\{a_1, \dots, a_n, T\}$, we have the product expansion $D(T) = \Delta\prod_i (T - a_i)$. On the other hand, by developing the determinant $D(T)$ along its last column, we obtain the equation $D(T) = \sum_p (-1)^{n-p}\Delta^{0,\dots,\hat{p},\dots,n}T^p$. Hence we have the equation,

$$\Delta\prod_{i=1}^{n}(T - a_i) = \sum_{p=0}^{n}(-1)^{n-p}\Delta^{0,\dots,\hat{p},\dots,n}T^p.$$

As noted in (5.9), the Formula (6.8.1) is a consequence.

(6.9) Definition. If f is any polynomial in $R[A]$, then the sum,

$$\sum_{\sigma\in\mathfrak{S}(A)} (\operatorname{sign}\sigma)\sigma f,$$

is an alternating polynomial. Indeed, the sum is R-linear as a function of f and if f is a monomial a^J, then, by (6.6.1), the sum is equal to the determinant Δ^J which is an alternating polynomial. Therefore, it follows from Corollary (6.7) that the equation,

$$\delta(f) := \frac{1}{\Delta}\sum_{\sigma\in\mathfrak{S}(A)} (\operatorname{sign}\sigma)\sigma f\,, \tag{6.9.1}$$

defines a map $\delta = \delta^A \colon R[A] \to \mathrm{Sym}_R[A]$. Obviously, the map δ is an R-linear operator. It is called the *symmetrization operator*. As noted above, the Schur polynomial s^J is the result of symmetrizing the monomial a^J,

$$s^J = \delta(a^J). \tag{6.9.2}$$

It is obvious from (6.9.1) that symmetrization is a $\mathrm{Sym}_R[A]$-linear operator, that is, if f and g are polynomials and g is symmetric, then

$$\delta(gf) = g\delta(f). \tag{6.9.2}$$

As a consequence, we obtain for a symmetric polynomial g and any multi index L the formula,

$$gs^L = \sum_I g_I s^{I+L}. \tag{6.9.3}$$

Indeed, the two sides of (6.9.3) are the results of symmetrizing $ga^L = \sum_I g_I a^{I+L}$.

(6.10) Note. The formula (6.9.3) expresses the product gs^L of a symmetric polynomial g and a Schur polynomial s^L as an R-linear combination of Schur polynomials. To get the expansion of gs^L in the basis s^J consisting of proper Schur polynomials we have to consider the non-zero terms in (6.9.3), that is, the terms for which the multi index $I + L$ has all entries different, and then, for the non-zero terms, we have to collect the coefficients for which $I + L$ is a permutation of a given strictly increasing multi index J. This collection of terms is often of combinatorial nature.

For instance, let m^K be the monomial symmetric polynomial. Then

$$m^K s^L = {\sum}'_I \, s^{I+L}, \tag{6.10.1}$$

where the sum is over all *different* permutations I of the entries in K. Indeed, the formula follows from (6.9.3) since $m^K = \sum' a^I$.

In particular, since the d'th elementary symmetric polynomial e_d, for $0 \le d \le n$, is the monomial symmetric polynomial $m^{1\dots10\dots0}$ (with 1 occurring d times), it follows that $e_d s^L$ is the sum (6.10.1) over all the $\binom{n}{d}$ permutations I of $(1, \dots, 1, 0, \dots, 0)$. Take $L := (0, 1, \dots, n-1)$. Then $s^L = 1$ and the formula is the expansion of e_d in terms of Schur polynomials. Clearly, $I + L$ has two equal entries unless $I = (0, \dots, 0, 1, \dots, 1)$. So the formula reduces to the formula of (6.8.1),

$$e_d = s^{0,\dots,n-d-1,n-d+1,\dots,n}. \tag{6.10.2}$$

Similarly, the d'th power sum p_d, for $d \ge 1$, is the monomial symmetric polynomial $m^{d0\dots0}$. Hence $p_d s^L$ is the sum (6.10.1) over the n permutations I of $(d, 0, \dots, 0)$. Take $L := (0, 1, \dots, n-1)$ to obtain the following expansion of p_d (over $0 \le i < n$):

$$p_d = \sum s^{0,1,\dots,i-1,d+i,i+1,\dots,n-1} = \sum_{i \ge n-d} (-1)^{n-i-1} s^{0,\dots,\hat{\imath},\dots,n-1,d+i}. \tag{6.10.3}$$

(6.11) Note. Take as L in (6.10.1) the strictly increasing multi index $E := (0, 1, \dots, n-1)$. Then the left side is the monomial symmetric polynomial m^K, and so, as noted above, from (6.10.1) we may obtain the coefficients $\gamma_{J,K}$ in the expansion of m^K in the basis of proper Schur polynomials s^J. Let us, for the moment, index the monomial symmetric polynomials by increasing multi indices, and assume that K is increasing. Consider a term s^{I+E} in (6.10.1) where I is a permutation of K, and let J be an arbitrary permutation of $I + E$. Then $J \geq K + E$, and equality implies that $I = K$. Indeed, with permutations τ, σ we have $I = \tau K$ and $J = \sigma(\tau K + E) = \sigma\tau K + \sigma E$. Since K and E are increasing, it follows that $\sigma\tau K \geq K$ and $\sigma E \geq E$, and hence $J \geq K + E$. Moreover, if $J = K + E$, then $\sigma\tau K = K$ and $\sigma E = E$. As E is strictly increasing, it follows first that $\sigma = 1_n$ and next that $\tau K = K$.

Consequently, the coefficients $\gamma_{J,K}$ are integers, they vanish when $J < K + E$, and $\gamma_{K+E,K} = 1$. Hence the expansion has the form,

$$m^K = s^{K+E} + \sum_{J > K+E} \gamma_{J,K} s^J, \tag{6.11.1}$$

over strictly increasing multi indices J with $\|J\| = \|K\| + \|E\|$ $(= \|K\| + n(n-1)/2)$.

The space of symmetric polynomials of degree d has two bases: The "monomial basis", formed by the monomial symmetric polynomials m^K, is indexed by increasing multi indices K with $\|K\| = d$, and the "Schur basis", formed by the proper Schur polynomials s^J, is indexed by strictly increasing multi indices J with $\|J\| = d + n(n-1)/2$. The map $K \mapsto K + E$ is a bijective order preserving map from the first index set to the second. According to this bijection, Equation (6.11.1) describes the base change from the second basis to the first. By (6.11.1) the base change matrix $\gamma_{J,K}$ is a lower triangular matrix with 1 in the diagonal, convenient for obtaining expansions of the Schur polynomials s^J in terms of the basis m^K. For $n = 4$, we have in degrees 1, 2, and 3 the equations,

$$m^1 = s^{0124}$$
$$m^{11} = s^{0134}, \quad m^2 = s^{0125} - s^{0134}$$
$$m^{111} = s^{0234}, \quad m^{21} = s^{0135} - 2s^{0234}, \quad m^3 = s^{0126} - s^{0135} + s^{0234}$$

(in the notation for m^K we take again decreasing K), and in degree 4,

$$m^{1111} = s^{1234}, \quad m^{211} = s^{0235} - 3s^{1234}, \quad m^{22} = s^{0145} - s^{0235} + s^{1234},$$
$$m^{31} = s^{0136} - s^{0145} - s^{0235} + 2s^{1234}, \quad m^4 = s^{0127} - s^{0136} + s^{0235} - s^{1234}.$$

By solving the equations, it follows in degrees 1, 2, and 3 that

$$s^{0124} = m^1$$
$$s^{0134} = m^{11}, \quad s^{0125} = m^2 + m^{11}$$
$$s^{0234} = m^{111}, \quad s^{0135} = m^{21} + 2m^{111}, \quad s^{0126} = m^3 + m^{21} + m^{111},$$

and in degree 4,

$$s^{1234} = m^{1111}, \quad s^{0235} = m^{211} + 3m^{1111}, \quad s^{0145} = m^{22} + m^{211} + 2m^{1111},$$
$$s^{0136} = m^{31} + m^{22} + 2m^{211} + 3m^{1111}, \quad s^{0127} = m^{4} + m^{31} + m^{22} + m^{211} + m^{1111}.$$

In fact, the coefficients of the proper Schur polynomials in terms of the basis m^K are always non-negative, and we will later give a combinatorial expression for the coefficients.

(6.12) Pieri's Formula. *Let s_d be the d'th complete symmetric polynomial. Then, for every strictly increasing multi index L, we have the expansion,*

$$s_d s^L = \sum{}' s^J, \tag{6.12.1}$$

where the sum is over all strictly increasing multi indices $J = (j_1, \dots, j_n)$ satisfying the inequalities,

$$l_1 \le j_1 < l_2 \le j_2 < \cdots < l_{n-1} \le j_{n-1} < l_n \le j_n,$$

and the equality $\|J\| = \|L\| + d$. In particular, we have the equation,

$$s_d = s^{0,\dots,n-2,n-1+d}. \tag{6.12.2}$$

Proof. Recall that $J \succeq L$ if $j_p \ge l_p$ for $p = 1, \dots, n$, or, equivalently, if $J = I + L$ with a multi index I of size n. Denote by $\mathcal{J}$ the set of multi indices J such that $J \succeq L$ and $\|J\| = \|L\| + d$. Since $s_d = \sum_{\|I\|=d} a^I$, it follows from (6.9.3) that $s_d s^L = \sum_{\|I\|=d} s^{I+L}$, or, equivalently,

$$s_d s^L = \sum_{J \in \mathcal{J}} s^J. \tag{6.12.3}$$

In (6.12.1), the sum is over all $J \in \mathcal{J}$ such that the following inequalities hold for $p = 1, \dots, n-1$:

$$j_p < l_{p+1}. \tag{*}$$

Therefore, to prove the Formula (6.12.1), we have to prove that the sum of the s^J, over those $J \in \mathcal{J}$ for which one of the inequalities (*) is false, is equal to zero.

For each $q = 1, \dots, n-1$, let $\mathcal{J}_q$ be the subset of $\mathcal{J}$ consisting of multi indices J such that the inequality (*) holds for all $p < q$ but not for $p = q$. It suffices to prove that the sum of the s^J for $J \in \mathcal{J}_q$ is equal to zero.

To prove the latter assertion, let τ be the simple transposition that interchanges in a multi index J the q'th and the $(q+1)$'st entry. If J belongs to $\mathcal{J}_q$, then $j_q \ge l_{q+1}$ by definition of $\mathcal{J}_q$, and $j_{q+1} \ge l_{q+1}$ since $J \succeq L$. As L is increasing, the two inequalities imply first that $\tau J \succeq L$, and next that $\tau J \in \mathcal{J}_q$. Hence τ defines an involution of the set $\mathcal{J}_q$. If $\tau J = J$, then J has two equal entries, and then $s^J = 0$. If $\tau J \ne J$, then $s^J + s^{\tau J} = 0$, since s^J is alternating in J. It follows that the sum $\sum_{J \in \mathcal{J}_q} s^J$ is equal to zero.

Thus Formula (6.12.1) holds. Clearly, the Formula (6.12.2) is the special case obtained when $L = (0, 1, \dots, n-1)$. □

(6.13) Notation. Consider the three bases for the R-module $\mathrm{Sym}_R[A]$ of symmetric polynomials: the *monomial basis* of monomial symmetric polynomials m^K, indexed by decreasing multi indices K, the *elementary basis* of products $e^I = e_1^{i_1} \cdots e_n^{i_n}$ of the elementary symmetric polynomials, indexed by arbitrary multi indices I, and the *Schur basis* of proper Schur polynomials s^J, indexed by strictly increasing multi indices J. In all three cases, the multi indices are assumed to be of size equal to the number n of letters of A. It will be convenient to introduce a notation where multi indices of arbitrary sizes are allowed.

First, as noted in (5.7), it is common for a decreasing multi index K to omit (some of) the trailing zeros in the notation m^K. More precisely, we define, for any decreasing multi index $K = (k_1, \dots, k_r)$, the polynomial m^K as follows: If $r < n$, then $m^K := m^{k_1,\dots,k_r,0,\dots,0}$ with $n - r$ trailing zeros. If $r > n$ and $k_{n+1} = \cdots = k_r = 0$, then $m^K := m^{k_1,\dots,k_n}$. Finally, if $r > n$ and some entry k_q with $q > n$ is positive, then $m^K := 0$. In all cases, m^K is homogeneous of degree $\|K\|$. In this notation, the monomial basis consists of the polynomials m^K where K is a decreasing multi index of size at most n and with no trailing zeros.

Next, for an arbitrary multi index $I = (i_1, \dots, i_r)$ of size r, we define $e^I := e_1^{i_1} \dots e_r^{i_r}$. The elementary symmetric polynomials e_d are defined for all d, and they vanish when $d > n$. So, $e^I = 0$ if and only if $r > n$ and some entry i_q for $q > n$ is positive. In all cases, e^I is homogeneous of degree equal to $i_1 + 2i_2 + \cdots + ri_r$. In this notation, the elementary basis consists of the products e^I where I is a multi index of size at most n and with no trailing zeros.

Consider finally Schur polynomials. For a strictly increasing multi index $J = (j_1, \dots, j_r)$ of size r, it will be convenient to say that multi indices of the following form, for some $t \geq 0$, are *extensions* of J:

$$\hat{J} := (0, 1, \dots, t-1, t + j_1, \dots, t + j_r).$$

If $r \leq n$, define $s^J := s^{\hat{J}}$, where $\hat{J}$ is the extension of J to a multi index of size n, obtained with $t := n - r$. Then s^J is a proper Schur polynmial, easily seen to have degree $\|J\| - r(r-1)/2$. Clearly, any strictly increasing multi index is the extension of a unique strictly increasing multi index of smaller size with all entries positive. With this notation, the Schur basis consists of the Schur polynomials s^J where J is a strictly increasing multi index of size at most n and with all entries positive. For $r > n$ we define s^J as follows: If J is an extension of a strictly increasing multi index J_0 of size n, then $s^J := s^{J_0}$; otherwise $s^J := 0$.

The empty sequence ( ) is allowed in all three cases, and, according to the definitions, $m^{()} = e^{()} = s^{()} = 1$.

For instance, in degree 4, the elements of the three bases are the following:

$$\begin{array}{ccccc}
m^{1111}, & m^{211}, & m^{22}, & m^{31}, & m^{4}; \\
e^{0001}, & e^{101}, & e^{02}, & e^{21}, & e^{4}; \\
s^{1234}, & s^{124}, & s^{23}, & s^{14}, & s^{4};
\end{array}$$

except that, when the number n of variables is less than 4, the polynomials with a multi index of size larger than n have to be discarded from the list.

Note that, in the extended notation for the proper Schur polynomials, the formulas of (6.10.2), (6.12.2), and (6.10.3), simplify to the following:

$$e_d = s^{1,\dots,d}, \quad s_d = s^d, \quad p_d = \sum_{0\le j<\max(n,d)} (-1)^j s^{1,\dots,j,d}.$$

(6.14) Notation. There is another natural way to index the products $e^I = e_1^{i_1}\cdots e_n^{i_n}$ of the elementary basis. If $L = (l_1,\dots,l_r)$ is multi index of any size r, let

$$e_L := e_{l_1}\cdots e_{l_r}.$$

Note that e_L is homogeneous of degree $\|L\|$. The products e_L are symmetric in the entries of L, and are often indexed by decreasing multi indices. If $l_1 > n$, then $e_L = 0$.

Clearly, for every multi index I of size n we have $e^I = e_L$, where

$$L := (\overbrace{n,\dots,n}^{i_n},\ \dots, \overbrace{2,\dots,2}^{i_2}, \overbrace{1,\dots,1}^{i_1}).$$

Hence the elementary basis consists of the products e_L where L is a decreasing multi index of integers in $\{1,\dots,n\}$, and of arbitrary size (including the empty multi index).

For instance, in degree 4 we have that

$$e^{0001} = e_4,\ e^{101} = e_1e_3 = e_{31},\ e^{02} = e_2^2 = e_{22},\ e^{21} = e_1^2e_2 = e_{211},\ e^4 = e_1^4 = e_{1111}.$$

(6.15) Remark. It is easy to express, for each decreasing multi index L of size r, the coefficients $\alpha_{L,K}$ in the expansion $e_L = \sum_K \alpha_{L,K} m^K$ of e_L in the monomial basis. Namely, *$\alpha_{L,K}$ is the number of $r\times n$ matrices with entries 0 or 1 and row sums $l_1,\dots,l_r$ and column sums $k_1,\dots,k_n$.* Indeed, the terms of e_l are the monomials $a_1^{p_1}\dots a_n^{p_n}$ where the p_j are 0 or 1 and $\sum_j p_j = l$. Hence each matrix of the said form corresponds to the selection of a term in each e_{l_i} for $i = 1,\dots,r$ such that the product of the selected terms is equal to a^K.

Similarly, if we define $s_L := s_{l_1}\cdots s_{l_r}$, then, in the expansion $s_L = \sum_K \beta_{L,K} m^K$, the coefficient *$\beta_{L,K}$ is the number of $r\times n$ matrices with non-negative integer entries and row sums $l_1,\dots,l_r$ and column sums $k_1,\dots,k_n$.*

As a consequence, the matrix of the $\alpha_{L,K}$, indexed by all decreasing multi indices of size n and degree d, is a symmetric matrix. Consider, for $d\le n$, the products e_L for all decreasing multi indices L of size n and degree at most d. They are simply the elementary symmetric polynomials e^I of degree at most d. Hence they form a basis for the R-module of symmetric polynomials of degree at most d, and the matrix $\alpha_{L,K}$ is the base change matrix from the basis of the e_L to the basis of the m^K for $\|K\|\le d$.

7. Determinantal Methods.

(7.1) Setup. Work with matrices and power series over a given commutative ground ring. We will follow a classical convention for matrices M. If I is an ordered set of row indices, we denote by M^I the matrix obtained from M by selecting its rows with indices in I, and if J is an ordered set of column indices, we denote by M_J the matrix obtained from M by selecting its columns with indices in J. In this notation, M^i is the i'th row of M, and M_j is the j'th column. In particular, M^i_j is the ij'th entry in M.

In general, if u is a power series we denote by u_i the coefficient of T^i, that is,

$$u = u_0 + u_1 T + u_2 T^2 + \cdots .$$

If $I = (i_1, \dots, i_r)$ is a multi index of size r, we denote by u_I the product,

$$u_I := u_{i_1} \cdots u_{i_r}.$$

The products u_I are symmetric in the entries of I. In particular, if K is the decreasing permutation of I, then $u_K = u_I$.

Note that the notation is in accordance with the definition of the series $s(T)$ and $e(T)$ in (5.14), but in conflict with the definition of the series $p(T)$ associated with power sums: The d'th coefficient of $p(T)$ is the power sum p_{d+1}.

(7.2) Notation. For a power series u, denote by $\langle u \rangle$ the infinite column of coefficients of u. More generally, for any finite or infinite sequence of r ($0 \le r \le \infty$) power series $u, v, w, \dots$, denote by $\langle u, v, w, \dots \rangle$ the $\infty \times r$ matrix with columns $\langle u \rangle, \langle v \rangle, \langle w \rangle, \dots$.

In this notation, associate with a given power series u the $\infty \times \infty$ matrix,

$$M(u) := \langle u, Tu, T^2u, \dots \rangle = \begin{pmatrix} u_0 & 0 & 0 & 0 & \dots \\ u_1 & u_0 & 0 & 0 & \dots \\ u_2 & u_1 & u_0 & 0 & \dots \\ u_3 & u_2 & u_1 & u_0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

The rows and columns in $M(u)$ are naturally indexed $0, 1, 2, \dots$. In particular, the ij'th entry in $M(u)$ is equal to u_{i-j} where, by convention, $u_k = 0$ if $k < 0$.

Clearly, for power series u, v, and w, the equation $uv = w$ is equivalent to any of the following two matrix equations:

$$M(u)\langle v \rangle = \langle w \rangle, \qquad M(u)M(v) = M(w). \tag{7.2.1}$$

(7.3) Lemma. *Let u be a power series with $u_0 = 1$. Assume for power series v and w that $uv = w$. Then, for $d = 0, 1, \dots$, the following equation holds:*

$$(-1)^d v_d = \begin{vmatrix} w_0 & u_0 & 0 & \dots & 0 \\ w_1 & u_1 & u_0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{d-1} & u_{d-1} & u_{d-2} & \dots & u_0 \\ w_d & u_d & u_{d-1} & \dots & u_1 \end{vmatrix}. \tag{7.3.1}$$

In particular, if $uv = 1$, then

$$(-1)^d v_d = \begin{vmatrix} u_1 & u_0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ u_{d-1} & u_{d-2} & \dots & u_0 \\ u_d & u_{d-1} & \dots & u_1 \end{vmatrix}. \tag{7.3.2}$$

Proof. Let U be the $(d+1) \times (d+1)$ matrix consisting of the first $d+1$ rows and columns of $M(u)$. The columns of U are $U_0, \dots, U_d$. In particular, the column U_0 consists of the first $(d+1)$ coefficients $u_0, \dots, u_d$ of u. Define the columns W_0 and V_0 similarly. The matrix $M(u)$ is a lower triangular matrix. Hence, from the first equation of (7.2.1), or directly, we obtain the matrix equation $UV_0 = W_0$. The latter equation, with the v_i for $i = 0, \dots, d$ as unknowns, is solved by Cramer's formula: Since $\det U = 1$, it follows that v_i is equal to the determinant of the matrix obtained from U by replacing the column U_i by W_0. In particular,

$$v_d = \det(U_0, \dots, U_{d-1}, W_0) = (-1)^d \det(W_0, U_0, \dots, U_{d-1}),$$

which is the asserted Formula (7.3.1). Clearly, Formula (7.3.2) is a special case. ∎

(7.4) Definition. Let $I = (i_1, \dots, i_r)$ and $J = (j_1, \dots, j_r)$ be multi indices of the same size r. Then $M(u)^I$, usually written $M^I(u)$, is an $r \times \infty$ matrix, and, with similar conventions, $M_J(u)$ is an $\infty \times r$ matrix, and $M^I_J(u)$ is an $r \times r$ matrix. Denote by u^I_J the determinant of $M^I_J(u)$, that is,

$$u^I_J := \begin{vmatrix} u_{i_1 - j_1} & \dots & u_{i_1 - j_r} \\ \vdots & & \vdots \\ u_{i_r - j_1} & \dots & u_{i_r - j_r} \end{vmatrix}.$$

Subsets with r elements of the non-negative integers will be identified with strictly increasing multi indices of size r. In particular, the interval $[r] := \{0, 1, \dots, r-1\}$ consisting of the first r non-negative integers will be identified with the multi index $(0, 1, \dots, r-1)$. For any multi index I of size r the matrix $M^I_{[r]}(u)$ is obtained by selecting from $M(u)$ the rows with indices in I and the first r columns, and we will sometimes use u^I as a notation for its determinant,

$$u^I = u^I_{[r]} := \begin{vmatrix} u_{i_1} & u_{i_1 - 1} & \dots & u_{i_1 - r + 1} \\ \vdots & \vdots & & \vdots \\ u_{i_r} & u_{i_r - 1} & \dots & u_{i_r - r + 1} \end{vmatrix}.$$

In this notation, the 1×1 determinant u^i, where the superscript i is an index, is the i'th coefficient u_i of u.

Note that the determinant u^I_J is alternating in I and alternating in J. In particular, the special determinant u^I is alternating in the multi index I. As a consequence, general properties of the determinants u^I_J may be deduced from properties valid for strictly increasing multi indices I and J.

The special determinants u^I should not be confused with the products u_I defined in (7.1).

(7.5) Proposition. *Let u be a power series and let I and J be strictly increasing multi indices of the same size r. View I and J as subsets of the interval $[N] = \{0, 1, \dots, N-1\}$ for some $N \gg 0$, and denote by $x \mapsto x^* = N - 1 - x$ the order reversing involution of $[N]$. Then:*

(1) (Vanishing) *The determinant u^I_J is non-zero only when $I \succeq J$, that is, when $i_p \geq j_p$ for $p = 1, \dots, r$. If u is a polynomial of degree at most n, then u^I_J is non-zero only when $n + j_p \geq i_p \geq j_p$ for $p = 1, \dots, r$.*

(2) (Extension) *The determinant u^I_J is unchanged if the same number is subtracted from all entries in I and in J. If $u_0 = 1$ and $i_p = j_p$ for $p = 1, \dots, t$, then $u^I_J = u^{I_0}_{J_0}$, where $I_0 = (i_{t+1}, \dots, i_r)$ and $J_0 = (j_{t+1}, \dots, j_r)$.*

(3) (Homogeneity) *If λ is an element in the ground ring and w is the power series defined by $w(T) = u(\lambda T)$, then*

$$w^I_J = \lambda^{\|I\| - \|J\|}\, u^I_J. \tag{7.5.1}$$

(4) (Symmetry) *Let I^* and J^* be the images of I and J, as subsets of $[N]$. Then,*

$$u^I_J = u^{J^*}_{I^*}. \tag{7.5.2}$$

(5) (Duality) *Let $\tilde I$ and $\tilde J$ be the complements of I and J, with respect to the interval $[N]$, and denote by I' and J' the images $I' := \tilde I^*$ and $J' := \tilde J^*$, as subsets of $[N]$. If $u_0 = 1$ and v is the power series defined by the equation $u(T)v(-T) = 1$, then*

$$u^I_J = v^{I'}_{J'}. \tag{7.5.3}$$

(6) (Multiplication) *If v and w are power series such that $uv = w$, then*

$$w^I_J = \sum_{I \succeq K \succeq J} u^I_K v^K_J, \tag{7.5.4}$$

where the sum is over strictly increasing multi indices K of size r.

Proof. (1) Consider the matrix $U := M^I_J(u)$. By definition, the pq'th entry in U is the coefficient $U^p_q = u_k$ where $k = i_p - j_q$. Entries of U above U^p_q and to the right of U^p_q are coefficients u_l with $l \leq k$. Assume that $i_p < j_p$ for some p. The diagonal element U^p_p is equal to u_k with $k < 0$. It follows that the largest rectangular block of U with the diagonal element U^p_p as its lower left corner is equal to zero. Hence the determinant $u^I_J = \det U$ vanishes.

Similarly, assume that u is a polynomial of degree at most n, and that $i_p > j_p + n$. Then the diagonal element U_p^p is equal to u_k with $k > n$. It follows that the largest rectangular block of U with the diagonal element U_p^p as its upper right corner is equal to zero. Hence the determinant $u_J^I = \det U$ vanishes.

(2) If the same number is subtracted from all entries of I and J, then the differences $i_p - j_q$ are unchanged. Hence the matrix $U = M_J^I(u)$ and its determinant are unchanged.

If $i_p = j_p$ for $p = 1, \dots, t$, then U is a block matrix,

$$U = \begin{pmatrix} T & 0 \\ * & U_0 \end{pmatrix},$$

where T is a lower triangular $t \times t$ matrix with the element u_0 in the diagonal, and $U_0 := M_{J_0}^{I_0}(u)$. Hence, if $u_0 = 1$, we have that $\det U = \det U_0$, and hence $u_J^I = u_{J_0}^{I_0}$.

(3) Since $w_i = \lambda^i u_i$, it follows that if the pq'th entry in $M_J^I(w)$ is non-zero, then it is equal to $\lambda^{i_p - j_q}$ times the pq'th entry of $M_J^I(u)$. Hence, in the usual expansion of a determinant as a signed sum of products, any non-zero product in the expansion of $\det M_J^I(w)$ is equal to $\lambda^{\|I\| - \|J\|}$ times the corresponding product in the expansion of $\det M_J^I(u)$. Thus Equation (7.5.1) holds.

(4) Denote by U' the matrix $M_{i_1^*, \dots, i_r^*}^{j_1^*, \dots, j_r^*}(u)$. Then the matrix $M_{I^*}^{J^*}(u)$ is obtained from U' by reversing first the order of the rows and next the order of the columns. Hence the two matrices have the same determinant, that is $u_{I^*}^{J^*} = \det U'$. The pq'th entry in U' is $u_{j_p^* - i_q^*}$. As $j_p^* - i_q^* = i_q - j_p$, it follows that U' is the transpose of the matrix $M_J^I(u)$. Hence $\det U' = u_J^I$. Thus Equation (7.5.2) holds.

(5) Let w be the power series defined by $w(T) = v(-T)$ so that, by hypothesis, $uw = 1$. Let U be the matrix consisting of the first N rows and columns of $M(u)$. It is a lower triangular matrix with $u_0 = 1$ in the diagonal. In particular, $\det U = 1$. Define W and V similarly from $M(w)$ and $M(v)$.

First, from the equation $uw = 1$ of power series, we obtain the matrix equation,

$$UW = 1;$$

Hence $W = U^{-1}$. The determinant u_J^I is the minor $\det U_J^I$ of U corresponding to the rows in I and the columns in J. It is the complementary minor to the determinant $\det U_{\tilde{J}}^{\tilde{I}}$. Now, since $\det U = 1$, it follows from Jacobi's Theorem, see Appendix (1.5), that the adjugate of U is the inverse matrix W, and that the complementary minor u_J^I is equal to the minor $\det W_{\tilde{I}}^{\tilde{J}}$ of the inverse matrix W multiplied by the signatures of the permutations $(I\,\tilde{I})$ and $(J\,\tilde{J})$, that is,

$$u_J^I = (-1)^{\ell(I\,\tilde{I})}(-1)^{\ell(J\,\tilde{J})} w_{\tilde{I}}^{\tilde{J}}. \tag{7.5.5}$$

The permutation $(I\,\tilde{I})$ of $(0, 1, \dots, N-1)$ may be brought into strictly increasing order using $i_r - (r-1) + i_{r-1} - (r-2) + \cdots + i_2 - 1 + i_1$ simple transpositions. Hence $\ell(I\tilde{I}) =$

$\|I\| - r(r-1)/2$, and $\ell(I\,\tilde I) + \ell(J\,\tilde J) \equiv \|I\| - \|J\| \pmod 2$. Thus we obtain for the product of the two signs in (7.5.5) the equation,

$$(-1)^{\ell(I\,\tilde I)+\ell(J\,\tilde J)} = (-1)^{\|I\|-\|J\|} = (-1)^{\|\tilde J\|-\|\tilde I\|}.$$

Therefore, by the homogeneity (7.5.1) with $\lambda := -1$, it follows from (7.5.5) that

$$u^I_J = v^{\tilde J}_{\tilde I}. \tag{7.5.6}$$

Finally, I' and J' are the images of $\tilde I$ and $\tilde J$ under the involution $x \mapsto x^*$. So, by the symmetry (7.5.2), the equation (7.5.3) follows from (7.5.6).

(6) Assume that an $r \times r$ matrix W is a product $W = UV$, where U is an $r \times N$ matrix and V is an $N \times r$ matrix. By the Cauchy–Binet Formula, see Appendix (1.2), we have the following equation for the determinant:

$$\det W = \sum_K \det U_K \det V^K, \tag{7.5.7}$$

where the sum is over all subsets K with r elements of the common set $[N]$ of indices for the columns of U and the rows of V.

Since $w = uv$, we have by (7.2.1) the matrix equation $M(w) = M(u)M(v)$. By extracting the equations for the rows in I and the columns in J, it follows that $M^I_J(w) = M^I(u)M_J(v)$. Moreover, in $M^I(u)$ only the first N columns are non-zero since I is a subset of $[N]$. It follows that $M^I_J(w) = M^I_{[N]}(u)M^{[N]}_J(v)$. Apply (7.5.7). The equation (7.5.4) is a consequence, since the product $u^I_K v^K_J$ is only non-zero when $I \succeq K \succeq J$. □

(7.6) Definition. Recall that if $I = (i_1, \dots, i_r)$ is a strictly increasing multi index of size r, then the extensions of I are the multi indices of the form,

$$\hat I := (0, 1, \dots, t-1, t+i_1, \dots, t+i_r).$$

Consider a second strictly increasing multi index J of the same size r, and the extension $\hat J$ of J (with the same t). If u is a power series with $u_0 = 1$, then it follows from (7.5)(2) that

$$u^I_J = u^{\hat I}_{\hat J}. \tag{7.6.1}$$

In particular, since the extension of $[r]$ is $[t+r]$, it follows that $u^I = u^{\hat I}$.

Assume that I is a subset of the interval $[N]$. The complement $\tilde I$ of I is a subset of $[N]$, and corresponds to a strictly increasing multi index $\tilde I = (\tilde\imath_1, \dots, \tilde\imath_t)$ with $t := N - r$. The multi index I' defined in (7.5)(5) is $(i'_1, \dots, i'_t) = (\tilde\imath_t{}^*, \dots, \tilde\imath_1{}^*)$. It is said to be *conjugate* to I. The conjugate of the multi index $[r]$ is equal to $[t]$. Hence, under the conditions of (7.5)(5), we have the equation,

$$u^I = v^{I'}. \tag{7.6.2}$$

Note that the definition of the conjugate depends on the choice of N. However, if N is enlarged, then the new conjugate multi index is an extension of the old. In particular, if u is a power series with $u_0 = 1$, then the determinant $u^{I'}$ is independent of the choice of N.

(7.7) Notation. Let $(\alpha_1, \dots, \alpha_n)$ be a set of n elements of the ground ring. Form, in the notation of (7.2), the $\infty \times n$ matrix,

$$V = V(\alpha_1, \dots, \alpha_n) := \Big\langle \frac{1}{1-\alpha_1 T}, \dots, \frac{1}{1-\alpha_n T} \Big\rangle = \begin{pmatrix} 1 & \dots & 1 \\ \alpha_1 & \dots & \alpha_n \\ \vdots & & \vdots \\ \alpha_1^i & \dots & \alpha_n^i \\ \vdots & & \vdots \end{pmatrix}.$$

The rows of V are naturally indexed $0, 1, 2 \dots$. The matrix V is the evaluation at $(\alpha_1, \dots, \alpha_n)$ of the matrix of (6.3.1). Hence, for any multi index $I = (i_1, \dots, i_n)$ of size n, the determinant $\det V^I(\alpha_1, \dots, \alpha_n)$ is the evaluation of the polynomial Δ^I,

$$\Delta^I(\alpha_1, \dots, \alpha_n) = \det V^I(\alpha_1, \dots, \alpha_n). \tag{7.7.1}$$

In particular, the determinant of the matrix $V^{[n]}$ consisting of the first n rows of V is the Vandermonde determinant $\Delta(\alpha_1, \dots, \alpha_n)$.

In addition, form the two power series,

$$s = s(\alpha_1, \dots, \alpha_n) := \prod (1-\alpha_i T)^{-1}, \quad e = e(\alpha_1, \dots, \alpha_n) := \prod (1+\alpha_i T)$$

(of which e is a polynomial of degree at most n). Finally, for multi indices I and J of the same size r, form the determinants,

$$s^I_J = s^I_J(\alpha_1, \dots, \alpha_n) := \det M^I_J\big(s(\alpha_1, \dots, \alpha_n)\big),$$
$$e^I_J = e^I_J(\alpha_1, \dots, \alpha_n) := \det M^I_J\big(e(\alpha_1, \dots, \alpha_n)\big).$$

Note that the determinants s^I_J and e^I_J are alternating in I and J.

We show in (7.9) that the determinant $s^I = s^I_{[r]}$, defined according to the general notation in (7.4), is in fact equal to the Schur polynomial s^I of (6.13), evaluated at $(\alpha_1, \dots, \alpha_n)$.

Note that the similar notation cannot be used for the determinant $e^I_{[r]}$ since e^I, for $r = n$, was defined as the power product $e_1^{i_1} \cdots e_n^{i_n}$ in (5.10).

(7.8) Corollary. *Assume that I and J are strictly increasing multi indices of the same size r, and view I and J as subsets of the interval $[N]$ for some $N \gg 0$. Then:*

(1) (Vanishing) *The determinant $s^I_J = s^I_J(\alpha_1, \dots, \alpha_n)$ is non-zero only when $I \succeq J$ and the following inequalities hold for all the entries in the conjugate multi indices I' and J':*

$$i'_q \le j'_q + n. \tag{7.8.1}$$

In particular, if $r > n$, then $s^I_{[r]}$ is non-zero only if I is an extension of a strictly increasing multi index I_0 of size n; moreover, if I is an extension of I_0, then $s^I_{[r]} = s^{I_0}_{[n]}$.

(2) *For* $n = 1$, *the determinant* $s^I_J(\alpha)$ *is non-zero only when the following inequalities hold:*

$$j_1 \le i_1 < j_2 \le i_2 < \cdots \le i_{r-1} < j_r \le i_r. \tag{7.8.2}$$

Moreover, if the inequalities (7.8.2) *hold, then* $s^I_J(\alpha) = \alpha^{\|I\|-\|J\|}$.

(3) (Duality) *The following equation holds,*

$$s^I_J(\alpha_1, \dots, \alpha_n) = e^{I'}_{J'}(\alpha_1, \dots, \alpha_n), \tag{7.8.3}$$

where I' *and* J' *are the conjugate multi indices.*

(4) (Multiplication) *If* $(\beta_1, \dots, \beta_m)$ *is a second set of elements of the ground ring, then*

$$s^I_J(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m) = \sum_{I \succeq K \succeq J} s^I_K(\alpha_1, \dots, \alpha_n) s^K_J(\beta_1, \dots, \beta_m), \tag{7.8.4}$$

where the sum is over strictly increasing multi indices K *of size* r.

(5) (Jacobi–Trudi's Formula) *If* I *is a multi index of size equal to the number* n *of the* α_i, *then*

$$s^I_{[n]}(\alpha_1, \dots, \alpha_n)\,\Delta = \Delta^I(\alpha_1, \dots, \alpha_n), \tag{7.8.5}$$

where $\Delta = \Delta(\alpha_1, \dots, \alpha_n)$ *is the Vandermonde determinant.*

Proof. The duality formula in (3) follows from (7.5)(5) since $s(T)e(-T) = 1$. Similarly, the multiplication formula in (4) follows from (7.5)(6) since $s(\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m) = s(\alpha_1, \dots, \alpha_n)s(\beta_1, \dots, \beta_m)$.

Given the duality formula, the first vanishing statement in (1) follows by applying (7.5)(1) to the left side and to the right side of (7.8.3), since e is a polynomial of degree at most n. Consider the special case $J = [r]$. Then $I \succeq [r]$ since I is assumed to be strictly increasing. Assume that I is a subset of the interval $[N]$, where $N = r + t$. Then $J' = [t]$. So the inequalities (7.8.1) are the inequalities $i'_q \le q - 1 + n$ for $q = 1, \dots, t$. Obviously, they hold for all q if and only if $i'_t \le t - 1 + n$, that is, if and only if

$$r - n \le \tilde{i}_1, \tag{7.8.6}$$

where $\tilde{i}_1$ is the first entry in the complement $\tilde{I}$. The condition (7.8.6) is vacuous if $r \le n$. If $r > n$, then (7.8.6) holds if and only if I is an extension of a strictly increasing multi index I_0 of size n. Moreover, if I is an extension of I_0 then $s^I_{[r]} = s^{I_0}_{[n]}$ by (7.6.1). Hence the special vanishing assertion in (1) holds.

To prove (2) assume that $n = 1$. Then $s = s(\alpha) = 1 + \alpha T + \alpha^2 T^2 + \cdots$. Hence $M(s)$ is the matrix whose ij'th entry is α^{i-j} with the (strange) convention that $\alpha^k = 0$ if the exponent k is negative. Assume that $I \succeq J$ and consider the matrix $S := M^I_J(s)$. Its pp'th diagonal entry is the power $\alpha^{i_p - j_p}$ since $i_p - j_p$ is non-negative. The inequalities (7.8.2) imply that all entries above the diagonal are zero. Thus, if the inequalities (7.8.2) hold, then $s^I_J(\alpha)$ is the product of the diagonal entries and hence $s^I_J(\alpha) = \alpha^{\|I\|-\|J\|}$. Assume that the inequalities

(7.8.2) do not hold, and let $q < r$ be the first index for which $i_q \geq j_{q+1}$. Consider the p'th entries in the q'th and the $(q+1)$'st column in S. By the choice of q, both entries vanish if $p < q$ and if $p \geq q$ then the entries are the powers $\alpha^{i_p - j_q}$ and $\alpha^{i_p - j_{q+1}}$. Hence the q'th column is equal to $\alpha^{j_{q+1} - j_q}$ times the $(q+1)$'th column. Therefore, the determinant $s^I_J(\alpha) = \det S$ vanishes. Hence (2) has been proved.

Finally, to prove (5), consider for $j = 1, \dots, n$ the product $d^{(j)} := \prod_{i \neq j}(1 - \alpha_i T)$. Then $d^{(j)}$ is a polynomial of degree at most $n-1$, and we have the equation of power series $sd^{(j)} = (1 - \alpha_j T)^{-1}$. Hence it follows from (7.2.1) that we have the matrix equation,

$$M(s)\langle d^{(j)}\rangle = \Big\langle \frac{1}{1 - \alpha_j T} \Big\rangle.$$

Therefore, by definition (7.7) of the matrix $V = V(\alpha_1, \dots, \alpha_n)$, we have the equation,

$$M(s)\langle d^{(1)}, \dots, d^{(n)}\rangle = V. \tag{7.8.7}$$

Let D be the $n \times n$ matrix consisting of the first n rows of the $\infty \times n$ matrix $\langle d^{(1)}, \dots, d^{(n)}\rangle$. The matrix D contains all the non-zero rows, since each polynomial $d^{(j)}$ is of degree at most $n-1$. Therefore, from (7.8.7) we obtain the matrix equation,

$$M_{[n]}(s) D = V. \tag{7.8.8}$$

In (7.8.8), extract the equation corresponding to the rows in I and take determinants. Since $\det V^I = \Delta^I$ by (7.7.1), we obtain the equation,

$$s^I_{[n]}(\alpha_1, \dots, \alpha_n) \det D = \Delta^I(\alpha_1, \dots, \alpha_n). \tag{7.8.9}$$

Take $I := [n]$ in (7.8.9). On the left the determinant $s^{[n]}_{[n]}$ is equal to 1, and on the right the determinant is the Vandermonde determinant $\Delta = \Delta(\alpha_1, \dots, \alpha_n)$. It follows that $\det D = \Delta$. Now (7.8.5), for any multi index I of size n, follows from (7.8.9). ∎

(7.9) Corollary. *For any strictly increasing multi index J of size r contained in an interval $[r+t]$, we have the equalities,*

$$s^J(\alpha_1, \dots, \alpha_n) = s^J_{[r]}(\alpha_1, \dots, \alpha_n) = e^{J'}_{[t]}(\alpha_1, \dots, \alpha_n), \tag{7.9.1}$$

where the left side is the Schur polynomial of (6.13) *evaluated at α and the right sides are the determinants of* (7.7).

Proof. The second equation is the duality formula of (7.8.3) for $J = (0, 1, \dots, r-1)$.

Clearly, to prove the first equation, we may assume that the ground ring is the polynomial ring $R[A]$ and $\alpha_i = a_i$. When the size r is equal to n, the equation follows from Jacobi–Trudi's formula, since $s^J = \Delta^J/\Delta$ by definition of the Schur polynomials. If $r < n$, then J has an extension $\hat{J}$ to a multi index of size n, and it follows from (7.6.1) and (6.13) that $s^J_{[r]} = s^{\hat{J}}_{[n]} = s^{\hat{J}} = s^J$. The argument is similar for $r > n$, using the last part of (7.8)(1). ∎

(7.10) Application. Assume that the base ring is the polynomial ring $R[A] = R[a_1, \dots, a_n]$ over the alphabet A. The definitions of (7.7) and the results of (7.8) apply to any set of polynomials α_i, and not necessarily with the number of α_i equal to the number n of letters of A. However, a natural choice is $(\alpha_1, \dots, \alpha_n) := (a_1, \dots, a_n)$. Then the power series $s(a_1, \dots, a_n)$ and $e(a_1, \dots, a_n)$ of (7.7) are the power series $s(A)$ and $e(A)$ of (5.14). Their d'th coefficients are, respectively, the complete symmetric polynomial $s_d = s_d(A)$ and the elementary symmetric polynomial $e_d = e_d(A)$. Consequently, the matrix $M\big(s(a_1, \dots, a_n)\big)$ has as ij'th entry the complete symmetric polynomial s_{i-j} (equal to zero if $i < j$), and the matrix $M\big(e(a_1, \dots, a_n)\big)$ has as ij'th entry the elementary symmetric polynomial e_{i-j} (equal to zero if $i < j$ or $i > j + n$). It follows that the determinants $s^I_J = s^I_J(A)$ and $e^I_J = e^I_J(A)$, for multi indices I and J of the same size r, are symmetric polynomials in the letters of A. They are alternating in I and in J. Moreover, since s_d and e_d are homogeneous of degree d, it follows that the polynomials s^I_J and e^I_J are homogeneous of degree $\|I\| - \|J\|$.

The polynomials s^I_J are called *skew Schur polynomials*. By (7.9.1), the skew Schur polynomial $s^I_{[r]}$, for a strictly increasing multi index I of size r, is equal to the Schur polynomial s^I of (6.13).

Note finally that the notation u_L of (7.1) for a product of the coefficients of a power series u in the cases $u = e$ and $u = s$ is in accordance with the notations e_L and s_L of (6.14) and (6.15).

The duality formula of (7.8) is an explicit expression of the skew Schur polynomial s^I_J as a determinant in the elementary symmetric polynomials e_d. In particular, with $J = [r]$ and with the conjugate I' of size t, duality is the following formula for the Schur polynomial s^I:

$$s^I = e^{I'}_{[t]} = \begin{vmatrix} e_{i'_1} & e_{i'_1-1} & \dots & e_{i'_1-t+1} \\ \vdots & \vdots & & \vdots \\ e_{i'_t} & e_{i'_t-1} & \dots & e_{i'_t-t+1} \end{vmatrix}. \tag{7.10.1}$$

(7.11) Note. As noted above, the determinant $s^I_J(\alpha_1, \dots, \alpha_m)$ of (7.7) is defined for any set of polynomials α_j in $R[A]$. Of course, for an arbitrary set of polynomials α_j, the determinant is not a symmetric polynomial.

As an example, consider an alphabet $(A, B) = \{a_1, \dots, a_n, b_1, \dots, b_m\}$ obtained as the union of A and the letters b_i of a second alphabet B. Take $R[A, B] = R[A][B]$ as ground ring. Consider for strictly increasing multi indices I and J of size r the skew Schur polynomial $s^I_J(A, B)$. It is a symmetric polynomial in the letters of (A, B). In particular, it is symmetric in the letters of A and in the letters of B. By (7.8)(4), we have the formula,

$$s^I_J(A, B) = \sum_{I \succeq K \succeq J} s^I_K(A) s^K_J(B), \tag{7.11.1}$$

where the sum is over strictly increasing multi indices K of size r. For $J := [r]$, it follows that

$$s^I(A, B) = \sum_{I \succeq K} s^I_K(A) s^K(B). \tag{7.11.2}$$

Take as r the number m of letters of B. View the two sides of (7.11.2) as polynomials in $R[A][B]$. They are symmetric polynomials in the letters of B, with coefficients that are polynomials in the letters of A. The polynomials $s^K(B)$ on the right side are the proper Schur polynomials, and they form an $R[A]$-basis for $\mathrm{Sym}_{R[A]}[B]$. Therefore, the equation (7.11.2) is the expansion of $s^I(A, B)$ in terms of the Schur basis. In other words, the skew Schur polynomials $s^I_K(A)$ may be defined as the coefficients of the Schur polynomial $s^I(A, B)$ expanded in the basis $s^K(B)$.

As a second example, consider for a fixed letter a_k of A the polynomial $s^I_J(a_k)$. It follows from (7.8)(2) that $s^I_J(a_k)$ is non-zero only if the following inequalities hold,

$$j_1 \le i_1 < j_2 \le i_2 < \cdots \le i_{r-1} < j_r \le i_r. \tag{7.11.3}$$

Moreover, if the inequalities (7.11.3) hold, then $s^I_J(a_k) = a_k^{\|I\|-\|J\|}$. For reasons that will become more transparent later we will say that I/J is a *horizontal strip* if the inequalities (7.11.3) hold.

Clearly, for strictly increasing multi indices I and J of the same size r, we obtain by repeated application of (7.11.1) the formula,

$$s^I_J(A) = \sum_{I=K_0\succeq K_1\succeq\cdots\succeq K_n=J} s^{K_0}_{K_1}(a_1)\cdots s^{K_{n-1}}_{K_n}(a_n), \tag{7.11.4}$$

where the sum is over strictly increasing multi indices K_p of size r. As just observed, the sum is unchanged if we restrict it by the condition that each K_{p-1}/K_p is a horizontal strip, and then the corresponding term in the sum is the monomial,

$$a_1^{k_1}\cdots a_n^{k_n} \quad \text{where } k_p := \|K_{p-1} - K_p\| \text{ for } p = 1, \dots, n. \tag{7.11.5}$$

If $I \succeq J$ are multi indices of the same size r, then a *tableau of shape I/J and biggest entry n* is a sequence $T = (K_0, \dots, K_n)$ of strictly increasing multi indices such that $I = K_0 \succeq K_1 \succeq \cdots \succeq K_n = J$ and such that K_{p-1}/K_p is a horizontal strip for $p = 1, \dots, n$. With each tableau T there is an *associated monomial a^T* defined as the monomial (7.11.5). The following formula is simply a fancy rewriting of (7.11.4):

$$s^I_J(A) = \sum_T a^T, \tag{7.11.6}$$

where the sum is over all tableaux T of shape I/J and biggest entry n. It is a consequence of the formula that *the skew Schur polynomial s^I_J is a sum of monomials*. Equivalently, if s^I_J is expanded in the basis of monomial symmetric polynomials m^K, then the coefficients are non-negative integers. More precisely, the coefficient to m^K is the number of tableaux T for which $a^T = a^K$.

(7.12) Special cases.
Consider for the complete symmetric polynomial $s_d = s_d(A)$ the following two equations:

$$s_d = s^{0,\dots,n-2,n-1+d} = s^d.$$

The first was proved in (SYM.6.12.2), as a consequence of Pieri's Formula, the second was taken as a definition in (SYM.6.13). On the other hand, if the two sides of the second equation are interpreted as Schur polynomials via the determinantal definition, then their equality results from a trivial determinant consideration, and, obviously, $s^d = s^d_{[1]} = s_d$.

To obtain an expression in terms of the elementary symmetric polynomials note that the multi index $I := (d)$ is contained in the interval $[d+1] = \{0, 1, \dots, d\}$, and the conjugate multi index is $I' = (1, \dots, d)$. Hence, by duality,

$$s_d = e^{1,2,\dots,d}_{0,1,\dots,d-1} = \begin{vmatrix} e_1 & 1 & 0 & \dots & 0 \\ e_2 & e_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ e_{d-1} & e_{d-2} & e_{d-3} & \dots & 1 \\ e_d & e_{d-1} & e_{d-2} & \dots & e_1 \end{vmatrix}. \tag{7.12.1}$$

For instance,

$$s_4 = \begin{vmatrix} e_1 & 1 & 0 & 0 \\ e_2 & e_1 & 1 & 0 \\ e_3 & e_2 & e_1 & 1 \\ e_4 & e_3 & e_2 & e_1 \end{vmatrix} = e_1^4 - 3e_1^2 e_2 + 2e_1 e_3 + e_2^2 - e_4\,.$$

Similarly, since $I'' = I$, we obtain the formula, equivalent to the formula in (6.13),

$$e_d = s^{1,2,\dots,d}. \tag{7.12.2}$$

Since $s(T)e(-T) = 1$, the formulas (7.12.1) and (7.12.2) could have been deduced directly from (7.3). As a direct application of (7.3), consider the power series $p = p(A)$ defined in (5.14). The d'th coefficient is the power sum $p_{d+1} = p_{d+1}(A)$. Since $e(T)p(-T) = e'(T)$, we obtain from (7.3.1) the formula,

$$p_d = \begin{vmatrix} e_1 & 1 & 0 & \dots & 0 \\ 2e_2 & e_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (d-1)e_{d-1} & e_{d-2} & e_{d-3} & \dots & 1 \\ de_d & e_{d-1} & e_{d-2} & \dots & e_1 \end{vmatrix}. \tag{7.12.3}$$

For instance,

$$p_4 = \begin{vmatrix} e_1 & 1 & 0 & 0 \\ 2e_2 & e_1 & 1 & 0 \\ 3e_3 & e_2 & e_1 & 1 \\ 4e_4 & e_3 & e_2 & e_1 \end{vmatrix} = e_1^4 - 4e_1^2 e_2 + 4e_1 e_3 + 2e_2^2 - 4e_4.$$

8. Base change.

(8.1) Setup. Several identities of symmetric polynomials are most easily expressed as equations in the power series ring $R[[A]]$ over the n letters of the alphabet A. Recall that an element f of $R[[A]]$ is an infinite sequence $f = (f_0, f_1, \dots)$ such that f_i is a homogeneous polynomial of degree i in $R[A]$. The polynomial f_i is the i'th *term* in f. If $f_i = 0$ for all $i < d$, then f is said to have *order* at least d. A formal series,

$$\sum_{\iota} F_\iota, \tag{8.1.1}$$

over any set of indices ι, of elements F_ι in $R[[A]]$ is called *convergent* if, for any d, all but a finite number of F_ι have order greater than d. When the series is convergent, we may view the sum (8.1.1) as the element in $R[[A]]$ whose d'th term is the sum of the d'th terms in the F_ι.

We will often use that two series in $R[[A]]$ are equal iff for all $N \gg 0$ they are equal modulo the ideal $(A)^N = (a_1, \dots, a_n)^N$ of series of order at least N.

(8.2) Proposition. *Let u be a power series in $R[[T]]$. For decreasing multi indices K of size n and strictly increasing multi indices I of size n consider the monomial symmetric polynomials $m^K(A)$ and the proper Schur polynomials $s^I(A)$. Let J be any multi index of size n. Then the following two formulas hold in $R[[A]]$:*

$$\prod_{a\in A} u(a) = \sum_K u_K m^K(A), \tag{8.2.1}$$

$$s^J(A)\prod_{a\in A} u(a) = \sum_I u^I_J s^I(A). \tag{8.2.2}$$

Proof. The two right sides are well-defined since the polynomials m^K and s^I are homogeneous of degrees, respectively, $\|K\|$ and $\|I\| - n(n-1)/2$.

The first formula is obtained by a simple multiplication of series,

$$\prod_{a\in A} u(a) = \prod_{q=1}^{n}\sum_{l=0}^{\infty} u_l a_q^l = \sum u_{l_1}\cdots u_{l_n} a_1^{l_1}\cdots a_n^{l_n},$$

where the last sum is over all multi indices $L = (l_1, \dots, l_n)$ of size n. Each multi index L is a permutation of a unique decreasing multi index K, and $u_L = u_K$. Hence the last sum is the right hand side of (8.2.1). Thus the first formula holds.

To prove the second formula, consider an element α in $R[[A]]$ without constant term. Then the evaluation $u(\alpha) = \sum_{j=0}^{\infty} u_j\alpha^i$ can be obtained by multiplying the infinite row $\langle u\rangle^{\mathrm{tr}} = (u_0, u_1, u_2, \dots)$ and the infinite column $V(\alpha)$ with entries $1, \alpha, \alpha^2, \dots$. Applied with $u := T^i u$, it follows that $u(\alpha)\alpha^i = \langle T^i u\rangle^{\mathrm{tr}} V(\alpha)$. Hence, in the notation of (7.2), we have the matrix equation,

$$u(\alpha)V(\alpha) = M(u)^{\mathrm{tr}} V(\alpha).$$

Applied with $\alpha := a_q$ for $q = 1, \dots, n$, we obtain the matrix equation,

$$V(a_1, \dots, a_n)\,\mathrm{diag}\big(u(a_1), \dots, u(a_n)\big) = M(u)^{\mathrm{tr}} V(a_1, \dots, a_n).$$

Extract the equations corresponding to the rows in J to obtain the matrix equation,

$$V^J(a_1, \dots, a_n)\mathrm{diag}\big(u(a_1), \dots, u(a_n)\big) = M_J(u)^{\mathrm{tr}} V(a_1, \dots, a_n). \tag{8.2.3}$$

Now it suffices to prove the following equation for the determinant of the right side of (8.2.3):

$$\det\Big(M_J(u)^{\mathrm{tr}} V(a_1, \dots, a_n)\Big) = \textstyle\sum_I u_J^I \Delta^I(a_1, \dots, a_n). \tag{8.2.4}$$

Indeed, if (8.2.4) holds, then (8.2.3) yields the following equation of determinants,

$$\Delta^J(a_1, \dots, a_n) \textstyle\prod_q u(a_q) = \sum_I u_J^I \Delta^I(a_1, \dots, a_n),$$

and, clearly, the latter equation implies (8.2.2) after division by $\Delta(a_1, \dots, a_n)$.

It remains to verify equation (8.2.4). It suffices to show that the equation holds modulo the ideal $(A)^N$ for $N \gg 0$. Take N larger than every coordinate of J. In the matrix $V(a_1, \dots, a_n)$ the k'th row consists of the powers a_i^k. So the class modulo $(A)^N$ of the matrix is unchanged if we replace all rows of index $k \geq N$ with zeros. It follows easily that we have the following congruence of matrices:

$$M_J(u)^{\mathrm{tr}} V(a_1, \dots, a_n) \equiv \big(M_J^{[N]}(u)\big)^{\mathrm{tr}} V^{[N]}(a_1, \dots, a_n) \quad \mathrm{mod}\ (A)^N.$$

Take determinants, and apply the Cauchy–Binet Formula, see Appendix (1.2), to the product on the right side to obtain the congruence,

$$\det\Big(M_J(u)^{\mathrm{tr}} V(a_1, \dots, a_n)\Big) \equiv \textstyle\sum_{I \subseteq [N]} u_J^I \Delta^I(a_1, \dots, a_n) \quad \mathrm{mod}\ (A)^N. \tag{8.2.5}$$

Finally, the sum on the right side of (8.2.5) is, modulo $(A)^N$, equal to the sum in (8.2.4), since the extra terms, for $I \not\subseteq [N]$, vanish modulo $(A)^N$.

Thus (8.2.4) has been verified modulo $(A)^N$, and the proof is complete. □

(8.3) Corollary. *Let B be a second alphabet with m letters. Let J be a multi index of size n. Then the following two formulas hold in $R[B][[A]]$:*

$$\prod_{a \in A,\, b \in B} \frac{1}{1 - ab} = \sum_K s_K(B) m^K(A), \tag{8.3.1}$$

$$s^J(A) \prod_{a \in A,\, b \in B} \frac{1}{1 - ab} = \sum_I s_J^I(B) s^I(A), \tag{8.3.2}$$

where the first sum is over all decreasing multi indices K of size n and the second sum is over all strictly increasing multi indices I of size n. In addition, the following formula holds in $R[A, B]$:

$$\prod_{a\in A,\, b\in B} (a+b) = \sum_{I\subseteq[n+m]} s^I(A)s^{\tilde{I}}(B)\,, \tag{8.3.3}$$

where the sum is over size-n subsets I and $\tilde{I} \subseteq [n+m]$ is the complement of I.

Proof. Replace in (8.2) R by $R[B]$. Clearly, the first two formulas follow from the Proposition by taking $u := s(B) = \prod_{b\in B}(1-bT)^{-1}$.

To prove the third formula, take $u := \prod_{b\in B}(T+b)$ and $J := (0, 1, \dots, n-1)$ in (8.2.2). We obtain the equation,

$$\prod_{a\in A,\, b\in B} (a+b) = \sum_I u^I_{[n]} s^I(A)\,. \tag{8.3.4}$$

The series u is a polynomial of degree m. In particular, if I is not contained in $[n+m]$, then the last row in $M^I_{[n]}(u)$ vanishes, and thus $u^I_{[n]} = 0$. Hence in (8.3.4) we may assume that the sum is over subsets $I \subseteq [n+m]$ with n elements.

Consider the series $e := e(B)(T)$ associated with the alphabet B, see (5.14.1). The i'th coefficient e_i is the i'th elementary symmetric polynomial in the letters of B. Clearly, $u_i = e_{m-i}$ for all i, and both sides vanish unless $0 \le i \le m$. As a consequence, if the order of the rows and the order of the columns in the matrix $M^{[n+m]}_{[n]}(u)$ are reversed, then the matrix $M^{[m+n]}_{[n]}(e)$ is the result. The reversions change $M^I_{[n]}(u)$ into $M^{I^*}_{[n]}(e)$, where I^* is the image of I under the involution $i \mapsto m+n-1-i$ of $[m+n]$. The two reversions do not change the determinant, and so $u^I_{[n]} = e^{I^*}_{[n]}$. Finally, by Duality (7.5)(5) applied to the series $e(B)$ and $s(B)$, we have $e^{I^*}_{[n]}(B) = s^{\tilde{I}}_{[m]}(B) = s^{\tilde{I}}(B)$. Thus (8.3.3) follows from (8.3.4). □

(8.4) Definition. Define in $\mathrm{Sym}_R[A]$ an R-bilinear form, denoted $(g, h) \mapsto (g \mid h)$, by the equations, for strictly increasing multi indices I and J of size n,

$$(s^I \mid s^J) := \delta_{I,J}\,,$$

where $\delta_{I,J}$ is Kronecker's δ. In other words, if symmetric polynomials g and h are expanded in the basis of proper Schur polynomials s^J, say $g = \sum \alpha_J s^J$ and $h = \sum \beta_I s^I$, then

$$(g \mid h) = \sum\nolimits_J \alpha_J \beta_J\,.$$

Clearly, the bilinear form is symmetric. It is called the *inner product* in $\mathrm{Sym}_R[A]$. It follows from the definition that the inner product $(g \mid s^J)$ is equal to the coefficient to s^J when the symmetric polynomial g is expanded in the Schur basis.

(8.5) Proposition. (1) *If K and L are decreasing multi indices of size n, then*

$$(s_L \mid m^K) = \delta_{L,K}\,. \tag{8.5.1}$$

(2) *If I, J, and K are strictly increasing multi indices of size n, then*

$$(s^I_J \mid s^K) = (s^I \mid s^J s^K). \tag{8.5.2}$$

Proof. (1) Consider the expansions of $s_L(A)$ and $m^K(A)$ in the basis of proper Schur polynomials,

$$s_L(A) = \sum_J \lambda_{L,J} s^J(A), \quad m^K(A) = \sum_J \mu_{K,J} s^J(A). \tag{8.5.3}$$

Then $(s_L \mid m^K) = \sum_J \lambda_{L,J}\mu_{K,J}$. In other words, if λ and μ denote the matrices of the $\lambda_{L,J}$ and $\mu_{L,J}$, for all L, J, then (1) holds if and only if the product matrix $\lambda\mu^{\mathrm{tr}}$ is the unit matrix 1. It is easy to prove the equivalence of the two equations $\lambda\mu^{\mathrm{tr}} = 1$ and $\mu^{\mathrm{tr}}\lambda = 1$. Indeed, this assertion would have been trivial if the matrices had finite order. To reduce this assertion, and assertions in the following paragraphs, to the case of finite order, it suffices to note that the bases consist of homogeneous polynomials, with a finite number of base elements of each degree. Each assertion may be reduced to a, possibly infinite, number of assertions each involving only a finite number of terms; in particular, to an assertion involving only polynomials with a given bound, say d, of the degree: so we may assume here that $\|K\|$, $\|L\| \le d$, $\|I\|$, $\|J\| \le d + n(n-1)/2$, etc.

With this note in mind, it suffices to prove that $\mu^{\mathrm{tr}}\lambda = 1$, that is, for all strictly increasing I and J,

$$\sum_K \mu_{K,I}\lambda_{K,J} = \delta_{I,J}. \tag{8.5.4}$$

To prove (8.5.4), let B be a second alphabet with n letters. Take $J := [n]$ in (8.3). It follows from the two equations (8.3.1) and (8.3.2) that

$$\sum_K s_K(B) m^K(A) = \sum_I s^I(B) s^I(A). \tag{8.5.5}$$

The expansions (8.5.3) hold when A is replaced by B. Insert the expansions of $s_K(B)$ and $m^K(A)$ in (8.5.5). The result is the equation in $R[[B, A]]$,

$$\sum_{K,I,J} \lambda_{K,J}\mu_{K,I} s^J(B) s^I(A) = \sum_I s^I(B) s^I(A). \tag{8.5.6}$$

It follows from (8.5.6), since the Schur polynomials $s^I(A)$ form a basis for the polynomials that are symmetric in the letters of A, that for any fixed strictly increasing multi index I we have the equation in $R[[B]]$,

$$\sum_{K,J} \lambda_{K,J}\mu_{K,I} s^J(B) = s^I(B). \tag{8.5.7}$$

Again, since the $s^I(B)$ form a basis for the polynomials that are symmetric in the letters of B, it follows from (8.5.7) that (8.5.4) holds. Hence (1) has been proved.

(2) The proof of (8.5.2) is similar. Consider the expansions of s^I_J and $s^J s^K$ in the basis of proper Schur polynomials,

$$s^I_J = \sum_K \lambda_{I,J,K} s^K, \quad s^J s^K = \sum_I \mu_{J,K,I} s^I, \tag{8.5.8}$$

where both sums are over strictly increasing multi indices of size n. Then $\lambda_{I,J,K} = (s^I_J \mid s^K)$ and $(s^I \mid s^J s^K) = \mu_{J,K,I}$. Thus (8.5.2) is the equation $\lambda_{I,J,K} = \mu_{J,K,I}$.

From the equation (8.3.2) and the same equation applied with $J := [n]$ and multiplied by $s^J(A)$ it follows that

$$\sum\nolimits_I s^I_J(B)s^I(A) = s^J(A)\sum\nolimits_K s^K(B)s^K(A).$$

Insert the expansions (8.5.8) to obtain the equation,

$$\sum\nolimits_{I,K} \lambda_{I,J,K} s^K(B)s^I(A) = \sum\nolimits_{I,K} \mu_{J,K,I} s^K(B)s^I(A). \tag{8.5.9}$$

As in the proof of (1), it follows from (8.5.9) that $\lambda_{I,J,K} = \mu_{J,K,I}$, which is the asserted equation (8.5.2). ∎

(8.6) Note. It follows from (8.5.2) that the problem of determining in "the Schur basis" (basis of proper Schur polynomials) the coefficients in the expansions of the skew Schur polynomials s^I_J is the same as the problem of determining the coefficients in the expansions of all products $s^J s^K$. The coefficients are in fact non-negative, and given by a combinatorial rule, called the Littlewood–Richardson rule.

(8.7) Corollary. *The products $s_L = s_L(A)$, for all decreasing multi indices L of size n, form an R-basis for* $\mathrm{Sym}_R[A]$.

Proof. The assertion follows from (8.5.1) since the m^K form a basis. More precisely, if a symmetric polynomial g is an R-linear combination of the products s_L, say

$$g = \sum\nolimits_L \alpha_L s_L, \tag{8.7.1}$$

then it follows from (8.5.1) that $\alpha_K = (g \mid m^K)$. Hence the coefficients in (8.7.1) are uniquely determined by g. To prove the existence, consider the inner products $\alpha_K := (g \mid m^K)$ for decreasing multi indices K of size n. The expansion of g in the basis of Schur polynomials involves only Schur polynomials s^J of degree at most equal to the degree of g. The expansion of m^K involves only Schur polynomials s^J of degree equal to the degree $\|K\|$. Hence the inner product α_K vanishes when $\|K\|$ is bigger than the degree of g. In particular, only a finite number of α_K are non-zero. We claim that the equation (8.7.1) holds. To prove it, consider the difference, $\tilde g := g - \sum_L \alpha_L s_L$. It follows from (8.5.1) that $(\tilde g \mid m^K) = 0$ for all m^K. Since the m^K form an R-basis for $\mathrm{Sym}_R[A]$, it follows that $(\tilde g \mid h) = 0$ for all symmetric polynomials h. In particular $(\tilde g \mid s^J) = 0$ for all Schur polynomials s^J. As a consequence, $\tilde g = 0$. Hence the equation (8.7.1) holds. ∎

(8.8) Definition. It follows from Theorem (5.10) that the R-algebra $\mathrm{Sym}_R[A]$ of symmetric polynomials is the free polynomial ring over R in the elementary symmetric polynomials $e_1, \dots, e_n$. In particular, there is a unique R-algebra endomorphism of $\mathrm{Sym}_R[A]$ such that $e_d \mapsto s_d$ for $d = 1, \dots, n$. By definition, if I is a multi index of size n, we have that

$$e^I = e_1^{i_1} \cdots e_n^{i_n} \mapsto s_1^{i_1} \cdots s_n^{i_n}. \tag{8.8.1}$$

Note that s^I is the notation of a Schur polynomial, and hence it can not be used as a notation for the right hand side of (8.8.1). In this context, it is common to denote, for $d = 0, 1, \dots,$ the d'th complete symmetric polynomial s_d also by h_d. Accordingly, we define, for a multi index I of arbitrary size r,

$$h^I := h_1^{i_1} \cdots h_r^{i_r}, \qquad h_I := h_{i_1} \cdots h_{i_r}. \tag{8.8.2}$$

With this notation, the endomorphism of $\mathrm{Sym}_R[A]$ is given by $e^I \mapsto h^I$ for multi indices I of size n. The endomorphism is denoted $g \mapsto g^*$. Note that e^I and e_I are defined for multi indices of any size r, but the equation,

$$(e^I)^* = h^I, \tag{8.8.3}$$

holds only if $i_p = 0$ for $n < p \le r$ (empty condition when $r \le n$). In particular, $e_d^* = h_d$ for $1 \le d \le n$ (in fact, for all $d \le n$, since $e_0 = h_0 = 1$ and $e_d = h_d = 0$ for $d < 0$).

(8.9) Lemma. *The endomorphism $g \mapsto g^*$ is an involution of the graded R-algebra $\mathrm{Sym}_R[A]$. Let K be a decreasing multi index and let J be a strictly increasing multi index, both of size r. Consider the following two equations:*

$$(e_K)^* = h_K, \quad \left(s^J\right)^* = s^{J'}. \tag{8.9.1}$$

(1) *If $k_p \le n$ for all entries in K, then the first equation holds.*
(2) *If $j_q \le n$ for all entries in J, then the second equation holds.*
(3) *In particular, both equations hold in degree at most n.*
(4) *Moreover, if f and g are symmetric polynomials of degree at most n, then*

$$(f^* \mid g^*) = (f \mid g). \tag{8.9.2}$$

Proof. The endomorphism is graded as e_d and $e_d^* = s_d$ are homogeneous of the same degree. To prove that it is an involution, use that the power series $e = \prod_a (1 + aT)$ and $s = \prod_a (1 - aT)^{-1}$ are related by the equation $e(T)s(-T) = 1$. In particular, the complete symmetric polynomials s_d, for $d = 1, \dots, n$, are determined from the elementary symmetric polynomials e_d by the congruence,

$$(1 + e_1 T + \cdots + e_n T^n)(1 + s_1(-T) + \cdots + s_n(-T)^n) \equiv 1 \pmod{T^{n+1}}.$$

Substitute $T := -T$ in the congruence, and apply the endomorphism $g \mapsto g^*$. On the left, the first factor is changed to the second. Therefore, since the constant 1 on the right side of the congruence is unchanged, the second factor is changed to the first. Hence we have, for $d = 1, \dots, n$, the equation $s_d^* = e_d$, that is, $e_d^{**} = e_d$. Since $\mathrm{Sym}_R[A]$ is generated as an R-algebra by $e_1, \dots, e_n$ and $g \mapsto g^*$ is an R-algebra endomorphism, it follows that $g^{**} = g$. Thus the endomorphism is an involution.

The two equations in (8.9.1) hold trivially if $r = 0$, so assume that $r \geq 1$. For Assertion (1), assume that $k_p \leq n$ for all p. Then each factor in e_K is of the form e_k with $1 \leq k \leq n$, and hence $e_k^* = s_k = h_k$ by definition. Clearly, the first equation is a consequence.

For Assertion (2), the multi index J is strictly increasing, and $j_r \leq n$. Thus J is a subset of the interval $[n+1]$. As J' may be determined with respect to this interval, the size of J' is t with $t + r = n + 1$, and $j'_t \leq n$. By Duality (7.5)(5), we have the equation,

$$s^J = \det M_{[t]}^{J'}(e)\,. \tag{8.9.3}$$

For any power series u the matrix $M_{[t]}^{J'}(u)$ has as its pq'th entry the coefficient u_d for $d = j'_p - (q-1)$ (where $u_d = 0$ when $d < 0$). The largest possible d is $d = j'_t$, obtained for $p = t$ and $q = 1$. As $j'_t \leq n$ it follows that every entry in the matrix equals u_d for some $d \leq n$. For these values of d we have $e_d^* = s_d$. Therefore, the involution $g \mapsto g^*$ maps $M_{[t]}^{J'}(e)$ to $M_{[t]}^{J'}(s)$. Consequently, by applying the involution $g \mapsto g^*$ to (8.9.3), we obtain the asserted second equation,

$$\bigl(s^J\bigr)^* = \det M_{[t]}^{J'}(s) = s^{J'}\,.$$

For Assertion (3), assume that the left sides of (8.9.1) are non-zero of degree at most n. For the first equation, note that, obviously, if the product $e_{k_1} \cdots e_{k_r}$ is non-zero, then $k_p \leq n$ for all p (independent of the degree of e_K). In particular, the first equation holds by (1). Consider the second equation. If J is an extension of a smaller multi index, then the two sides are unchanged if J is replaced by the smaller multi index. Hence we may assume that J is not an extension of a smaller multi index, that is, we may assume that $j_1 > 0$. Since J is strictly increasing, it follows that $j_p \geq p$ for $p = 1, \dots, r$. The degree of s^J is at most n, that is, $n \geq \|J\| - r(r-1)/2$. Hence, by the inequalities $j_p \geq p$ for $p = 1, \dots, r-1$ it follows that

$$n \geq \|J\| - r(r-1)/2 = \sum\nolimits_{p=1}^{r} \bigl(j_p - (p-1)\bigr) \geq (r-1) + j_r - (r-1) = j_r.$$

Hence $j_r \leq n$. As J is increasing, the asserted equality is now a consequence of (2).

Finally, to prove Assertion (4), note that the proper Schur polynomials s^J of degree at most n form a basis for the module of symmetric polynomials of degree at most n. It follows from the second equation of (8.9.1) that in this basis, the involution $g \mapsto g^*$ is a permutation of the basis elements. By definition, the proper Schur polynomials form an orthonormal basis with respect to the inner product. Therefore (8.9.2) holds, and (4) has been proved. □

(8.10) Definition. Up to now we have found several bases for the R-module $\mathrm{Sym}_R[A]$ of symmetric polynomials. The bases $m^\bullet = \{m^K\}$ of (5.8) and $h_\bullet = \{h_K\} = \{s_K\}$ of (8.7) are indexed by decreasing multi indices K of size n. The basis $e^\bullet = \{e^I\}$ of (5.10) is indexed by arbitrary multi indices I of size n. The basis $s^\bullet = \{s^J\}$ of (6.7) is indexed by strictly increasing multi indices of size n. Since $g \mapsto g^*$ is an automorphism of $\mathrm{Sym}_R[A]$, it follows that the products $h^I = (e^I)^*$, for multi indices I of size n, form a basis $h^\bullet$.

Consider an arbitrary basis $g^\bullet = \{g^I\}$ (where I runs through some suitable index set) of $\mathrm{Sym}_R[A]$. For any symmetric polynomial f, denote by $C(f, g^\bullet)$ the row of coefficients in the expansion of f in the basis $g^\bullet$. In a similar notation, denote by $(f \mid g^\bullet)$ the row of inner products $(f \mid g^I)$. For instance, with respect to the basis $s^\bullet$ of proper Schur polynomials we have, as noted in (8.4),

$$C(f, s^\bullet) = (f \mid s^\bullet). \tag{8.10.1}$$

Similarly, if $f^\bullet = \{f^K\}$ is a second basis, denote by $C(f^\bullet, g^\bullet)$ the matrix whose K'th row is $C(f^K, g^\bullet)$, and denote by $(f^\bullet \mid g^\bullet)$ the matrix whose K'th row is $(f^K \mid g^\bullet)$. For instance, it follows from (8.10.1) that

$$C(f^\bullet, s^\bullet) = (f^\bullet \mid s^\bullet). \tag{8.10.2}$$

Moreover, it follows from (8.5.1) that

$$C(f^\bullet, m^\bullet) = (f^\bullet \mid h_\bullet), \quad C(f^\bullet, h_\bullet) = (f^\bullet \mid m^\bullet). \tag{8.10.3}$$

The matrices $C(f^\bullet, g^\bullet)$ are infinite matrices. In general, it is assumed that the symmetric polynomials of a basis are homogeneous. Then the part of the basis consisting of polynomials of fixed degree d is a finite basis for the R-module of homogeneous symmetric polynomials of degree d. Accordingly, the matrix $C(f^\bullet, g^\bullet)$ may be viewed as a sequence of quadratic matrices where the part in degree d is obtained from the parts of $f^\bullet$ and $g^\bullet$ in degree d.

Note that the two bases, $h_\bullet = \{h_K\}$ indexed by decreasing multi indices K of size n and $h^\bullet = \{h^I\}$ indexed by arbitrary multi indices of size n, agree in degree at most n, but not in degree bigger than n. For instance, h_{n+1} is part of the first basis and not of the second, and h_1^{n+1} is part of the second and not of the first.

Note also that the products e_K, for decreasing multi indices K of size n, do not form a basis, since $e_K = 0$ if $k_1 > n$. However, the products e_K, for decreasing multi indices K of size n and $\|K\| \le n$, form a basis $e_\bullet$ for the symmetric polynomials of degree at most n, equal to the part of degree at most n of the basis $e^\bullet$.

(8.11) Definition. Consider in particular the matrix $C := C(s^\bullet, m^\bullet)$. Its IL'th entry C_{IL}, for a strictly increasing multi index I and a decreasing multi index L, both of size n, is determined by the expansion,

$$s^I = \sum\nolimits_L C_{IL} m^L, \tag{8.11.1}$$

of the Schur polynomial s^I in terms of the basis of monomial symmetric functions. It follows from (7.10) that the entries C_{IL} are non-negative integers, determined combinatorically as a number of tableaux with certain properties. The numbers C_{IL} are called the *Kostka numbers*.

The Kostka number $C_{I,L}$ and, more generally, the coefficient $C_{I,L}$ in the expansion (8.11.1) of s^I for a strictly increasing multi index I of any size r, may be determined as follows:

There is a bijective correspondence between strictly increasing multi indices of size r and weakly decreasing multi indices of size r, given by $J \mapsto \bar{J}$, where $\bar{J} = (j_r - (r-1), \dots, j_2 - 1, j_1)$. Hence a sequence $K_0, \dots, K_n$ of strictly increasing multi indices of size r corresponds to an $r \times (n+1)$ matrix T of non-negative integers whose q'th column, for $q = 0, \dots, n$, is the weakly decreasing sequence $\bar{K}_q$. The relations $K_0 \succeq \cdots \succeq K_n$ correspond to the condition that the entries in each row of T are weakly decreasing. The condition that each K_{q-1}/K_q is a horizontal strip corresponds to the condition that the entries in each skew diagonal of T (southwest to northeast) are weakly decreasing. The two conditions on T, on the rows and on the skew diagonals, are called the *tableau conditions*. Note that the tableau conditions imply that the entries in each column of T, which is not the last column, are weakly increasing. Hence a tableau of shape $I/[r]$, as defined in (7.11), can be identified with a matrix T satisfying the tableau conditions and such that the first column of T is equal to $\bar{I}$ and the last column consists of zeros. For each tableau, let t_q denote the sum of the entries in the q'th column of the matrix. Clearly, $\|K_{q-1} - K_q\| = t_{q-1} - t_q$. Hence it follows from (7.11.1) that

$$s^I = \sum\nolimits_T a_1^{t_0-t_1} \cdots a_n^{t_{n-1}-t_n}.$$

In particular, the coefficient C_{IL} is equal to the number of matrices T satisfying the above conditions and the equation

$$(t_0 - t_1, \dots, t_{n-1} - t_n) = L. \tag{8.11.2}$$

(8.12) Lemma. *Let $I = (i_1, \dots, i_n)$ be a strictly increasing multi index and let $L = (l_1, \dots, l_n)$ be a decreasing multi index. Form the decreasing multi index $\bar{I} = (i_n - (n-1), \dots, i_2 - 1, i_1)$. Then the Kostka number C_{IL} vanishes unless $\|\bar{I}\| = \|L\|$ and the following inequalities hold:*

$$\bar{i}_1 + \cdots + \bar{i}_q \geq l_1 + \cdots + l_q \quad \text{for } q = 1, \dots, n. \tag{8.12.1}$$

Moreover, if all the inequalities are equalities, that is, if $L = \bar{I}$, then $C_{IL} = 1$.

Proof. The Kostka number C_{IL} is the number of $n \times (n+1)$ matrices T satisfying the conditions of (8.11).

Assume that $C_{IL} \neq 0$. Then there is a matrix $T = (t_{pq})$ satisfying the conditions. The first column T_0 of T is $\bar{I}$, and the last column T_n of T consists of zeros. Clearly, the tableau conditions on the skew diagonals of T imply that the last q entries in the q'th column are equal to 0. The column sum t_0 is equal to $\|\bar{I}\|$. Hence it follows from (8.11.2) that

$$l_1 + \cdots + l_q = t_0 - t_q = \sum\nolimits_{p=1}^{n} \bar{i}_p - \sum\nolimits_{p=1}^{n-q} t_{pq}. \tag{8.12.2}$$

Since the entries along the skew diagonal are weakly increasing, it follows for a term t_{pq} in the last sum that

$$t_{pq} \geq t_{p+1,q-1} \geq \cdots \geq t_{p+q,0} = \bar{i}_{p+q}. \tag{8.12.3}$$

Hence the inequalities (8.12.1) follow from (8.12.2) and (8.12.3).

Clearly, if the inequalities are equalities, then T is unique and determined by equalities in (8.12.3). Hence the last assertion of the Lemma holds. $\square$

(8.13) Proposition. *Let $C = C(s^\bullet, m^\bullet)$ be the matrix of Kostka numbers. Let S be the quadratic matrix indexed by strictly increasing multi indices I, J of size n given by*

$$S_{IJ} = \begin{cases} 1 & \text{if } I \text{ and } J' \text{ are extensions of the same multi index,} \\ 0 & \text{otherwise.} \end{cases}$$

Then S is a symmetric matrix. Moreover, the following formulas hold:

$$C(s^\bullet, m^\bullet) = C, \tag{8.13.1}$$

$$C(h_\bullet, s^\bullet) = C^{\mathrm{tr}}, \tag{8.13.2}$$

$$C(h_\bullet, m^\bullet) = C^{\mathrm{tr}} C, \tag{8.13.3}$$

and, in degree at most n,

$$C(e_\bullet, s^\bullet) = C^{\mathrm{tr}} S, \tag{8.13.4}$$

$$C(e_\bullet, m^\bullet) = C^{\mathrm{tr}} S C. \tag{8.13.5}$$

In particular, in degree at most n, the matrices $C(h_\bullet, m^\bullet)$ and $C(e_\bullet, m^\bullet)$ are symmetric.

Proof. The matrix S is symmetric because conjugation $J \mapsto J'$ is an involution. The formula (8.13.1) is the definition of the matrix C. As noted in (8.10), we have that

$$C(h_\bullet, s^\bullet) = (h_\bullet \mid s^\bullet) = (s^\bullet \mid h_\bullet)^{\mathrm{tr}} = C(s^\bullet, m^\bullet)^{\mathrm{tr}}.$$

Hence (8.13.2) holds.

Clearly, $C(h_\bullet, m^\bullet) = C(h_\bullet, s^\bullet) C(s^\bullet, m^\bullet)$. Hence (8.13.3) follows from (8.13.1) and (8.13.2).

Restrict to the parts of degree at most n. Then it follows from (8.9)(3) that $S = C(s^\bullet, (s^\bullet)^*)$. Therefore, by (8.9), (4) and (3), we have that

$$C(e_\bullet, s^\bullet) = C((e_\bullet)^*, (s^\bullet)^*) = C((e_\bullet)^*, s^\bullet)\, C(s^\bullet, (s^\bullet)^*) = C(h_\bullet, s^\bullet) S.$$

Hence (8.13.4) follows from (8.13.2). Finally, (8.13.5) follows from (8.13.1) and (8.13.4) since $C(e_\bullet, m^\bullet) = C(e_\bullet, s^\bullet)\, C(s^\bullet, m^\bullet)$. □

(8.14) Remark. The R-bases found for $\mathrm{Sym}_R[A]$ are independent of R. For instance, if $R \subseteq S$ is a subring, then $\mathrm{Sym}_R[A] \subseteq \mathrm{Sym}_S[A]$ is a subring, and the basis of, say, monomial symmetric polynomials is an R-basis for $\mathrm{Sym}_R[A]$ and an S-basis for $\mathrm{Sym}_S[A]$. It follows from standard linear algebra, that if a family $(f_j)_{j \in J}$ of symmetric polynomials in $\mathrm{Sym}_R[A]$ is an R-basis, then the family is also an S-basis for $\mathrm{Sym}_S[A]$.

9. Partitions.

(9.1) Definition. A *partition* is a decreasing sequence $\lambda = (\lambda_1, \lambda_2, \dots)$ of non-negative integers containing only a finite number of positive terms. The term λ_i is called the *i'th part* of the partition λ.

The number of positive parts is called the *length* of the partition, the sum of the parts is called the *degree* of the partition and denoted $\|\lambda\|$. The *biggest part* of the partition is the first part λ_1, since the sequence is decreasing.

There are several convenient notations for partitions. First, we may indicate a partition by giving any finite subsequence containing all the positive parts, in particular the finite sequence containing only the positive parts. For instance, each of the sequences (7, 7, 3, 3, 3, 1) and (7, 7, 3, 3, 3, 1, 0, 0) represent the partition,

$$\lambda = (7, 7, 3, 3, 3, 1, 0, \dots). \tag{9.1.1}$$

The length of λ is 6, the degree of λ is 24, and the biggest part of λ is 7.

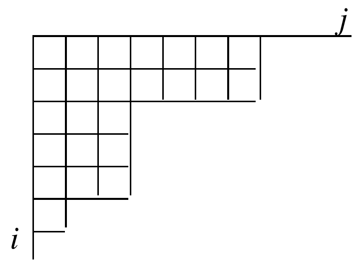
In this notation, the *zero-partition* $(0, 0, \dots)$, also called the *empty partition*, is represented by any finite sequence of zeros, in particular by the empty sequence ( ).

Next, a partition λ may be given by its *type*, that is, by the numbers $m_p = m_p(\lambda)$ counting, for $p = 1, 2, \dots$, the number of parts of λ that are equal to p. The type is often indicated by the "formal" product $1^{m_1}2^{m_2}\cdots$. For instance, the partition λ of (9.1.1) may be given by its type $1^1 3^3 7^2$ (or $7^2 3^3 1^1$).

Note that a positive integer d, both as the sequence (d) and as the type d^1, represents the partition $(d, 0, 0, \dots)$.

Third, a partition λ may be given by its *Ferrers diagram* D_λ. The diagram D_λ is the set of points with integer coordinates (i, j) such that $1 \le i$ and $1 \le j \le \lambda_i$. The diagram will always be pictured in a system of matrix coordinates where the first index i is a row index and the second j is a column index. Moreover, the point (i, j) will be pictured as the unit box with (i, j) as the lower right vertex.

For instance, the diagram of the partition in (9.1.1) is the following:



(9.2) Definition. There are several natural orders among partitions. First, we write $\lambda < \mu$ if $\lambda \ne \mu$ and if either $\|\lambda\| < \|\mu\|$ or if $\|\lambda\| = \|\mu\|$ and, for the first i for which $\lambda_i \ne \mu_i$ we have that $\lambda_i < \mu_i$. The relation is a total order on the set of partitions.

Next, we write $\lambda \prec \mu$ if $\lambda_i \le \mu_i$ for all i with strict inequality for at least one i. The relation is a partial order. With respect to the Ferrers diagrams, we have that $\lambda \preceq \mu$ if and only $D_\lambda \subseteq D_\mu$. If $\lambda \preceq \mu$, then the difference set $D_\mu - D_\lambda$ (of boxes) is called the *skew diagram* of μ/λ.

Finally, we write $\lambda \ll \mu$ if $\|\lambda\| = \|\mu\|$ and, for $q = 1, 2, \dots,$

$$\lambda_1 + \cdots + \lambda_q \le \mu_1 + \cdots + \mu_q.$$

Clearly, if $\lambda \ll \mu$, then $\lambda \le \mu$. The relation "$\ll$" is a reflexive (partial) order on the set of partitions; it is called the *dominance order* or the *majorization order*.

(9.3) Definition. For any partition λ, define the *conjugate partition* λ' by

$$\lambda'_p := \#\{i \mid \lambda_i \ge p\}.$$

For instance, for the partition λ of (9.1.1), the conjugate is the partition $\lambda' = (6, 5, 5, 2, 2, 2, 2) = 2^4 5^2 6^1$.

Note that the Ferrers diagram of λ' is obtained by reflecting the diagram of λ in the diagonal $i = j$, that is, the diagram $D_{\lambda'}$ is the transpose, D_λ^{tr}, of D_λ.

Clearly, we have that $\lambda'' = \lambda$. The biggest part of λ' is the length of λ, and the length of λ' is the biggest part of λ.

(9.4). There are several natural ways to associate partitions with multi indices.

(1) To a decreasing multi index K of size r, associate a partition according to the map,

$$(k_1, \dots, k_r) \mapsto (k_1, \dots, k_r, 0, 0, \dots). \tag{1}$$

Clearly, the map (1) defines a bijective correspondence between decreasing multi indices of size r and partitions of length at most r. Moreover, two decreasing multi indices of different size define the same partition if and only if the longer is obtained from the shorter by adding a trailing sequence of zeros.

Obviously, if λ and κ are the partitions associated via the map (1) to decreasing multi indices L and K of the same size, then $\lambda < \kappa$ if and only if $L < K$ and $\lambda \prec \kappa$ if and only if $L \prec K$.

(2) To any multi index I of size r, associate a partition according to the map,

$$(i_1, \dots, i_r) \mapsto (i_1 + \cdots + i_r, \dots, i_{r-1} + i_r, i_r, 0, 0, \dots). \tag{2}$$

Clearly, the map (2) defines a bijective correspondence between multi indices of size r and partitions of length at most r. Moreover, two multi indices of different size define the same partition if and only if the longer is obtained from the shorter by adding a trailing sequence of zeros.

(2') To any multi index I of size r, associate a partition according to the map,

$$(i_1, \dots, i_r) \mapsto (\overbrace{r, \dots, r}^{i_r}, \dots, \overbrace{2, \dots, 2}^{i_2}, \overbrace{1, \dots, 1}^{i_1}, 0, \dots). \tag{2'}$$

In other words, the associated partition is given by the type $1^{i_1} \cdots r^{i_r}$. Clearly, the map (2') defines a bijective correspondence between multi indices of size r and partitions of biggest part at most r. Moreover, two multi indices of different size define the same partition if and only if the longer is obtained from the shorter by adding a trailing sequence of zeros.

Clearly, if I is a multi index, then the two partitions associated with I via the maps (2) and (2') are conjugate.

(3) To a strictly increasing multi index J of size r, associate a partition according to the map,

$$(j_1, \dots, j_r) \mapsto (j_r - (r-1), \dots, j_2 - 1, j_1, 0, 0, \dots). \tag{3}$$

Clearly, the map (3) defines a bijective correspondence between strictly increasing multi indices of size r and partitions of length at most r. Moreover, two strictly increasing multi indices of different size define the same partition if and only if the longer is an extension of the shorter.

(9.5) Lemma. *If λ is the partition associated to a strictly increasing multi index J of size r via the map* (3), *then λ' is associated to the conjugate strictly increasing multi index J'.*

Proof. Indeed, assume that J is a subsequence of $[r + t]$. Then the complement $\tilde{J}$ and the conjugate J' are strictly increasing subsequences of $[r + t]$ of size t. Clearly, the conjugate of the partition λ is given by the equation,

$$\lambda'_p = \#\{q \in [1, r] \mid j_q - (q-1) \geq p\}.$$

Since $j_q - q \leq j_r - r \leq r + t - 1 - r = t - 1$, it follows that $\lambda'_p = 0$ for $p > t$. Fix p with $1 \leq p \leq t$. Clearly, for $1 \leq q \leq r$ and $q - 1 \leq i \leq t + q - 1$ the following relations are equivalent:

$$j_q \geq i, \quad \#J \cap [i] \leq q - 1, \quad \#\tilde{J} \cap [i] \geq i - (q-1), \quad \tilde{j}_{i-(q-1)} < i.$$

It follows in particular, with $i := p+q-1$, that $j_q - (q-1) \geq p$ if and only if $\tilde{j}_p < p+q-1$. Hence λ'_p is the number of integers q such that $\tilde{j}_p - (p-1) < q \leq r$. Here $\tilde{j}_p - (p-1) \leq r$, since $\tilde{j}_p - p \leq \tilde{j}_t - t \leq r + t - 1 - t = r - 1$. Thus λ'_p is the difference,

$$\lambda'_p = r - \big(\tilde{j}_p - (p-1)\big). \tag{9.5.1}$$

Clearly, the right hand side of (9.5.1) is the partition associated to $(j'_1, \dots, j'_t)$. ∎

(9.6) Definition. For any partition λ of length at most n, define

$$m_\lambda := m^K,$$

where K is the decreasing multi index of size n corresponding to λ via the map (1). In other words,

$$m_\lambda := m^{\lambda_1, \dots, \lambda_n}.$$

Define $m_\lambda := 0$ if the length of λ is bigger than n. Note that $m_d = p_d$ is the d'th power sum, and $m_{1\dots 1} = e_d$, or, in the type notation, $m_{1^d} = e_d$.

For any partition λ, define

$$e_\lambda := e_K, \quad h_\lambda := h_K$$

where K is a decreasing multi index corresponding to λ via the map (1). In other words,

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

Note that the products are finite since $e_0 = h_0 = 1$. The products are defined for arbitrary partitions λ. The product e_λ vanishes if the biggest part of λ is strictly greater than n, since $e_d = 0$ when $d > n$. Clearly, if I is any multi index and λ is associated to I via the map (2'), then $e^I = e_\lambda$.

For any partition λ of length at most n, define

$$\Delta_\lambda := \Delta^J, \qquad s_\lambda := s^J,$$

where J is the strictly increasing multi index of size n corresponding to λ via the map (3). In other words,

$$\Delta_\lambda = \Delta^{\lambda_n, \lambda_{n-1}+1, \dots, \lambda_1+(n-1)}, \qquad s_\lambda = s^{\lambda_n, \lambda_{n-1}+1, \dots, \lambda_1+(n-1)}.$$

Note that the empty partition ( ) corresponds to the sequence $(0, 1, \dots, n-1)$. Hence $\Delta_{()}$ is the Vandermonde determinant, and $s_{()} = s^{0,1,\dots,n-1} = 1$. Under the correspondence, the partition d corresponds to the sequence $(0, 1, \dots, n-2, n-1+d)$. Hence it follows that, with d as partition, we have that s_d is the d'th complete symmetric polynomial.

Consider finally any two partitions λ and μ of length at most n, and strictly increasing multi indices I and J of the same size associated to λ and μ via the map (3). To any power series $u \in R[A][[T]]$ we associated in (7.4) a determinant u^I_J. Indexed by partitions the determinant is sometimes denoted by $u_{\lambda/\mu}$. For instance, the skew Schur polynomial s^I_J is also denoted $s_{\lambda/\mu}$.

(9.7). In the language of partitions, the notion of a tableau is the following: Denote by $J \mapsto \bar{J}$ the bijection (3) from strictly increasing multi indices of size r to partitions of length at most r. Under this correspondence, we have that $I \succeq J$ if and only if $D_{\bar{I}} \supseteq D_{\bar{J}}$. Moreover, if $I \succeq J$, then I/J is a horizontal strip as defined in (7.11) if and only if the skew diagram $D_{\bar{I}} - D_{\bar{J}}$ has no more than one box in each column.

Let $I \succeq J$ be strictly increasing multi indices of size r corresponding, via the map (3), to partitions λ and μ. Then a sequence $I = K_0 \succeq K_1 \succeq \cdots \succeq K_n = J$ of strictly increasing multi indices K_r corresponds to a sequence of partitions, $\lambda = \kappa_0 \succeq \kappa_1 \succeq \cdots \succeq \kappa_n = \mu$. The sequence of partitions may be represented by the skew diagrams D_q of κ_{q-1}/κ_q for $q = 1, \dots, n$. The skew diagram D_q is a subdiagram of the skew diagram D of λ/μ, and it may be visualized by inserting the number q in all boxes of D_q. Note that the inserted numbers increase along the rows and along the columns of D. Moreover, the condition that each κ_{q-1}/κ_q is a horizontal strip means that no number q can occur more than once in each column. In other words, the condition means that the inserted numbers are strictly increasing in each column of D. It follows that a tableau of shape I/J and biggest entry n, as defined in (7.11), can be identified with an insertion of numbers from $\{1,\dots,n\}$ in the boxes of the diagram of λ/μ such that the inserted numbers increase weakly in each row and increase strictly in each column.

Under the identifications, the degree $\|K_{q-1}-K_q\|$ is equal to the degree $\|\kappa_{q-1}-\kappa_q\|$, and hence equal to the number of boxes in the skew diagram D_q. Hence the degree $\|K_{q-1}-K_q\|$ is equal to the number, denoted $k_q(T)$, of times the number q occurs in the tableau T. The following formula is therefore a rewriting of (7.11.6):

$$s_{\lambda/\mu} = \sum_T a_1^{k_1(T)} \cdots a_n^{k_n(T)},$$

where the sum is over all tableaux T of shape λ/μ and biggest entry n. In particular, if K is a given decreasing multi index of size n, then the coefficient to m^K in the expansion of $s_{\lambda/\mu}$ in the basis of monomial symmetric polynomials is equal to the number of tableaux T of shape λ/μ for which

$$K = (k_1(T),\dots,k_n(T)).$$

10. Applications of Determinantal Methods.

(10.1). Let n be a nonnegative integer. Let $(\alpha_1, \dots, \alpha_n)$ be a sequence of n elements of the ground ring. Let notations be as in (7.7). Set $S_{[n]}(\alpha_1, \dots, \alpha_n) = M_{[n]}(s)$, where $s = s(\alpha_1, \dots, \alpha_n)$ as in (7.7). In the proof of Jacobi–Trudi's formula (7.8)(5), we proved the formula (7.8.8):

$$S_{[n]}(\alpha_1, \dots, \alpha_n)D = V(\alpha_1, \dots, \alpha_n), \tag{10.1.1}$$

where $D = D(\alpha_1, \dots, \alpha_n)$ is an $n \times n$ matrix with determinant equal to the Vandermonde determinant $\Delta = \Delta(\alpha_1, \dots, \alpha_n)$. The formula has far reaching consequences.

Consider first a power series u which is a polynomial. Then, for every element α in the ground ring, the value $u(\alpha)$ is a well defined. Recall that $V(\alpha) = \langle (1-\alpha T)^{-1} \rangle$ is the infinite column with entries $1, \alpha, \alpha^2, \dots$. Hence, if the column $V(\alpha)$ is multiplied from the left by the infinite row whose entries are the coefficients of $T^i u$, the result is $\alpha^i u(\alpha)$. In other words, we have the following matrix equation,

$$M(u)^{\mathrm{tr}} V(\alpha) = V(\alpha) u(\alpha).$$

As a consequence, by multiplying (10.1.1) from the left by the matrix $M(u)^{\mathrm{tr}}$ we obtain the equation,

$$M(u)^{\mathrm{tr}} S_{[n]}(\alpha_1, \dots, \alpha_n)D = V(\alpha_1, \dots, \alpha_n)\mathrm{diag}(u(\alpha_1), \dots, u(\alpha_n)). \tag{10.1.2}$$

Let J be a strictly increasing multi index of size n, and extract from (10.1.2) the equations corresponding to the rows in J. The result is the equation,

$$M_J(u)^{\mathrm{tr}} S_{[n]}(\alpha_1, \dots, \alpha_n)D = V^J(\alpha_1, \dots, \alpha_n)\mathrm{diag}(u(\alpha_1), \dots, u(\alpha_n)). \tag{10.1.3}$$

Finally, take determinants in the equation (10.1.3). The determinant of D is the Vandermonde determinant Δ. The determinant of the product $M_J(u)^{\mathrm{tr}} S_{[n]}$ is developed as in the proof of the multiplication formula (7.5)(6). The result is the equation,

$$\Delta \sum_{I \succeq J} u^I_J s^I_{[n]}(\alpha_1, \dots, \alpha_n) = \Delta^J(\alpha_1, \dots, \alpha_n) \prod_i u(\alpha_i). \tag{10.1.4}$$

From (10.1.4) it follows, since $\Delta^J = \Delta s^J$ by definition of the Schur polynomials, that

$$\sum_{I \succeq J} u^I_J s^I(\alpha_1, \dots, \alpha_n) = s^J(\alpha_1, \dots, \alpha_n) \prod_i u(\alpha_i). \tag{10.1.5}$$

Indeed, the equation follows from (10.1.4) when the ground ring is the polynomial ring $R[A]$ and $\alpha_i := a_i$, because then the Vandermonde determinant is a regular element. Therefore, since (10.1.5) is of universal nature, it holds for general ground rings.

Equation (10.1.5) holds when u is a polynomial. For some ground rings, the equation holds even when u is a power series. Assume for simplicity that the ground ring is the power series ring $R[[A]]$ over the alphabet A. If α is an element of $R[[A]]$ of positive order (that is, without constant term), then evaluation $u(\alpha)$ is the element of $R[[A]]$ defined formally by

$$u(\alpha) := u_0 + u_1\alpha + u_2\alpha^2 + \dots,$$

In fact, in the sum only finitely many terms have order less than a given number, and so the infinite sum is a well defined element of $R[[A]]$.

In this setup, if $\alpha_1, \dots, \alpha_n$ are power series in $R[[A]]$ without constant terms, then (10.1.5) holds. Indeed, (10.1.5) is an equation of power series in $R[[A]]$, and so it suffices to show that its two sides agree in every degree. However, in a fixed degree, the equation involves only finitely many terms of u. Hence, since the equation holds for polynomials u, it holds for arbitrary power series.

As a standard application, replace R by a power series ring $R[[B]]$ over a second alphabet B. Take $u := s(B)$ and $\alpha_i := a_i$. Then formula (10.1.5), for a strictly increasing multi index J of size n equal to the number of letters of A, is the following equation in $R[[A, B]]$, identical to (8.3.2):

$$\sum_{I \succeq J} s^I_J(B) s^I(A) = s^J(A) \prod_{a \in A,\, b \in B} \frac{1}{1 - ab}, \tag{10.1.6}$$

where the sum is over strictly increasing multi indices I of size n. In particular, with $J = [n]$, we obtain the formula,

$$\sum_I s^I(B) s^I(A) = \prod_{a \in A,\, b \in B} \frac{1}{1 - ab}. \tag{10.1.7}$$

Difference operators

1. Operators on rational functions.

(1.1) Setup. Fix a commutative ring R and a finite alphabet $A = \{a_1, \dots, a_n\}$ of $n \geq 0$ letters. Denote by $R(A)$ the ring of *rational functions* in the letters of A, that is, the ring of all fractions f/g where f and g are polynomials in $R[A]$ and g is a regular element in $R[A]$.

Clearly, the action of the symmetric group $\mathfrak{S}(A)$ on $R[A]$ extends to an action on $R(A)$. A rational function f in $R(A)$ is called *symmetric* if it is invariant under the action of $\mathfrak{S}(A)$, that is, if

$$\sigma(f) = f \text{ for all } \sigma \in \mathfrak{S}(A). \tag{1.1.1}$$

The symmetric rational functions in $R(A)$ form an R-subalgebra, denoted $\mathrm{Sym}_R(A)$.

A rational function f is called *anti-symmetric* if it is semi-invariant under the action of $\mathfrak{S}(A)$ in the sense that

$$\sigma(f) = \mathrm{sign}(\sigma) f \text{ for all } \sigma \in \mathfrak{S}(A). \tag{1.1.2}$$

Clearly, the anti-symmetric rational functions in $R(A)$ form a module over the ring $\mathrm{Sym}_R(A)$ of symmetric rational functions. In particular, a product of a symmetric rational function and an anti-symmetric rational function is an anti-symmetric rational function. Similarly, a product of two anti-symmetric rational functions is a symmetric rational function.

It follows that if h is an anti-symmetric polynomial and regular in $R[A]$ (for instance, h could be the Vandermonde determinant $\Delta = \prod_{a<b}(b - a)$), then multiplication by h defines an isomorphism from the R-module $\mathrm{Sym}_R(A)$ of symmetric rational functions onto the R-module of anti-symmetric rational functions.

(1.2) Remark. If g is any polynomial in $R[A]$, then by multiplying g by the product of the polynomials $\sigma(g)$ for all permutations $\sigma \neq 1$, we obtain a symmetric polynomial. It follows easily that any rational function in $R(A)$ can be written as a fraction f/h where h is a symmetric polynomial, and regular in $R[A]$. A symmetric polynomial regular in $R[A]$ is of course also regular in $\mathrm{Sym}_R[A]$. Conversely, a symmetric polynomial regular in $\mathrm{Sym}_R[A]$ is also regular in $R[A]$, since, by the following Proposition (1.3), $R[A]$ is a free $\mathrm{Sym}_R[A]$-module.

It follows that the ring of symmetric functions $\mathrm{Sym}_R(A)$ is the total fraction ring of its subring $\mathrm{Sym}_R[A]$, that is, the symmetric functions are the fractions of the form f/g where f and g are symmetric polynomials and g is regular in $\mathrm{Sym}_R[A]$.

(1.3) Proposition. *Let E be the sequence $(0, 1, \dots, n-1)$. Then the monomials a^J for $J \preceq E$ form a basis both for the algebra $R(A)$ of rational functions as a module over its subring $\mathrm{Sym}_R(A)$ of symmetric functions, and for the algebra $R[A]$ of polynomials as a module over its subring $\mathrm{Sym}_R[A]$ of symmetric polynomials.*

Proof. As noted in (1.2), every rational function in $R(A)$ is a fraction f/g where g is a symmetric polynomial. As a consequence, the first assertion of the Lemma follows from the second.

The second assertion will be proved by induction on the number n of letters of A. Clearly, the assertion holds when $n = 1$. Assume that $n > 1$ and consider the alphabet $\bar{A} := \{a_1, \dots, a_{n-1}\}$. Then $R[A] = R[a_n][\bar{A}]$. The subring $\mathrm{Sym}_{R[a_n]}[\bar{A}]$ consists of the polynomials of $R[A]$ that are symmetric in the first $n-1$ letters. By induction, $R[A]$ is a free module over its subring $\mathrm{Sym}_{R[a_n]}[\bar{A}]$, with a basis formed by the monomials $\bar{a}^{\bar{J}}$ for $\bar{J} \preceq \bar{E}$, where $\bar{E} = (0, 1, \dots, n-2)$. Therefore, it suffices to prove that the ring $\mathrm{Sym}_{R[a_n]}[\bar{A}]$ is a free module over its subring $\mathrm{Sym}_R[A]$, with a basis formed by the powers $1, a_n, \dots, a_n^{n-1}$.

In other words, it suffices to prove that any polynomial p in $\mathrm{Sym}_{R[a_n]}[\bar{A}]$ has a unique expansion,

$$p = q_0 + q_1 a_n + \cdots + q_{n-1} a_n^{n-1}, \tag{1}$$

where $q_i \in \mathrm{Sym}_R[A]$.

Assume that p is a polynomial in $\mathrm{Sym}_{R[a_n]}[\bar{A}]$. For $i = 1, \dots, n$, let $p_i := \mu(p)$ where μ is any permutation in $\mathfrak{S}(A)$ such that $\mu(a_n) = a_i$. The polynomial p_i is independent of the choice of μ, because p is symmetric in the letters $a_1, \dots, a_{n-1}$. Consider the following system of n equations,

$$p_i = q_0 + q_1 a_i + \cdots + q_{n-1} a_i^{n-1} \quad \text{for } i = 1, \dots, n, \tag{2}$$

with unknown functions q_i in $R(A)$. The determinant of the matrix of coefficients is the Vandermonde determinant Δ. Hence the system (2) has a unique solution $(q_0, \dots, q_{n-1})$ with q_i in $R(A)$.

The n'th equation in (2) is the equation (1), because $p_n = p$. Clearly, if the equation (1) holds with polynomials q_i that are symmetric in the letters of A, then the equations (2) hold. Hence it suffices to prove for the solutions q_i to the system of equations (2) that each q_i is a polynomial and symmetric in the letters of A. By Cramer's rule, the solution q_i is the fraction Q_i/Δ, where the denominator is the Vandermonde determinant and the numerator is the determinant,

$$Q_i := \begin{vmatrix} 1 & \dots & a_1^{i-1} & p_1 & a_1^{i+1} & \dots & a_1^{n-1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 1 & \dots & a_n^{i-1} & p_n & a_n^{i+1} & \dots & a_n^{n-1} \end{vmatrix}.$$

Let $\tau := \varepsilon_{a_j, a_k}$ be the transposition that interchanges two different letters a_j and a_k of A. It follows from the definition of the polynomials p_i that $\tau(p_i) = p_i$ when i is different from k and j and $\tau(p_k) = p_j$. Hence, when τ is applied to the determinant Q_i, the k'th and the

j'th row of the determinant are interchanged, and if we substitute $a_k = a_j$ in the determinant then its j'th and k'th row become equal. Therefore, the determinant Q_i is an alternating polynomial. As a consequence, the quotient $q_i = Q_i/\Delta$ is a polynomial and symmetric in the letters of A.

Thus the Proposition has been proved. ∎

(1.4) Corollary. *Let $\varphi\colon R[A] \to R[A]$ be a $\mathrm{Sym}_R[A]$-linear map. Then for every polynomial $f \in R[A]$ we have the inequality for the orders of f and $\varphi(f)$,*

$$\operatorname{ord}\varphi(f) \geq \operatorname{ord} f - n(n-1)/2. \tag{1.4.1}$$

Proof. Clearly, we may assume that f is homogeneous of, say, degree d. By the Proposition, f is a sum of terms $q_J a^J$ where $J \preceq E$ and with uniquely determined coefficients $q_J \in \mathrm{Sym}_R[A]$. As $\mathrm{Sym}_R[A] \subseteq R[A]$ is a homogeneous subring, the coefficient q_J is homogeneous of degree $d - \|J\|$. Hence $\varphi(f)$ is the sum of the terms $q_J\varphi(a^J)$, and the order of $q_J\varphi(a^J)$ is at least $d - \|J\|$. So all the terms have order at least $d - \|E\| = d - n(n-1)/2$. Clearly, the inequality (1.4.1) is a consequence. ∎

(1.5) Note. The relevance of the previous result, and its proof, was pointed out by Darij Grinberg. The inequality in (1.4.1) is best possible since for instance the symmetrization operator δ^A of (2.3) is a homogeneous $\mathrm{Sym}_R[A]$-linear map $R[A] \to R[A]$ lowering the degree by $n(n-1)/2$. It follows also that any homogeneous $\mathrm{Sym}_R[A]$-basis for $R[A]$ contains a basis element of degree at least $n(n-1)/2$.

2. The simple difference operators.

(2.1) Definition. Consider the *(twisted) group algebra* $R(A)[\mathfrak{S}(A)]$ of the ring of rational functions $R(A)$. As a left-$R(A)$-module, the group algebra is freely generated by the permutations in $\mathfrak{S}(A)$, that is, the elements of the algebra are $R(A)$-linear combinations of permutations,

$$\alpha = \sum_\sigma f_\sigma \sigma, \tag{2.1.1}$$

and the multiplication in the algebra is given by the rule, for $f \in R(A)$, $\sigma \in \mathfrak{S}(A)$,

$$\sigma \cdot f = \sigma(f)\sigma\,. \tag{2.1.2}$$

The twisted group algebra $R(A)[\mathfrak{S}(A)]$ contains the ring of rational functions $R(A)$, and it contains the group $\mathfrak{S}(A)$. Note that σf, for $\sigma \in \mathfrak{S}(A)$ and $f \in R(A)$, can be interpreted both as the product of σ and f in the group algebra and as the function obtained from f be the action of σ. When the interpretation is not clear from the context, we write $\sigma \cdot f$ for the product in the group algebra (defined by (2.1.2)) and $\sigma(f)$ for the function obtained from the action.

A permutation μ in $\mathfrak{S}(A)$ is invertible in the group algebra. Hence it induces an inner automorphism $\alpha \mapsto \mu\alpha\mu^{-1}$ of the group algebra, called *conjugation* by μ. On the subring of rational functions, conjugation is the map $f \mapsto \mu(f)$.

In addition, the group algebra has a *canonical involution* $\alpha \mapsto \alpha^*$. It is the anti-automorphism of the group algebra defined by

$$(f\sigma)^* := \operatorname{sign}(\sigma)\sigma^{-1}\cdot f = \operatorname{sign}(\sigma)\sigma^{-1}(f)\sigma^{-1}. \tag{2.1.3}$$

Note that the involution is an anti-automorphism, that is, it reverses the order of the factors in a product. It is equal to the identity on the subring $R(A)$ of rational functions. In particular, $(\sigma f)^* = f\sigma^* = \operatorname{sign}(\sigma) f\sigma^{-1}$.

(2.2) Definition. In these notes, actions are assumed to be left actions, and modules are assumed to be left modules unless the contrary is explicitly stated. The group algebra $R(A)[\mathfrak{S}(A)]$ acts naturally on the R-module $R(A)$. More precisely, to the element $\alpha \in R(A)[\mathfrak{S}(A)]$ with the expansion (2.1.1) we associate the operator on $R(A)$ defined by

$$\alpha(g) := \sum_\sigma f_\sigma \sigma(g), \qquad g \in R(A). \tag{2.2.1}$$

The action is faithful, that is, if an element α of the group algebra operates as the zero map on $R(A)$, then $\alpha = 0$. Equivalently, if a sum $\alpha = \sum_{i=1}^k f_i\sigma_i$, where the σ_i are k different permutations, defines the zero operator (2.2.1), then the functions f_i are equal to zero. Indeed, by a standard argument of Galois theory, the assertion is proved by induction on k. It holds when $k = 1$. Assume that $k > 1$ and that α defines the zero operator. Then, for every function $f \in R(A)$, we have that $\sigma_k(f)\alpha$ and $\alpha \cdot f$ defines the zero operator. Hence the difference $\sigma_k(f)\alpha - \alpha{\cdot}f$ defines the zero operator. Clearly, the difference is the sum $\sum_{i=1}^{k-1}[\sigma_k(f) - \sigma_i(f)]f_i\sigma_i$. Hence, by induction, we have for $i = 1, \dots, k-1$ the equation $\big(\sigma_k(f) - \sigma_i(f)\big)f_i = 0$. The equation, for all $f \in R(A)$, implies that $f_i = 0$, since we may take as f a letter $a \in A$ such that $\sigma_k(a)$ and $\sigma_i(a)$ are two different letters of A. Hence $f_i = 0$ for $i = 1, \dots, k-1$. Clearly, then also $f_k = 0$.

The algebra of all operators on $R(A)$ of the form $g \mapsto \alpha(g)$ where α is an element of the twisted group algebra will be denoted $\mathcal{E}_R(A)$. The elements of the algebra $\mathcal{E}_R(A)$ will simply be called *operators* on $R(A)$. By the result just proved, operators can be identified with elements of the twisted group algebra. Note that all operators are $\mathrm{Sym}_R(A)$-linear. The subalgebra of $\mathcal{E}_R(A)$ consisting of operators that map the subring $R[A]$ into itself will be denoted $\mathcal{E}_R[A]$. Obviously, the algebra $\mathcal{E}_R[A]$ contains the ring of polynomials $R[A]$ and the group of permutations $\mathfrak{S}(A)$, that is, the algebra contains the twisted group algebra $R[A][\mathfrak{S}(A)]$.

It will be proved in Chapter SCHUB that the algebra of operators $\mathcal{E}_R(A)$ is the full ring of all $\mathrm{Sym}_R(A)$-linear endomorphisms of $R(A)$.

Clearly, the subalgebra $\mathcal{E}_R[A]$ is invariant under the conjugations. It will be proved in Chapter SCHUB that $\mathcal{E}_R[A]$ is invariant under the canonical involution.

(2.3) Definition. The *symmetrization operator* $\delta = \delta^A$ of $R(A)$ is the following operator:

$$\delta^A := \sum_{\sigma\in\mathfrak{S}(A)} \sigma{\cdot}\frac{1}{\Delta} = \frac{1}{\Delta}\sum_{\sigma\in\mathfrak{S}(A)} (\operatorname{sign}\sigma)\sigma.$$

The two expressions are equal, because $\sigma(\Delta) = (-1)^{\ell(\sigma)}\Delta$. It is clear from the first expression that the values of the operator δ^A are symmetric functions. When the sum in the second expression is applied to a polynomial, the result is an alternating polynomial and hence divisible by Δ. Therefore the operator δ^A belongs to $\mathcal{E}_R[A]$.

Clearly, the first sum is transformed into the second by the canonical involution. Hence the operator δ^A is invariant under the canonical involution. Moreover, for any permutation $\mu \in \mathfrak{S}(A)$ we have $\mu\delta^A = \delta^A$ and $\delta^A\mu = (\operatorname{sign}\mu)\,\delta^A$; in particular, under conjugation by μ we have $\mu\delta^A\mu^{-1} = (\operatorname{sign}\mu)\,\delta^A$.

(2.4) Definition. For $1 \le p < n$, let ε_p be the simple transposition that interchanges a_p and a_{p+1}. Define the *simple difference operator*,

$$\partial^p := \frac{1}{a_{p+1} - a_p}(1 - \varepsilon_p). \tag{2.4.1}$$

Clearly, for each polynomial f in $R[A]$ we have that the difference $f - \varepsilon_p(f)$ vanishes when we substitute $a_{p+1} = a_p$. Hence the difference is divisible by $a_{p+1} - a_p$. Therefore the operator ∂^p belongs to the subring $\mathcal{E}_R[A]$. Moreover, the operator ∂^p is invariant under the canonical involution:

$$(\partial^p)^* = \partial^p. \tag{2.4.2}$$

Indeed, if we let $\Delta_p := a_{p+1} - a_p$, then $\partial^p = 1/\Delta_p - (1/\Delta_p)\varepsilon_p$. Under the canonical involution, the function $1/\Delta_p$ is invariant and ε_p is changed into $-\varepsilon_p$. Hence $(1/\Delta_p)\varepsilon_p$ is changed to $-\varepsilon_p{\cdot}(1/\Delta_p) = (1/\Delta_p)\varepsilon_p$, and consequently the equation (2.4.2) holds.

It will be convenient to define $\partial_p := -\partial^p$. In addition, define operators π^p and ψ^p, and π_p and ψ_p, by the equations,

$$\pi^p = \partial^p{\cdot}a_{p+1}, \quad \psi^p = a_p\partial^p, \quad \pi_p = \partial_p{\cdot}a_p, \quad \psi_p = a_{p+1}\partial_p. \tag{2.4.3}$$

Clearly, under conjugation by ε_p the function $1/\Delta_p$ changes sign and the operator $1 - \varepsilon_p$ is invariant. Therefore, under conjugation by ε_p the three operators ∂^p, π^p and ψ^p are changed into ∂_p, π_p and ψ_p. Similarly, under conjugation by ω the functions a_p and a_{p+1} are mapped to a_{n-p+1} and a_{n-p}, and ε_p is mapped to ε_{n-p}. Hence we obtain the formulas,

$$\omega\partial^p\omega = \partial_{n-p}, \quad \omega\pi^p\omega = \pi_{n-p}, \quad \omega\psi^p\omega = \psi_{n-p}. \tag{2.4.4}$$

Since the canonical involution reverses the order of the factors, we obtain from (2.4.2) and (2.4.3) the following formulas:

$$(\partial^p)^* = -\partial_p, \quad (\pi^p)^* = -\psi_p, \quad (\psi^p)^* = -\pi_p. \tag{2.4.5}$$

Finally, by combining the previous two sets of formulas we obtain the following:

$$\omega(\partial^p)^*\omega = -\partial^{n-p}, \quad \omega(\pi^p)^*\omega = -\psi^{n-p}, \quad \omega(\psi^p)^*\omega = -\pi^{n-p}. \tag{2.4.6}$$

(2.5) Observation. It is immediate from the definition that $\partial^p(a_{p+1}) = 1$ and $\partial^p(a_p) = -1$. As a consequence, $\pi^p(1) = 1$. Note that the operator ∂^p is of degree -1 and π^p and ψ^p are of degree 0 in the variables of A.

(2.6) Lemma. *The operators of* (2.4) *are linear with respect to polynomials that are symmetric in the variables a_p and a_{p+1}. Moreover, the image of ∂^p is symmetric in these variables, and ∂^p vanishes on polynomials that are symmetric in these variables. Finally, the image of π^p is symmetric in the variables a_p and a_{p+1} and $\pi^p(1) = 1$.*

Proof. All assertions result directly from the definition. ∎

(2.7) The Leibnitz Formula. *The operator ∂^p is a ε_p-derivation, that is, for rational functions f and g in $R(A)$ we have that*

$$\partial^p(gf) = \partial^p(g)f + \varepsilon_p(g)\partial^p(f).$$

Proof. The assertion follows by a direct calculation. ∎

(2.8) Lemma. *The following equations hold:*

$$1 = \pi^p - \psi^p, \quad \partial^p\partial^p = 0, \quad \pi^p\pi^p = \pi^p, \quad \psi^p\psi^p = -\psi^p.$$

Proof. It follows from the Leibnitz formula that $\partial^p(a_{p+1}f) = f + a_p\partial^p(f)$. Hence the first equation of the Lemma holds. The second equation follows from the second assertion of Lemma (2.6). The third equation follows from the third assertion of Lemma (2.6). Finally, the last equation is a consequence of the first and the third. ∎

(2.9) Lemma. *Let* $E := (0, 1, 2, \dots, n-1)$ *and* $E_1 := (0, 0, 1, \dots, n-2)$. *Then the following equation of operators holds:*

$$a^{E_1}\pi^1 \cdots \pi^{n-1} = \partial^1 \cdots \partial^{n-1} \cdot a^E.$$

Proof. The monomial a^{E_1} is equal to $a^E/(a_2 \cdots a_n)$. Define more generally, for $p = 1, \dots, n$, the monomial $a^{E_p} := a^E/(a_{p+1} \cdots a_n)$. Then the following equation holds for $p < n$:

$$a^{E_p}\pi^p = \partial^p \cdot a^{E_{p+1}}. \tag{1}$$

Indeed, the left side of the equation is equal to $\pi^p \cdot a^{E_p}$ because a^{E_p} is symmetric in the variables a_p and a_{p+1}, and $\pi^p a^{E_p}$ is equal to the right side by definition of π^p and a^{E_p}.

Clearly, the equations (1) for $p < n$ imply that the following product of operators, for $p = 1, \dots, n$, is independent of p:

$$\partial^1 \cdots \partial^{p-1} a^{E_p} \pi^p \cdots \pi^{n-1}.$$

Finally, the asserted equation of the Lemma is the equality of the latter products for $p = 1$ and $p = n$. □

3. General difference operators.

(3.1) Definition. Define, by induction on the number of letters n of the alphabet A, two operators $\partial^\omega = \partial^{\omega_A}$ and $\pi^\omega = \pi^{\omega_A}$ as follows. If $n = 1$, then both operators are equal to 1. If $n > 1$, define

$$\partial^{\omega_A} := \partial^{\omega_{A_1}} \partial^1 \cdots \partial^{n-1} \text{ and } \pi^{\omega_A} := \pi^{\omega_{A_1}} \pi^1 \cdots \pi^{n-1},$$

where A_1 is the alphabet with the $n-1$ letters $a_2, \dots, a_n$. In addition, define the operator $\psi^\omega = \psi^{\omega_A}$ by the equation,

$$\psi^\omega := (-1)^{n(n-1)/2} \omega (\pi^\omega)^* \omega.$$

The appearence of the maximal permutation $\omega = \omega_A$ as an exponent in the operators ∂^ω, π^ω and ψ^ω will be explained in (3.5) where we give a more flexible definition of all three operators; we show in particular that the operator ψ^ω satisfies the equation $\psi^{\omega_A} = \psi^{\omega_{A_1}} \psi^1 \cdots \psi^{n-1}$, analogous to the formulas used for the inductive definition of ∂^ω and π^ω.

(3.2) Example. If $n = 3$, then

$$\partial^\omega = \partial^2 \partial^1 \partial^2, \quad \pi^\omega = \pi^2 \pi^1 \pi^2, \quad \psi^\omega = \psi^1 \psi^2 \psi^1.$$

Indeed, the first two equations follow from the inductive definition in (3.1). As $\omega(\pi^p)^*\omega = -\psi^{n-p}$, see (DIFF.2.4.6), the third asserted equation follows by applying the involution $\alpha \mapsto \omega\alpha^*\omega$ to the second equation.

(3.3) Theorem. *Consider the symmetrization operator δ^A of* (DIFF.2.3),

$$\delta^A = \sum_{\sigma \in \mathfrak{S}(A)} \sigma \cdot \frac{1}{\Delta} = \frac{1}{\Delta} \sum_{\sigma \in \mathfrak{S}(A)} (\operatorname{sign} \sigma) \sigma. \tag{3.3.1}$$

Then, for any monomial a^J we have that

$$\delta^A(a^J) = \Delta^J / \Delta\,. \tag{3.3.2}$$

In particular, if $E = (0, 1, \dots, n-1)$, then $\delta^A(a^E) = 1$ and $\delta^A(a^J) = 0$ when $J \prec E$. Finally, the following three operator equations hold:

$$\partial^\omega = \delta^A, \quad \pi^\omega = \partial^\omega \cdot a^E, \quad \psi^\omega = \omega(a^E) \partial^\omega. \tag{3.3.3}$$

Proof. The equality of the expressions in (3.3.1) was observed in the definition (DIFF.2.3) of δ^A. Clearly, when the second sum in (3.3.1) is applied to a monomial a^J the result is the determinant Δ^J, see (SYM.6.6). Hence the equation (3.3.2) follows from the second expression for δ^A.

By definition of the Vandermonde determinant we have that $\Delta^E = \Delta$. Moreover, if $J \prec E$, then J has two equal entries, and consequently $\Delta^J = 0$. Hence we have obtained for $\delta^A(a^J)$ the special values given in the Theorem.

Consider the three operator equations of (3.3.3). The second equation follows by induction on the number of letters of A from Lemma (DIFF.2.9) and the recursive definitions of ∂^ω and π^ω in (3.1). Indeed, let $A_1 := A - \{a_1\}$ be the alphabet of (3.1), and let E_1 be the sequence of (DIFF.2.9). Clearly, by induction, we may assume that $\pi^{\omega_{A_1}} = \partial^{\omega_{A_1}} a^{E_1}$. Hence, by the recursive definitions of ∂^ω and π^ω, it suffices to prove the following equation,

$$\partial^{\omega_{A_1}} a^{E_1} \pi^1 \cdots \pi^{n-1} = \partial^{\omega_{A_1}} \partial^1 \cdots \partial^{n-1} a^E.$$

The latter equation follows immediately from Lemma (DIFF.2.9).

Consider next the first equation of (3.3.3). The two sides are operators of $\mathcal{E}_R(A)$. In particular, they are $\mathrm{Sym}_R(A)$-linear operators. Therefore, by Proposition (DIFF.1.3), it suffices to prove that the two operators yield the same value when they are evaluated on a monomial a^J, where $J \preceq E$. The values $\delta^A(a^J)$ were found in the first part of the Theorem. The other operator ∂^ω is a product of $n(n-1)/2$ simple operators ∂^p and each operator ∂^p lowers the degree by 1. Therefore, if $J \prec E$, then $\partial^\omega(a^J) = 0$, because the degree of a^J is strictly less than $n(n-1)/2$. Moreover, for $J = E$ we have that $\partial^\omega(a^E) = \pi^\omega(1)$ by the second equation of (3.3.3), and $\pi^\omega(1) = 1$, because π^ω is a composition of simple operators π^p and $\pi^p(1) = 1$ for all p. Hence, the equality $\partial^\omega(a^J) = \delta^A(a^J)$ holds for $J \preceq E$. Consequently, the two operators are equal and the first equation of (3.3.3) has been proved.

It remains to prove the third equation of (3.3.3). The following equation holds:

$$\omega(\partial^\omega)^*\omega^{-1} = (-1)^{n(n-1)/2}\partial^\omega. \tag{3.3.4}$$

Indeed, it follows from the observations in the last two lines of Definition (2.3) that the equation holds for the operator δ^A and we have proved that $\partial^\omega = \delta^A$. Apply the involution $\alpha \mapsto (-1)^{n(n-1)/2}\omega\alpha^*\omega^{-1}$ to the second equation of (3.3.3). We obtain the equation,

$$(-1)^{n(n-1)/2}\omega(\pi^\omega)^*\omega^{-1} = (-1)^{n(n-1)/2}\omega(a^E)\omega(\partial^\omega)^*\omega^{-1}.$$

The left hand side is ψ^ω by definition and the right hand side is $\omega(a^E)\partial^\omega$ by (3.3.4). Therefore the third equation of (3.3.3) holds.

Thus all the assertions of the Theorem have been proved. ∎

(3.4) Corollary. *Each of the three sets of operators,*

$$\{\partial^1, \ldots, \partial^{n-1}\}, \qquad \{\pi^1, \ldots, \pi^{n-1}\}, \qquad \{\psi^1, \ldots, \psi^{n-1}\},$$

satisfies the Coxeter–Moore relations.

Proof. The first of the Coxeter–Moore relations, for any of the three sets, are satisfied because the operator ∂^p commutes with ∂^q and with a_q and a_{q+1} when $|p - q| > 1$.

Clearly, to verify the second Coxeter–Moore relation, it suffices to consider an alphabet A with 3 letters a_1, a_2, a_3. The equations, for the three sets of operators, are the following:

$$\partial^2\partial^1\partial^2 = \partial^1\partial^2\partial^1, \quad \pi^2\pi^1\pi^2 = \pi^1\pi^2\pi^1, \quad \psi^2\psi^1\psi^2 = \psi^1\psi^2\psi^1. \tag{1}$$

Consider the first equation. The left side is equal to ∂^ω. Hence, by (3.3.3), we obtain the equation $\partial^2\partial^1\partial^2 = \delta^A$. Now apply the involution $\alpha \mapsto \omega\alpha^*\omega$ to the latter equation. On the left, the result is $-\partial^1\partial^2\partial^1$ by the first equation of (DIFF.2.4.6), and on the right the result is $-\delta^A$ as observed in the last paragraph of (DIFF.2.3). Hence we obtain the equation $\partial^1\partial^2\partial^1 = \delta^A$. So we have proved that both sides of the first equation of (1) are equal to δ^A. In particular, the first equation holds.

Consider the second equation in (1). The left side is the operator π^ω and hence, by (3.3.3), the left side is equal to the operator $\partial^2\partial^1\partial^2 a_2 a_3^2$. So, by the Coxeter relations proved for ∂^p, the left side is equal to the operator $\partial^1\partial^2\partial^1 a_2 a_3^2$. Hence it suffices to prove that the latter operator is equal to the right side, that is, it suffices to prove the following equation:

$$\partial^1\partial^2\partial^1 a_2 a_3^2 = \partial^1 a_2 \partial^2 a_3 \partial^1 a_2.$$

The latter equation is easily verified. Indeed, on the right side, a_3 commutes with ∂^1. Next apply the equation $a_2\partial^2 = \partial^2 a_3 - 1$ from the proof of Lemma (DIFF.2.8). Finally, use the equation $\partial^1\partial^1 = 0$ of Lemma (DIFF.2.8) to obtain the left hand side. Hence the second equation in (1) has been proved.

Finally, the third equation in (1) follows by applying the involution $\alpha \mapsto \omega\alpha^*\omega^{-1}$ to the second equation, since $\omega(\pi^p)^*\omega^{-1} = -\psi^{n-p}$ by Formula (DIFF.2.4.6).

Thus the second set of relations have been verified for all three sets of operators, and the proof is completed. ∎

(3.5) Definition. Let μ be a permutation of A. Define the corresponding *difference operator* ∂^μ by the following equation:

$$\partial^\mu := \partial^{i_1} \cdots \partial^{i_r}, \tag{3.5.1}$$

where $(\varepsilon_{i_1}, \dots, \varepsilon_{i_r})$ is any minimal presentation of μ. Since the operators ∂^p satisfy the Coxeter-Moore relations by Corollary (3.4), it follows from Proposition (SYM.2.6) that the operator ∂^μ is well defined, that is, the right hand side of the equation is independent of the choice of the minimal presentation of μ.

Define similarly operators π^μ and ψ^μ by the equations,

$$\pi^\mu := \pi^{i_1} \cdots \pi^{i_r} \quad \text{and} \quad \psi^\mu := \psi^{i_1} \cdots \psi^{i_r}.$$

Again, it follows from Corollary (3.4) that the operators are well defined. Finally, using the equations (DIFF.2.4.4), it follows from Corollary (3.4) that also each of the three sets of operators, $\{\partial_1, \dots, \partial_{n-1}\}$, $\{\pi_1, \dots, \pi_{n-1}\}$, and $\{\psi_1, \dots, \psi_{n-1}\}$, satisfies the Coxeter-Moore relations. Hence we obtain operators ∂_μ, π_μ, and ψ_μ, defined by the equations,

$$\partial_\mu := \partial_{i_1} \cdots \partial_{i_r}, \quad \pi_\mu := \pi_{i_1} \cdots \pi_{i_r}, \quad \text{and} \quad \psi_\mu := \psi_{i_1} \cdots \psi_{i_r}.$$

With $\mu := \omega_A$, we obtain the operators defined in (3.1):

$$\partial^\omega = \partial^{\omega_A}, \quad \pi^\omega = \pi^{\omega_A}, \quad \psi^\omega = \psi^{\omega_A},$$

where the left sides are the operators of (3.1). Indeed, a minimal presentation of ω_A is obtained recursively from $\omega_A = \omega_{A_1}\varepsilon_1 \cdots \varepsilon_{n-1}$. Using this, the first two equations follow from the recursive definitions in (3.1). The third equation is a consequence of the definition of ψ^ω in (3.1) and the general formula $\omega(\pi^\mu)^*\omega = (-1)^{\ell(\mu)}\psi^{\omega\mu^{-1}\omega}$ proved in the following Lemma.

(3.6) Lemma. *Let μ be a permutation of A. Then*

$$\partial^\mu = (-1)^{\ell(\mu)}\partial_\mu \text{ and } (\partial^\mu)^* = \partial^{\mu^{-1}}. \tag{3.6.1}$$

Moreover, the following nine formulas hold:

$$\begin{aligned}
&(\partial^\mu)^* = (-1)^{\ell(\mu)}\partial_{\mu^{-1}}, \quad \omega\partial^\mu\omega = \partial_{\omega\mu\omega}, \quad \omega(\partial^\mu)^*\omega = (-1)^{\ell(\mu)}\partial^{\omega\mu^{-1}\omega},\\
&(\pi^\mu)^* = (-1)^{\ell(\mu)}\psi_{\mu^{-1}}, \quad \omega\pi^\mu\omega = \pi_{\omega\mu\omega}, \quad \omega(\pi^\mu)^*\omega = (-1)^{\ell(\mu)}\psi^{\omega\mu^{-1}\omega},\\
&(\psi^\mu)^* = (-1)^{\ell(\mu)}\pi_{\mu^{-1}}, \quad \omega\psi^\mu\omega = \psi_{\omega\mu\omega}, \quad \omega(\psi^\mu)^*\omega = (-1)^{\ell(\mu)}\pi^{\omega\mu^{-1}\omega}.
\end{aligned}$$

Proof. The first equation in (3.6.1) follows from the definition of ∂^μ and ∂_μ, because $\partial^p = -\partial_p$. The second equation in (3.6.1) follows from the first equation and the first of the nine formulas.

To prove the nine formulas, let $(\varepsilon_{i_1}, \ldots, \varepsilon_{i_r})$ be a minimal presentation of μ. Consider the equation of Definition (3.5):

$$\partial^\mu = \partial^{i_1} \cdots \partial^{i_r}. \tag{1}$$

Apply the canonical involution $\alpha \mapsto \alpha^*$. The canonical involution reverses the order of the factors in a product, and $(\partial^p)^* = -\partial_p$ by the first equation of (DIFF.2.4.5). Hence, from Equation (1) we obtain the equation,

$$(\partial^\mu)^* = (-1)^r \partial_{i_r} \cdots \partial_{i_1}. \tag{2}$$

As the reversed sequence $(\varepsilon_{i_r}, \ldots, \varepsilon_{i_1})$ is a minimal presentation of μ^{-1}, it follows that the right side (2) is equal to $(-1)^{\ell(\mu)}\partial_{\mu^{-1}}$. Hence the first of the nine formulas holds.

To prove the second of the nine formulas, apply the conjugation $\alpha \mapsto \omega\alpha\omega$ to Equation (1). Conjugation is a homomorphism, and $\omega\partial^p\omega = \partial_{n-p}$ by the first equation of (DIFF.2.4.4). Hence from Equation (1) we obtain the equation,

$$\omega\partial^\mu\omega = \partial_{n-i_1} \cdots \partial_{n-i_r}. \tag{3}$$

As the sequence $(\varepsilon_{n-i_1}, \ldots, \varepsilon_{n-i_r})$ is a minimal presentation of $\omega\mu\omega$, it follows that the right side of (3) is equal to $\partial_{\omega\mu\omega}$. Hence the second of the nine formulas holds.

Clearly, the third of the nine formulas is consequence of the first two formulas. Finally, the proofs of remaining two sets of three formulas are entirely analogous to the proof of the first three formulas. ∎

(3.7) Proposition. *Let μ and ν be permutations. Then:*

(1) *If $\ell(\mu)+\ell(\nu)=\ell(\mu\nu)$, then $\partial^\mu\partial^\nu=\partial^{\mu\nu}$, $\pi^\mu\pi^\nu=\pi^{\mu\nu}$ and $\psi^\mu\psi^\nu=\psi^{\mu\nu}$.*
(2) *If $\ell(\mu)+\ell(\nu)>\ell(\mu\nu)$, then $\partial^\mu\partial^\nu=0$.*

Proof. Let $(\varepsilon_{i_1},\dots,\varepsilon_{i_r})$ and $(\varepsilon_{j_1},\dots,\varepsilon_{j_s})$ be minimal presentations of μ and ν. Then the concatenated sequence $(\varepsilon_{i_1},\dots,\varepsilon_{i_r},\varepsilon_{j_1},\dots,\varepsilon_{j_s})$ is a presentation of $\mu\nu$. Moreover, the latter presentation is minimal if and only if $\ell(\mu)+\ell(\nu)=\ell(\mu\nu)$. Hence assertion (1) follows from the Definition (3.5).

If the concatenated sequence is not minimal, then, by Proposition (SYM.2.6), it is Coxeter–Moore equivalent to a sequence $(\varepsilon_{k_1},\dots,\varepsilon_{k_t})$ where two consecutive k_i's are equal. It follows from Corollary (3.4) that $\partial^\mu\partial^\nu$ is equal to the product $\partial^{k_1}\cdots\partial^{k_t}$. The latter product vanishes by the second equation of Lemma (DIFF.2.8). Therefore assertion (2) holds. □

(3.8) Corollary. *For all μ in $\mathfrak{S}(A)$ we have that*

$$\partial^\omega=\partial^{\omega\mu^{-1}}\partial^\mu=\partial^\mu\partial^{\mu^{-1}\omega},$$
$$\pi^\omega=\pi^{\omega\mu^{-1}}\pi^\mu=\pi^\mu\pi^{\mu^{-1}\omega},\quad \psi^\omega=\psi^{\omega\mu^{-1}}\psi^\mu=\psi^\mu\psi^{\mu^{-1}\omega}.$$

Proof. It follows from assertions (5) and (6) in Lemma (SYM.1.3) that $\ell(\omega)=\ell(\omega\mu^{-1})+\ell(\mu)$. Hence the formulas of the Corollary follow from the first assertion of the Proposition. □

(3.9) Lemma. *For a multi index $K=(k_1,\dots,k_n)$, let $R[A]_{\preceq K}$ denote the R-submodule of $R[A]$ generated by the monomials a^J for $J\preceq K$. Assume that $|k_{p+1}-k_p|\le 1$ for all $p<n$. Then the R-submodule $R[A]_{\preceq K}$ is invariant under the operators ∂^μ for all permutations μ in $\mathfrak{S}(A)$.*

Proof. The operator ∂^μ is a composition of the simple difference operators ∂^p. Therefore we may assume that $\mu=\varepsilon_p$ for some $p<n$. It suffices to prove for a given multi index $J\preceq K$ that the value $\partial^p(a^J)$ is an R-linear combination of monomials a^I for $I\preceq K$. Moreover, since ∂^p is linear with respect to polynomials that do not depend on the variables a_p and a_{p+1}, we may assume that a_p and a_{p+1} are the only letters of the alphabet, that is, we may assume that $n=2$ and $\mu=\varepsilon_1$. Then $K=(k_1,k_2)$, and $J=(j_1,j_2)\preceq K$, and we consider the monomials $a^I=a_1^{i_1}a_2^{i_2}$ occurring in the value $\partial^1(a^J)$.

Let k denote the larger of k_1 and k_2, and let j denote the larger of j_1 and j_2. By the hypothesis on K, we have that $(k-1,k-1)\preceq K$, and by the assumption on J we have that $j\le k$. Hence $(j-1,j-1)\preceq K$. Therefore, to prove the assertion, it suffices to prove for any of the monomials a^I occurring in $\partial^1(a^J)$ that $I\preceq(j-1,j-1)$.

Clearly, the expansion of $\partial^1(a^J)$ is given by the formula,

$$\partial^1 a^J=\frac{a_1^{j_1}a_2^{j_2}-a_2^{j_1}a_1^{j_2}}{a_2-a_1}=\pm\sum_{i_1,i_2}{}' a_1^{i_1}a_2^{i_2},$$

where the sum is over all pairs $I = (i_1, i_2)$ such that i_1 and i_2 are less than or equal to $j-1$, and $i_1 + i_2 = j_1 + j_2 - 1$. In particular, for each of the occurring monomials a^I we have that $I \preceq (j-1, j-1)$, as asserted.

Thus we have proved the Lemma. ▯

(3.10) Lemma. *Given invertible rational functions $f_1, \dots, f_{n-1}$ and $g_1, \dots, g_{n-1}$ of $R(A)$. Define operators ξ_i for $i = 1, \dots, n-1$ in $\mathcal{E}_R(A)$ by $\xi_i = f_i(1 - \varepsilon_i)g_i$. Then, for every permutation μ in $\mathfrak{S}(A)$ and every minimal presentation $(\varepsilon_{i_1}, \dots, \varepsilon_{i_r})$ of μ we have an expansion of operators,*

$$\xi_{i_1} \cdots \xi_{i_r} = h_\mu \mu + \sum_{\ell(\nu) < \ell(\mu)} h_\nu \nu,$$

with rational functions h_ν. Moreover, the coefficient h_μ to μ in the expansion is an invertible function in $R(A)$.

As a consequence, each of the three sets of operators, $\{\partial^\mu\}$, $\{\pi^\mu\}$, and $\{\psi^\mu\}$ for μ in $\mathfrak{S}(A)$, is an $R(A)$-basis for the algebra $\mathcal{E}_R(A)$ of all operators as a left module over its subring $R(A)$.

Proof. Use the rule $\tau \cdot g = \tau(g)\tau$ in the algebra $\mathcal{E}_R(A)$ to develop the expression $\xi_{i_1} \cdots \xi_{i_r} = f_{i_1}(1 - \varepsilon_{i_1})g_{i_1} \cdots f_{i_r}(1 - \varepsilon_{i_r})g_{i_r}$ as an $R(A)$-linear combination of monomials in $\varepsilon_1, \dots, \varepsilon_{n-1}$. Clearly, each occurring monomial is of the form $\varepsilon_{j_1} \cdots \varepsilon_{j_q}$ where $(j_1, \dots, j_q)$ is a subsequence of $(i_1, \dots, i_r)$. In particular, the product $\varepsilon_{j_1} \cdots \varepsilon_{j_q}$ is either equal to μ or it has length strictly less than $\ell(\mu)$. [In fact, the product is less than or equal to μ in the Bruhat–Ehresman order by Proposition (SYM.3.7?).] Moreover, the coefficient h_μ is equal to the sign $(-1)^r$ multiplied by a product of functions obtained by applying suitable permutations to the f_i and g_i. Hence h_μ is invertible in $R(A)$. Thus the first two assertions of the Lemma hold.

Since the operators ∂^μ, π^μ and ψ^μ are all of the form $\xi_{i_1} \cdots \xi_{i_r}$ for particular choices of invertible f_i and g_i, it follows from the expansion, by induction on the length, that every permutation μ belongs to the $R(A)$-module generated by any of these sets of operators. However, the $R(A)$-module of operators can be identified with the twisted group algebra $R(A)[\mathfrak{S}(A)]$ and so the permutations μ in $\mathfrak{S}(A)$ form a basis for $\mathcal{E}_R(A)$ as an $R(A)$–module. Moreover, the number of operators in any of the three sets is equal to the number of permutations. Hence any of the three sets is a basis. ▯

(3.11) Note. We will prove in Section (SCHUB.2.4) that the operators ∂^μ for μ in $\mathfrak{S}(A)$ form a basis for the algebra of operators $\mathcal{E}_R[A]$ as a module over the ring of polynomials $R[A]$.

(3.12) Lemma. *Let μ be a permutation in $\mathfrak{S}(A)$ and let g be a rational function which is symmetric in the letters $a_2, \dots, a_n$. Then the derivative $\partial^\mu(g)$ vanishes unless $\mu = \varepsilon_{q-1} \cdots \varepsilon_1$ for some $q = 1, \dots, n$. Moreover, the derivative $\partial^{q-1} \cdots \partial^1(g)$ is symmetric in the letters $a_1, \dots, a_q$ and in the letters $a_{q+1}, \dots, a_n$.*

Proof. Clearly, a function f is symmetric if and only if $\partial^p(f) = 0$ for $p = 1, \dots, n-1$.

Consider, for $q = 1, \dots, n$, the permutation $\sigma_q := \varepsilon_{q-1} \cdots \varepsilon_1$. [Thus σ_q is the identity if $q = 1$ and the q-cycle $(q, \dots, 2, 1)$ for $q > 1$.] Then, for $p = 1, \dots, n-1$ and $p \neq q-1$, we have the equation,

$$\varepsilon_p\sigma_q = \begin{cases} \sigma_q\varepsilon_p & \text{if } p > q, \\ \sigma_{q+1} & \text{if } p = q, \\ \sigma_q\varepsilon_{p+1} & \text{if } p < q-1. \end{cases} \tag{3.12.1}$$

The equations are easily verified. For instance, the equation $\varepsilon_q\sigma_q = \sigma_{q+1}$ follows from the definition of the σ_q, and the remaining equations follow from the Coxeter–Moore relations.

Consider a minimal presentation of μ. If μ is not of the form σ_q, then there is a unique $q \geq 1$ so that the presentation is of the form,

$$\mu = \varepsilon_{p_1}\cdots\varepsilon_{p_s}\varepsilon_{q-1}\cdots\varepsilon_1,$$

where $p_s \neq q$. Since the presentation is minimal, it follows that $p_s \neq q-1$. Hence, by Equation (3.12.1), there is a second minimal presentation of the form $\mu = \varepsilon_{q_1}\cdots\varepsilon_{q_r}$ where $q_r \geq 2$. Therefore, since g is symmetric in the letters $a_2,\dots,a_n$, it follows that $\partial^\mu(g) = 0$.

To prove the second assertion of the Lemma, let $h := \partial^{\sigma_q}(g)$. Consider, for $1 \leq p \leq n-1$, the derivative $\partial^p(h)$. For $p = q-1$ we have that $\partial^{q-1}(h) = 0$, since $\partial^{q-1}\partial^{q-1} = 0$. For $p < q-1$, it follows from Equation (3.12.1) that $\partial^p(h) = \partial^{\sigma_q}\partial^{p+1}(g)$, and $\partial^{p+1}(g) = 0$ since g is symmetric in $a_2,\dots,a_n$. Thus $\partial^p(h) = 0$ for $p = 1,\dots,q-1$, and hence h is symmetric in the letters $a_1,\dots,a_q$. Similarly, if $p > q$, it follows that $\partial^p(h) = \partial^{\sigma_q}\partial^p(g) = 0$, and hence h is symmetric in the letters $a_{q+1},\dots,a_n$. ∎

(3.13) Definition. Consider an operator $\alpha \in \mathcal{E}_R[A]$. As an element of $\mathcal{E}_R(A)$ it defines a $\mathrm{Sym}_R(A)$-linear map $R(A) \to R(A)$. By definition, α induces a $\mathrm{Sym}_R[A]$-linear action on polynomials, also denoted $\alpha\colon R[A] \to R[A]$. The former α is the unique extension of the latter α to a $\mathrm{Sym}_R(A)$-linear action on rational functions.

In addition α acts naturally on the power series: If $f = \sum f_i$ is a power series in $R[[A]]$, with f_i homogeneous of degree i, then there is a well defined power series $\alpha(f)$ given as the convergent series,

$$\alpha(f) := \sum_i \alpha(f_i).$$

Indeed, since $\alpha\colon R[A] \to R[A]$ is $\mathrm{Sym}_R[A]$-linear, it follows from Corollary (1.4) that $\alpha(f_i)$ is of order at least $i - n(n-1)/2$.

In particular, the difference operators ∂^μ act on the power series ring $R[[A]]$. Note that the operator ∂^μ is homogeneous and lowers the degree by $d = \ell(\mu)$, that is, the homogeneous term of degree i in $\partial^\mu(f)$ is equal to $\partial^\mu(f_{i+d})$.

Clearly, the Leibnitz Formula of (2.7) holds for power series f, g in $R[[A]]$.

(3.14) Example. For $1 \leq q < n$, we have in $R(A)$ and in $R[[A]]$ the equations,

$$\partial_q\Big(\frac{1}{1-a_q}\Big) = \frac{1}{(1-a_q)(1-a_{q+1})}, \tag{3.14.1}$$

$$\partial_q\cdots\partial_1\Big(\frac{1}{1-a_1}\Big) = \frac{1}{(1-a_1)\cdots(1-a_{q+1})}. \tag{3.14.2}$$

Indeed, the first formula is obtained by applying the Leibnitz rule to the equation $(1-a_q)^{-1}(1-a_q) = 1$, and the second formula follows by induction on q from the first.

(3.15) Example. For a polynomial g depending only on the first letter a_1, it follows from (3.12) that the derivatives $\partial_\mu(g)$ vanish unless $\mu = \varepsilon_{q-1}\cdots\varepsilon_1$ for some $q = 1,\dots,n$. For a polynomial g of the form,

$$g := (a_1 - b_1)\cdots(a_1 - b_m), \tag{3.15.1}$$

a formula for the derivative $\partial_{q-1}\cdots\partial_1(g)$ may be obtained as follows:

Assume first that the b_i are variables, that is, let $B = \{b_1,\dots,b_m\}$ be a second alphabet with m letters, and replace the ground ring with the polynomial ring $R[B]$. Form, for $q = 1,\dots,n$, the power series in $R[[A,B]]$,

$$S^{(q)} = S(a_1,\dots,a_q;B) := \frac{(1-b_1)\cdots(1-b_m)}{(1-a_1)\cdots(1-a_q)}, \tag{3.15.2}$$

and denote by $s_d(a_1,\dots,a_q;B)$ the homogeneous term of degree d in $S^{(q)}$. Then

$$\partial_{q-1}\cdots\partial_1(g) = s_{m-q+1}(a_1,\dots,a_q;B). \tag{3.15.3}$$

Indeed, the numerator in $S^{(q)}$ is the polynomial,

$$(1-b_1)\cdots(1-b_m) = 1 - e_1 + \cdots + (-1)^m e_m,$$

where $e_i := e_i(B)$. Hence, for $q = 1$, we have the equation,

$$S^{(1)} = (1 - e_1 + \cdots + (-1)^m e_m)(1 + a_1 + a_1^2 + \cdots).$$

It follows that the homogeneous term of degree d in $S^{(1)}$, for $d \geq m$, is the polynomial,

$$a_1^d - e_1 a_1^{d-1} + \cdots + (-1)^m e_m a_1^{d-m} = a_1^{d-m}(a_1 - b_1)\cdots(a_1 - b_m).$$

In particular, for $d = m$, it follows that g is the homogeneous term of degree m in $S^{(1)}$. It follows from (3.14.2) that

$$\partial_{q-1}\cdots\partial_1 S^{(1)} = S^{(q)}.$$

Therefore, by taking the homogeneous terms of degree $m - q + 1$, we obtain the asserted formula (3.15.3).

In (3.15.3), the b_i are assumed to be variables. The formula for the general case when the b_i are arbitrary elements of R is obtained by specializing. Note however, that in the collection of the homogeneous terms in the $S^{(q)}$, the b_i have to be considered as homogeneous of degree 1.

For instance, let $b_1 = \cdots = b_m = 0$. Then $g = a_1^m$ and the power series $S^{(q)}$ of (3.15.2) specializes to the series,

$$S(a_1,\dots,a_q) = \frac{1}{(1-a_1)\cdots(1-a_q)}.$$

It follows that

$$\partial_{q-1}\cdots\partial_1(a_1^m) = s_{m-q+1}(a_1,\dots,a_q)$$

is the $(m-q+1)$'th complete symmetric polynomial in the letters $a_1,\dots,a_q$.

4. The bilinear form.

(4.1) Definition. Define a bilinear form on the algebra $R(A)$ of rational functions by the following equation:

$$\langle f, g\rangle := \sum_{\sigma\in\mathfrak{S}(A)} \sigma\Big(\frac{fg}{\Delta}\Big) = \frac{1}{\Delta}\sum_{\sigma\in\mathfrak{S}(A)} (\operatorname{sign}\sigma)\sigma(fg). \tag{4.1.1}$$

By Definition (DIFF.2.3) of the symmetrization operator δ^A, the two sums are equal, and equal to the value $\delta^A(fg)$. Moreover, by Theorem (DIFF.3.3) and Definition (DIFF.3.5), the following equation holds:

$$\langle f, g\rangle = \partial^\omega(fg). \tag{4.1.2}$$

The bilinear form is called the *inner product* on $R(A)$.

(4.2) Lemma. *The values of the inner product* (4.1) *are symmetric functions in* $R(A)$. *Moreover, the inner product is symmetric and* $\operatorname{Sym}_R(A)$*-bilinear. Furthermore, if* α *is any operator in* $\mathcal{E}_R(A)$*, then*

$$\langle \alpha(f), g\rangle = \langle f, \alpha^*(g)\rangle, \tag{4.2.1}$$

where $\alpha \mapsto \alpha^*$ *is the canonical involution of* (DIFF.2.1). *In particular, for any permutation* μ *in* $\mathfrak{S}(A)$*, we have that*

$$\langle \mu(f), g\rangle = (\operatorname{sign}\mu)\langle f, \mu^{-1}(g)\rangle \quad \textit{and} \quad \langle \partial^\mu(f), g\rangle = \langle f, \partial^{\mu^{-1}}(g)\rangle. \tag{4.2.2}$$

Finally, if f *and* g *are polynomials in* $R[A]$*, then* $\langle f, g\rangle$ *belongs to* $\operatorname{Sym}_R[A]$.

Proof. We have observed in (DIFF.2.3) the value $\langle f, g\rangle = \delta^A(fg)$ is a symmetric function. Moreover, it is clear that the inner product is symmetric in f and g, and $\operatorname{Sym}_R(A)$-bilinear.

By additivity of the inner product it suffices to prove equation (4.2.1) when α is a product $h\mu$ of a function h in $R(A)$ and a permutation μ in $\mathfrak{S}(A)$. Moreover, since the canonical involution reverses the orders of the factors in a product, it suffices to treat separately the two cases: $\alpha = h$ and $\alpha = \mu$. In the first case the equation is obvious: $\langle hf, g\rangle = \delta^A(hfg) = \langle f, hg\rangle$. In the second case the equation is the first equation of (4.2.2), and it follows immediately by rearranging the terms in second sum in (4.1.1).

By the second equation in (DIFF.3.6.1), we have that $(\partial^\mu)^* = \partial^{\mu^{-1}}$. Hence the second equation of (4.2.2) is a special case of the general equation (4.2.1).

The final assertion of the Lemma is a consequence of the equation $\langle f, g\rangle = \delta^A(fg)$, since the symmetrization operator δ^A belongs to $\mathcal{E}_R[A]$, see Definition (DIFF.2.3). ∎

(4.3) Note. The inner product is non-degenerate, that is, if $\langle f, g\rangle = 0$ for all g, then $f = 0$. Indeed, it follows from Equation (4.1.1) that $\langle f, g\rangle$ is the result of evaluating the operator $\sum_\sigma \sigma f = \sum_\sigma \sigma(f)\sigma$ on the function g/Δ. If the result is 0 for all g, then the operator is zero and consequently, by (DIFF.2.2), zero as an element in the twisted group algebra $R(A)[\mathfrak{S}(A)]$. Hence $f = 0$.

We prove in (SCHUB.2.4) that if $\langle f, g\rangle$ is a polynomial for all polynomials g, then f is a polynomial.

(4.4) Lemma. *Let I and J be multi indices of size n such that $I \preceq (n-1, \dots, 1, 0)$ and $J \preceq (0, 1, \dots, n-1)$. Then the inner product $\langle a^I, a^J \rangle$ is equal to 0 unless all entries in the sequence $I + J$ are different. In the exceptional case, the sequence $I + J$ is a permutation of the sequence $(0, 1, \dots, n-1)$, and $\langle a^I, a^J \rangle$ is equal to the signature of the latter permutation.*

Proof. By definition we have that $\langle a^I, a^J \rangle = \delta^A(a^{I+J})$. Moreover, as observed in Theorem (DIFF.3.3), we have that $\delta^A(a^K) = \Delta^K/\Delta$. From the assumptions on I and J it follows that all entries in the sequence $I + J$ are non-negative integers between 0 and $n-1$. The assertion of the Lemma is a consequence, because Δ^K is alternating in K and equal to Δ when $K = (0, 1, \dots, n-1)$. ∎

5. The Möbius transformation.

Schubert polynomials

1. Double Schubert polynomials.

(1.1) Setup. Fix an alphabet $A = \{a_1, \dots, a_n\}$ with $n \geq 0$ letters, and consider the algebras $R[A]$ of polynomials and $R(A)$ of rational functions in the letters of A. For a multi index I of size n, we denote by $R[A]_{\preceq I}$ the R-module of polynomials generated by the monomials a^J for $J \preceq I$. It follows from Lemma (DIFF.3.9) that the two R-modules, $R[A]_{\preceq 0,1,\dots,n-1}$ and $R[A]_{\preceq n-1,\dots,1,0}$, are invariant under the operators ∂^μ for $\mu \in \mathfrak{S}(A)$. Recall that $\partial_\mu = (\operatorname{sign}\mu)\partial^\mu$.

Fix a set $B = (b_1, \dots, b_n)$ of n elements in the ground ring R. If f is a polynomial in $R[A]$, we denote by $f(B)$ the value obtained by specializing the letters of A to the elements of B.

Consider the following two polynomials in $R[A]$:

$$X^B := \prod_{p+q \leq n} (a_p - b_q) \quad \text{and} \quad Y^B := \prod_{p>q} (a_p - b_q). \tag{1.1.1}$$

Note that $Y^B = \omega(X^B)$.

When the ground ring is a ring of polynomials over an alphabet B with the n letters b_i, we write $X(A,B) := X^B$ and $Y(A,B) := Y^B$.

(1.2) Lemma. *The following assertions hold:*

(1) *The value $Y^B(B)$ is equal to $\Delta(B)$. Moreover, for a permutation $\mu \neq 1$ in $\mathfrak{S}(A)$, the value $\mu(Y^B)(B)$ is equal to 0.*

(2) *The polynomials $\partial^\omega(Y^B)$ and $\partial_\omega(X^B)$ are equal to 1. Moreover, for a permutation $\mu \neq \omega$ in $\mathfrak{S}(A)$, the value $\partial_\mu(X^B)(B)$ is equal to 0.*

(3) *If f is a polynomial in $R[A]_{\preceq n-1,\dots,1,0}$, then*

$$\langle Y^B, f \rangle = f(B).$$

Proof. In (3) we may, by linearity of the inner product, assume that f is a monomial a^I for $I \preceq (n-1, \dots, 1, 0)$ in which case the equation asserts that $\langle Y^B, a^I \rangle = b^I$. It suffices to prove the latter equation and the equations of (1) and (2) in the special case where the ground ring is a polynomial ring $R[B]$ over an alphabet B with n letters. Indeed, if the equations hold in the special case, then the equations in the general case follow by specializing the letters of

B to the given sequence of elements in R. In the polynomial ring $R[B]$, the Vandermonde determinant $\Delta(B)$ is not a zero divisor. Therefore, in the remaining part of the proof we may assume that the value $\Delta(B)$ is a not a zero divisor in R.

Consider the assertions of (1). It is obvious from the definition (1.1.1) of Y^B that $Y^B(B) = \Delta(B)$. The polynomial $\mu(Y^B)$ for a permutation μ is the product of the factors $\mu(a_p) - b_q$ for $p > q$. Assume that $\mu \neq 1$. Then there exist indices $p > q$ such that $\mu(a_p) = a_q$. Clearly, the corresponding factor $\mu(a_p) - b_q$ specializes to 0, and consequently $\mu(Y^B)(B) = 0$. Hence the assertions of (1) have been proved.

Consider the first assertion of (2). The operator ∂^ω is the symmetrization operator δ^A. Hence $\partial^\omega = (\operatorname{sign}\omega)\partial^\omega\omega = \partial_\omega\omega$. As $X^B = \omega(Y^B)$, it follows that $\partial_\omega(X^B) = \partial^\omega(Y^B)$. Hence it suffices to prove that $\partial^\omega(Y^B) = 1$. Let $E := (0, 1, \ldots, n-1)$. Clearly, by expanding the product defining Y^B we obtain an R-linear combination of monomials a^J where $J \preceq E$, and the coefficient of a^E is equal to 1. The values of the operator $\partial^\omega = \delta^A$ on the monomials a^J for $J \preceq E$ were determined in Theorem (DIFF.3.3). It follows that $\partial^\omega(Y^B) = 1$.

To prove the second assertion of (2), assume that $\mu \neq \omega$. It follows from Lemma (DIFF.3.10) that there is an expansion of operators,

$$\partial_\mu = \sum_{\ell(\nu)\leq\ell(\mu)} h_\nu \nu,$$

with rational functions h_ν. Since $\mu \neq \omega$, it follows that all permutations ν in the expansion are different from ω. Clearly, the denominators in the rational functions h_ν are products of factors of the form $a_p - a_q$ for $p \neq q$. Therefore, when N is sufficiently big, the operator $\Delta^N\partial_\mu$ is an $R[A]$-linear combination of permutations $\nu \neq \omega$. It follows from part (1) that if $\nu \neq \omega$, then $\nu(X^B)(B) = \nu\omega(Y^B)(B) = 0$. Consequently, when the operator $\Delta^N\partial_\mu$ is applied to X^B and A is specialized to B, we obtain the equation $\Delta^N(B)\partial_\mu(X^B)(B) = 0$. Since $\Delta(B)$ is assumed to be a non-zero divisor, the latter equation implies that $\partial_\mu(X^B)(B) = 0$. Hence the assertions of (2) have been proved.

Consider assertion (3). The polynomial Y^B is an R-linear combination of monomials a^J, where $J \preceq (0, 1, \ldots, n-1)$, and f is an R-linear combination of monomials a^I, where $I \preceq (n-1, \ldots, 1, 0)$. From the latter conditions on I and J, it follows from Lemma (DIFF.4.4) that the inner product $\langle a^J, a^I\rangle$ is either 0 or ± 1. As a consequence, the inner product $\langle Y^B, f\rangle$ belongs to R.

On the other hand, from the definition of the inner product we obtain the equation,

$$\Delta\langle Y^B, f\rangle = \sum_{\mu\in\mathfrak{S}(A)} (\operatorname{sign}\mu)\mu(Y^B f).$$

Specialize A to B in the latter equation. On the left hand side, the first factor Δ specializes to $\Delta(B)$ and the second factor is left unchanged since it belongs to R. On the right hand side, by (1), the terms $\mu(Y^B f) = \mu(Y^B)\mu(f)$ corresponding to $\mu \neq 1$ specialize to 0 and the term corresponding to $\mu = 1$ specializes to $\Delta(B)f(B)$. Hence the specialization yields the following equation:

$$\Delta(B)\langle Y^B, f\rangle = \Delta(B)f(B).$$

As $\Delta(B)$ is assumed to be a non-zero divisor in R, the latter equation implies the equation of Assertion (3).

Hence all assertions of the Lemma have been proved. □

(1.3) Definition. In the setup of (1.1), define a family of polynomials X^B_μ indexed by permutations μ in $\mathfrak{S}(A)$, by the equation,

$$X^B_\mu = \partial_{\mu^{-1}\omega}(X^B). \tag{1.3.1}$$

When the ground ring is of the form $R[B]$, where B is a second alphabet with n letters, the resulting polynomials are called *double Schubert polynomials*, and they are often denoted $X_\mu(A, B)$. The double Schubert polynomials belong to the ring of polynomials $R[A, B]$.

It will be convenient to define an auxiliary family of polynomials,

$$Y^B_\mu = \partial_{\mu^{-1}}(Y^B) = \partial_{\mu^{-1}}(\omega X^B). \tag{1.3.2}$$

By Lemma (DIFF.3.6) we have that $\omega\partial_\mu\omega = (\operatorname{sign}\mu)\partial_{\omega\mu\omega}$. Use this equation for $\mu := \mu^{-1}$ and the definitions to transform the right side of (1.3.2) into the right side of the following equation:

$$Y^B_\mu = (\operatorname{sign}\mu)\omega(X^B_{\mu\omega}). \tag{1.3.3}$$

Clearly, $X^B_\omega = X^B$ and $Y^B_1 = Y^B$. Moreover, $X^B_1 = 1$ by Lemma (1.2)(2), and then $Y^B_\omega = (-1)^{n(n-1)/2}$ by (1.3.3).

(1.4) Remark. Note the following rule for the calculation of X^B_μ: represent μ as the sequence of indices $(i_1 \dots i_n)$ where $\mu(a_p) = a_{i_p}$. Rearrange by simple transpositions the elements in sequence so that the sequence becomes strictly decreasing, that is, solve the equation $\mu\varepsilon_{p_1}\cdots\varepsilon_{p_r} = \omega$ with a minimal number r. Then $X^B_\mu = \partial_{p_1}\cdots\partial_{p_r}(X^B)$.

(1.5) Example. For $n = 3$ we obtain, omitting the superscript B, the following polynomials:

$$\begin{aligned}
&(321) = \omega, && X_{321} = X = (a_1 - b_1)(a_1 - b_2)(a_2 - b_1),\\
&(312)\varepsilon_2 = \omega, && X_{312} = \partial_2 X = (a_1 - b_1)(a_1 - b_2),\\
&(231)\varepsilon_1 = \omega, && X_{231} = \partial_1 X = (a_1 - b_1)(a_2 - b_1),\\
&(132)\varepsilon_1\varepsilon_2 = \omega, && X_{132} = \partial_1\partial_2 X = a_1 - b_1 + a_2 - b_2,\\
&(213)\varepsilon_2\varepsilon_1 = \omega, && X_{213} = \partial_2\partial_1 X = a_1 - b_1,\\
&(123)\varepsilon_2\varepsilon_1\varepsilon_2 = \omega, && X_{123} = \partial_2\partial_1\partial_2 X = 1.
\end{aligned}$$

(1.6) Lemma. *The polynomial X^B_μ belongs to the R-module $R[A]_{\le n-1,\dots,1,0}$. Its degree is equal to $\ell(\mu)$. Moreover, the polynomial X^B_1 is equal to 1, and if $\mu \ne 1$, then the value*

$X^B_\mu(B)$ is equal to 0. Finally, if μ and ν are any permutations in $\mathfrak{S}(A)$, then the following equation holds:

$$\partial_\nu(X^B_\mu) = \begin{cases} X^B_{\mu\nu^{-1}} & \text{if } \ell(\mu) = \ell(\mu\nu^{-1}) + \ell(\nu), \\ 0 & \text{otherwise.} \end{cases} \tag{1.6.1}$$

Proof. It is obvious from the definition that the polynomial X^B belongs to the submodule $R[A]_{\preceq n-1,\dots,1,0}$. As noted in (1.1), the submodule is invariant under the operators ∂_ν. Therefore, the first assertion of the Lemma holds.

The polynomial X^B has degree equal to $\ell(\omega)$ and the difference operator ∂_ν lowers the degree by $\ell(\nu)$. Therefore the degree of X^B_μ is at most equal to $\ell(\omega) - \ell(\mu^{-1}\omega) = \ell(\mu)$. It follows from the last equation of the Lemma, applied with $\nu = \mu$, that $\partial_\mu(X^B_\mu) = X^B_1 = 1$. Hence, when the last equation of the Lemma has been proved, it follows that $\partial_\mu X^B_\mu \neq 0$ and therefore, the degree of X^B_μ is equal to $\ell(\mu)$.

The equation $X^B_\mu(B) = 0$ for $\mu \neq 1$ follows from the definition of X^B_μ and Lemma (1.2)(2).

Consider finally Equation (1.6.1). By definition, the left hand side is equal to the result of applying the operator $\partial_\nu\partial_{\mu^{-1}\omega}$ to the polynomial X^B. By Proposition (DIFF.3.7), the latter operator is equal to zero unless the following condition holds:

$$\ell(\nu) + \ell(\mu^{-1}\omega) = \ell(\nu\mu^{-1}\omega). \tag{1.6.2}$$

When condition (1.6.2) holds, the operator is equal to $\partial_{\nu\mu^{-1}\omega}$, and consequently, when applied to X^B the result is the polynomial $X^B_{\mu\nu^{-1}}$. Moreover, the condition for the first case in Equation (1.6.1) is equivalent to the condition (1.6.2), as it follows from Lemma (SYM.1.3). Hence the final equation of the Lemma holds. ∎

(1.7) Theorem. *The polynomials X^B_μ for permutations μ in $\mathfrak{S}(A)$ form a basis for the ring $R[A]$ of polynomials as a module over the subring* $\mathrm{Sym}_R[A]$ *of symmetric polynomials, and a basis for the ring $R(A)$ of rational functions as a module over the subring* $\mathrm{Sym}_R(A)$ *of symmetric functions. In the expansion of a function f in the latter basis, the coefficient to X^B_μ is given by the inner products,*

$$\langle Y^B_\mu, f\rangle = \langle \partial_{\mu^{-1}}(Y^B), f\rangle = \langle Y^B, \partial_\mu(f)\rangle. \tag{1.7.1}$$

Moreover, if μ and ν are permutations in $\mathfrak{S}(A)$, then we have the equation,

$$\langle Y^B_\mu, X^B_\nu\rangle = \begin{cases} 1 & \text{if } \mu = \nu, \\ 0 & \text{otherwise.} \end{cases} \tag{1.7.2}$$

Finally, the polynomials X^B_μ form a basis for the R-module $R[A]_{\preceq n-1,\dots,1,0}$ and, for every polynomial f in the latter module, the coefficient (1.7.1) *of X^B_μ satisfies the equation,*

$$\langle Y^B_\mu, f\rangle = \partial_\mu(f)(B), \tag{1.7.3}$$

that is, we have the Newton interpolation formula,

$$f = \sum_{\mu \in \mathfrak{S}(A)} \partial_\mu(f)(B) X^B_\mu. \tag{1.7.4}$$

Proof. In (1.7.1), the first equation follows from the definition (1.3.2), and the second equation holds, because the operators $\partial_{\mu^{-1}}$ and ∂_μ are adjoint with respect to the inner product by Lemma (DIFF.4.2). Hence the inner products in (1.7.1) are equal.

Let N be the R-submodule $N := R[A]_{\preceq n-1,\dots,1,0}$ of $R[A]$. If f belongs to N, then it follows from Lemma (1.2)(3) and the equations of (1.7.1) that $\langle Y^B_\mu, f\rangle = \langle Y^B, \partial_\mu(f)\rangle = \partial_\mu(f)(B)$. Hence Equation (1.7.3) holds.

By Lemma (1.6) we can apply Equation (1.7.3) with $f := X^B_\nu$. As a consequence,

$$\langle Y^B_\mu, X^B_\nu\rangle = \partial_\mu(X^B_\nu)(B). \tag{1.7.5}$$

Assume first that $\mu = \nu$. Then, by Lemma (1.6), the polynomial $\partial_\mu(X^B_\nu)$ is equal to 1. In particular, the value on the right hand side of (1.7.5) is equal to 1. Assume next that $\mu \neq \nu$. Then, again by Lemma (1.6), the polynomial $\partial_\mu(X^B_\nu)$ is either equal to 0 or it is of the form X^B_τ, with $\tau \neq 1$. In the latter case, the value $X^B_\tau(B)$ is equal to 0 by Lemma (1.6). Therefore, for $\mu \neq \nu$, the value on the right side of (1.7.5) is equal to 0. Hence Equation (1.7.2) follows from (1.7.5).

By a standard argument of linear algebra, the remaining assertions of the Theorem are consequences of the equations (1.7.2). Indeed, let $d = n!$ be the cardinality of $\mathfrak{S}(A)$. Let M denote the algebra of rational functions $R(A)$ as a module over the ring $S := \mathrm{Sym}_R(A)$. Define, for each of the d elements ν in $\mathfrak{S}(A)$, an S-linear form $\check{X}_\nu$ on the module M:

$$\check{X}_\nu(f) := \langle Y^B_\nu, f\rangle.$$

Now, the d elements X^B_μ of the S-module M define an S-linear map $X\colon S^d \to M$, and the d linear forms $\check{X}_\nu$ define an S-linear map $\check{X}\colon M \to S^d$. The equation (1.7.2) asserts that $\check{X}X = 1$. By Proposition (DIFF.1.3), the S-module M can be identified with S^d via the basis formed by the d monomials a^I for $I \preceq (n-1,\dots,1,0)$. Under the latter identification, the linear maps X and $\check{X}$ are $d \times d$ matrices. Consequently, the equation $\check{X}X = 1$ implies that the matrix X is invertible. Hence the polynomials X^B_μ form an S-basis for M. Moreover, the linear forms $\check{X}_\nu$ form the dual basis.

Clearly, in the expansion of a rational function f of $R(A)$ in terms of the basis X^B_μ, the coefficient of X^B_μ is given by evaluation of the linear form $\check{X}_\mu$ on f. In other words, the coefficient is given by the expressions of (1.7.1). Hence we have proved the second assertion of the Theorem.

By the same standard argument, applied to the ring of polynomials $R[A]$ as a module over the subring $\mathrm{Sym}_R[A]$, it follows that the polynomials X^B_μ form a $\mathrm{Sym}_R[A]$-basis for $R[A]$. Hence the first assertion of the Theorem has been proved.

Consider finally the R-module N. It contains the polynomials X_μ^B, as noted in Lemma (1.6). Moreover, on the module N the forms $\check{X}_\nu$ take values in R, by (1.7.3). Therefore, again by the same standard argument, the polynomials X_μ^B form an R-basis for N and, in the expansion of a polynomial f of N, the coefficient to X_μ^B is given by the value (1.7.3). Therefore, the interpolation formula of Newton holds. □

(1.8) Example. For $1 \le p < n$ we have the equation,

$$X_{\varepsilon_p}(A, B) = a_1 + \cdots + a_p - b_1 - \cdots - b_p. \tag{1.8.1}$$

The equation follows by applying the interpolation formula to the polynomial f on the right hand side. Clearly, $f(B) = 0$. If $\mu \neq 1$, then $\partial_\mu f$ is equal to 0 unless $\mu = \varepsilon_p$ in which case $\partial_{\varepsilon_p}(f) = 1$.

(1.9) Example. For $0 \le p < n$ we have the following equation,

$$X_{\varepsilon_p \cdots \varepsilon_1}(A, B) = (a_1 - b_1)(a_1 - b_2) \cdots (a_1 - b_p). \tag{1.9.1}$$

Indeed, we may assume that the ground ring is a polynomial ring $R[B]$, and that $b_1, \dots, b_n$ are the letters of B. Denote by g the polynomial in $R[A, B]$ on the right hand side of (1.9.1).

It follows from Example (DIFF.3.15) that $\partial_\mu(g) = 0$ unless μ is of the form $\varepsilon_q \cdots \varepsilon_1$. Moreover, $\partial_q \cdots \partial_1(g)$ is the $(p-q)$'th homogeneous term in the power series in $R[[A, B]]$,

$$S(a_1, \dots, a_{q+1}; B) = \frac{(1 - b_1) \cdots (1 - b_p)}{(1 - a_1) \cdots (1 - a_{q+1})}. \tag{1.9.2}$$

If $q > p$, then $p - q$ is negative, and the homogeneous term vanishes. If $q = p$, then the term is equal to 1. Assume that $q < p$. Then it follows that the value $\partial_q \cdots \partial_1(g)(B)$ is the $(p-q)$'th homogeneous term in the series obtained from (1.9.2) by substituting $a_i := b_i$. The series obtained is the polynomial,

$$(1 - b_{q+2}) \cdots (1 - b_p),$$

which is of degree $p - q - 1$. In particular, its term of degree $p - q$ vanishes.

Hence, the value $\partial_\mu(g)(B)$ is equal to 1 for $\mu = \varepsilon_p \cdots \varepsilon_1$, and equal to zero otherwise. Therefore, the asserted equation $g = X_{\varepsilon_p \cdots \varepsilon_1}(A, B)$ follows from (1.7.4).

(1.10) Note. *The Newton interpolation in one variable.* The monomial a_1^p, for $p < n$, is obtained from the polynomial g of (1.9) by specializing the letters of B to 0. The power series (1.9.2) specializes to the series $S(a_1, \dots, a_{q+1})$. It follows from (1.9) that $\partial_\mu(a_1^p)$ vanishes unless μ is of the form $\varepsilon_q \cdots \varepsilon_1$. Moreover,

$$\partial_q \cdots \partial_1(a_1^p) = s_{p-q}(a_1, \dots, a_{q+1})$$

is the $(p-q)$'th complete symmetric function in the letters $a_1, \dots, a_{q+1}$.

Now, let f be a polynomial in the variable a_1 only, and of degree at most $n-1$. Then, by R-linearity of ∂_μ, it follows that $\partial_\mu(f)$ vanishes unless the permutation μ is of the form $\mu = \varepsilon_q \cdots \varepsilon_1$ for some $q = 0, \dots, n-1$. Therefore, for any sequence B of elements in the ground ring, we obtain from the Newton interpolation formula (1.7.4) and Equation (1.9.1) the following formula,

$$f = \sum_{q=0}^{n-1} \partial_q \cdots \partial_1(f)(B)\,(a_1 - b_1)\cdots(a_1 - b_q). \tag{1.10.1}$$

(1.11) Note. This note presupposes knowledge of multi Schur polynomials, see SCHUR.2. Assume that the ground ring is $R[B]$, where $B = \{b_1, \dots, b_n\}$ is a second alphabet. Then the polynomials $X(A, B)$ and $Y(A, B)$ can be expressed as multi Schur functions. In fact, with the words $A_p := a_1 + \cdots + a_p$ and $B_p := b_1 + \cdots + b_p$ for $p = 1, \dots, n$, we have the formulas,

$$S^{1,3,\dots,2n-3}(A_{n-1} - B_1, \dots, A_1 - B_{n-1}) = X(A, B), \tag{1.11.1}$$

$$S^{n-1,\dots,n-1}(A_n - B_{n-1}, \dots, A_2 - B_1, A_1) = Y(A, B). \tag{1.11.2}$$

Indeed, from the Factorization formula (SCHUR.2.8), applied with $r := n-1$ to the sequence $a_{n-1}, \dots, a_2, a_1$ of letters, we obtain for the left side of (1.11.1) the factorization,

$$S_1(a_{n-1}-B_1)\, S_2(a_{n-2}-B_2) \cdots S_{n-1}(a_1-B_{n-1}) = \prod_{p+q\le n} (a_p - b_q),$$

and, by the same formula, applied with $r := n$ to the sequence $a_n, \dots, a_1$ of letters and the sequence $B_{n-1}, \dots, B_1, B_0 := 0$ of words, we obtain for the left side of (1.11.2) the factorization,

$$S_{n-1}(a_n-B_{n-1}) \cdots S_1(a_2-B_1) S_0(a_1) = \prod_{p>q} (a_p - b_q).$$

2. Simple Schubert polynomials.

(2.1) Definition. We denote by X_μ and Y_μ the polynomials X_μ^B and Y_μ^B corresponding to the sequence $B = (0, \dots, 0)$. We call the polynomials X_μ the *Schubert polynomials*.

By Definition (SCHUB.1.3) we have that $X_\mu = \partial_{\mu^{-1}\omega} X = \operatorname{sign}(\mu\omega)\omega(Y_{\mu\omega})$, where X and Y are the monomials,

$$X := \prod_{i+j\leq n} a_i = a_1^{n-1}\cdots a_n^0 \quad \text{and} \quad Y := \prod_{i>j} a_i = a_1^0 \cdots a_n^{n-1}. \tag{2.1.1}$$

(2.2) Example. For $n = 3$, it follows from the computation in (SCHUB.1.5) that $X_{321} = a_1^2 a_2$, $X_{312} = a_1^2$, $X_{231} = a_1 a_2$, $X_{132} = a_1 + a_2$, $X_{213} = a_1$, and $X_{123} = 1$.

(2.3) Remark. Note that Theorem (SCHUB.1.7) applies to the special sequence $B = (0, \dots, 0)$. In particular, the Schubert polynomials X_μ for μ in $\mathfrak{S}(A)$ form a basis for the algebra of rational functions $R(A)$ as a module over the ring of symmetric functions $\mathrm{Sym}_R(A)$, they form a basis for the algebra of polynomials $R[A]$ as a module over the ring of symmetric polynomials $\mathrm{Sym}_R[A]$, and they form an R-basis for the R-submodule generated by the monomials a^I for $I \preceq (n-1, \dots, 1, 0)$. As it follows from the proofs of the following results, the basis of Schubert polynomials has extremely good properties with respect to the general difference operators ∂_μ and with respect to the inner product in $R(A)$.

(2.4) Proposition. *The algebra $\mathcal{E}_R(A)$ of operators on $R(A)$ is the full ring of endomorphisms of $R(A)$ as a module over the ring of symmetric functions* $\mathrm{Sym}_R(A)$. *Similarly, the subalgebra $\mathcal{E}_R[A]$ is the full ring of endomorphisms of $R[A]$ as a module over the ring of symmetric polynomials* $\mathrm{Sym}_R[A]$. *Moreover, the general difference operators ∂_μ for μ in $\mathfrak{S}(A)$ form a basis of $\mathcal{E}_R(A)$ as a module over $R(A)$, and a basis for $\mathcal{E}_R[A]$ as a module over $R[A]$. Furthermore, if f is a rational function in $R(A)$ such that the inner product $\langle g, f\rangle$ is a polynomial for every polynomial g in $R[A]$, then f is a polynomial in $R[A]$. Finally, the subring $\mathcal{E}_R[A]$ of $\mathcal{E}_R(A)$ is invariant under the canonical involution of* (DIFF.2.1).

Proof. Let M denote the algebra of rational functions $R(A)$ as a module over the ring $S := \mathrm{Sym}_R(A)$. The Schubert polynomials X_μ form an S-basis for M. Hence, if $\{\check{X}_\mu\}$ denotes the dual basis for the module of S-linear forms on M, every S-linear endomorphism of M is of the form $f \mapsto \sum_\mu \check{X}_\mu(f) h_\mu$ for rational functions h_μ in M. Therefore, to prove the first assertion of the Proposition, it suffices to prove for every permutation μ and every rational function h that the map $f \mapsto h\check{X}_\mu(f)$ belongs to the ring $\mathcal{E}_R(A)$. Now, by the equation (SCHUB.1.7.2), for the dual basis $\check{X}_\mu$ we have the equation,

$$\check{X}_\mu(f) = \langle Y_\mu, f\rangle.$$

Moreover, by Definition (DIFF.4.1) of the inner product, $\langle Y_\mu, f\rangle = \delta^A(Y_\mu f)$. Hence the map $f \mapsto h\check{X}_\mu(f)$ is the operator $h\delta^A \cdot Y_\mu$ belonging to $\mathcal{E}_R(A)$. Thus the first assertion of the Proposition has been proved. The proof of the second assertion is entirely similar.

Consider the general difference operators ∂_μ. It was proved in Lemma (DIFF.3.10) that the ∂_μ form an $R(A)$-basis for $\mathcal{E}_R(A)$. Hence, every operator $\alpha \in \mathcal{E}_R(A)$ has a unique expansion,

$$\alpha = \sum_{\mu \in \mathfrak{S}(A)} \alpha_\mu \partial_\mu,$$

with uniquely determined coefficients α_μ in $R(A)$. If $\alpha \in \mathcal{E}_R[A]$, then the coefficients α_μ are polynomials. Indeed, to prove that $\alpha_\nu \in R[A]$ we may, by induction on $\ell(\nu)$, assume that $\alpha_\mu \in R[A]$ for all permutations μ with $\ell(\mu) < \ell(\nu)$. By subtracting from α the sum of the $\alpha_\mu \partial_\mu$ for $\ell(\mu) < \ell(\nu)$, we may assume that $\alpha_\mu = 0$ for $\ell(\mu) < \ell(\nu)$. By (SCHUB.1.6), if $\ell(\mu) \geq \ell(\nu)$ and $\mu \neq \nu$, then $\partial_\mu(X_\nu) = 0$; moreover, $\partial_\nu(X_\nu) = 1$. Hence $\alpha(X_\nu) = \alpha_\nu \partial_\nu(X_\nu) = \alpha_\nu$. Consequently, the coefficient α_ν is a polynomial. Thus the third assertion of the proposition holds.

Let f be a rational function in $R(A)$ such that the inner product $\langle g, f\rangle$ is a polynomial for every polynomial g in $R[A]$. Expand f in terms of the Schubert polynomials, $f = \sum_\mu f_\mu X_\mu$, where the coefficients f_μ are symmetric rational functions. Now, by Theorem (SCHUB.1.7), the coefficient f_μ is equal to the inner product $\langle Y_\mu, f\rangle$, and hence the coefficient is a polynomial. Therefore, the function f is a polynomial.

Finally, the last assertion of the Proposition follows from the expansion of an operator α in terms of the ∂_μ. Indeed, if $\alpha = \sum_\mu \alpha_\mu \partial_\mu$ where the α_μ are rational functions in $R(A)$, then $\alpha^* = \sum_\mu (\partial_\mu)^* \alpha_\mu$. By Lemma (DIFF.3.6), we have that $(\partial_\mu)^* = \partial_{\mu^{-1}}$. As proved above, if α belongs to $\mathcal{E}_R[A]$ then the functions α_μ are polynomials, and hence it follows from the expression for α^* that α^* belongs to $\mathcal{E}_R[A]$.

Hence all assertions of the Proposition have been proved. ∎

(2.5) The Cauchy Formula. *Assume that B and C are two sequences with n elements in the ground ring R. Then, for every permutation μ in $\mathfrak{S}(A)$ we have the equation of values,*

$$X_\mu^B(C) = (\operatorname{sign}\mu) X_{\mu^{-1}}^C(B). \tag{2.5.1}$$

Moreover, the expansion of X_μ^B in terms of the basis X_ν^C is given by the formula,

$$X_\mu^B = \sum_{\substack{\nu \in \mathfrak{S}(A)\\ \ell(\mu\nu^{-1})+\ell(\nu)=\ell(\mu)}} X_{\mu\nu^{-1}}^B(C) X_\nu^C. \tag{2.5.2}$$

Proof.

Consider Equation (2.5.1). By Lemma (1.2)(3), the left side is equal to $\langle Y^C, X_\mu^B\rangle$. On the right side, we have $X_{\mu^{-1}}^C = \partial_{\mu\omega}(X^C)$. Hence, by (1.7.3), the right side of (2.5.1) is equal to $(\operatorname{sign}\mu)\langle Y_{\mu\omega}^B, X^C\rangle$. Consequently, Equation (2.5.1) is equivalent to the following equation:

$$\langle Y^C, X_\mu^B\rangle = (\operatorname{sign}\mu)\langle Y_{\mu\omega}^B, X^C\rangle. \tag{2.5.3}$$

On the right side, $X^C = \omega(Y^C)$, and $(\operatorname{sign}\mu)Y^B_{\mu\omega} = (\operatorname{sign}\omega)\omega(X^B_\mu)$ by (1.3.3). In general, $\langle\omega(f), \omega(g)\rangle = (\operatorname{sign}\omega)\langle f, g\rangle$ by Lemma (DIFF.4.2), and $\langle f, g\rangle = \langle g, f\rangle$. Therefore the right side of Equation (2.5.3) is equal to the left side. Hence Equation (2.5.1) holds.

To prove (2.5.2), we apply Newton's interpolation formula (1.7.4) to the function $f := X^B_\mu$ and obtain the equation,

$$X^B_\mu = \sum_{\nu\in\mathfrak{S}(A)} \partial_\nu(X^B_\mu)(C)X^C_\nu.$$

It follows from (1.6.1) that the polynomial $\partial_\nu(X^B_\mu)$ is equal to zero unless the condition for the summation index in (2.5.2) is satisfied; moreover, if the latter condition is satisfied, then the polynomial is equal to $X^B_{\mu\nu^{-1}}$. Therefore Equation (2.5.2) holds. ▯

(2.6) Corollary. *Let B be a set of n elements in R. Then the following formula holds:*

$$X^B = \sum_{\mu\in\mathfrak{S}(A)} (\operatorname{sign}\mu)X_\mu(B)X_{\mu\omega}. \tag{2.6.1}$$

Proof. Apply the Cauchy formula (2.5.2) with $\mu := \omega$ and $C := (0, \dots, 0)$. In the sum on the right hand side, the condition $\ell(\omega\nu^{-1}) + \ell(\nu) = \ell(\omega)$ on the summation index ν is satisfied for all permutations ν, by (SYM.1.3). Hence we obtain the equation,

$$X^B_\omega = \sum_{\nu\in\mathfrak{S}(A)} X^B_{\omega\nu^{-1}}(0)X_\nu. \tag{2.6.2}$$

In the terms of the sum on the right hand side, let $\nu = \mu\omega$ and form the summation over μ. By Equation (2.5.1), we have that $X^B_{\mu^{-1}}(0) = (\operatorname{sign}\mu)X_\mu(B)$. Therefore, Equation (2.6.1) follows from Equation (2.6.2). ▯

(2.7) Example. Let $n = 3$. The right side of (2.6.1) is the sum,

$$\begin{aligned}
&X_\omega(A)X_1(B) - X_{\varepsilon_2\varepsilon_1}(A)X_{\varepsilon_1}(B) - X_{\varepsilon_1\varepsilon_2}X_{\varepsilon_2}(B)\\
&+ X_{\varepsilon_2}(A)X_{\varepsilon_1\varepsilon_2}(B) + X_{\varepsilon_1}(A)X_{\varepsilon_2\varepsilon_1}(B) - X_1(A)X_\omega(B)\\
&= a_1^2a_2 - a_1a_2(b_1+b_2) - a_1^2b_1 + a_1b_1^2 + (a_1+a_2)b_1b_2 - b_1^2b_2.
\end{aligned}$$

The formula is the reduction of this sum to

$$X(A, B) = (a_1 - b_1)(a_1 - b_2)(a_2 - b_1).$$

Partially symmetric functions

1. Partially symmetric functions.

(1.1) Setup. Assume in this chapter that the alphabet A is *partitioned*, that is, assume that A is the union of a set $\mathbf{A} = (A_1|\dots|A_r)$ of r disjoint alphabets A_i such that, for all $i < j$, if $a \in A_i$ and $b \in A_j$, then $a < b$. The A_i are often called the *intervals*. The partitioning may be given by listing, for $i = 1, \dots, r$, the size n_i of the i'th interval A_i, or by listing, for $i = 1, \dots, r-1$, the index p_i of the last letter in A_i:

$$A = \{\overbrace{a_1, \dots, a_{p_1}}^{A_1}, \overbrace{a_{p_1+1}, \dots, a_{p_2}}^{A_2}, \ \dots\ , \overbrace{a_{p_{r-1}+1}, \dots, a_n}^{A_r}\},$$

With $p_0 := 0$ and $p_r := n$, we have $|A_i| = n_i = p_i - p_{i-1}$.

The *Young subgroup* $\mathfrak{S}(\mathbf{A}) = \mathfrak{S}(A_1|\dots|A_r)$ corresponding to $\mathbf{A}$ is the subgroup of permutations in $\mathfrak{S}(A)$ that leave every A_i invariant. The group $\mathfrak{S}(A_i)$ will be identified with the subgroup of $\mathfrak{S}(\mathbf{A})$ consisting of the permutations that are equal to the identity on letters outside A_i. Clearly, any permutation μ in $\mathfrak{S}(\mathbf{A})$ is a product $\mu = \mu_1 \cdots \mu_r$ with uniquely determined factors $\mu_i \in \mathfrak{S}(A_i)$. Moreover, as different factors μ_i commute, it follows that the Young subgroup is the product,

$$\mathfrak{S}(\mathbf{A}) = \mathfrak{S}(A_1) \times \cdots \times \mathfrak{S}(A_r).$$

Obviously, the Young subgroup $\mathfrak{S}(\mathbf{A})$ is generated by the simple transpositions ε_p for p different from the p_i. In fact, if $\mu = \mu_1 \dots \mu_r$ with $\mu_i \in \mathfrak{S}(A_i)$ then the length of μ is the sum of the lengths of the μ_i, and a (minimal) presentation of μ is obtained by concatenating (minimal) presentations of the μ_i.

Denote by $\mathfrak{T}(\mathbf{A})$ the subset of $\mathfrak{S}(A)$ consisting of the permutations that are increasing on every interval A_i. The subset $\mathfrak{T}(\mathbf{A})$ is a system of representatives for the set $\mathfrak{S}(A)/\mathfrak{S}(\mathbf{A})$ of left cosets modulo the Young subgroup since, obviously, for every permutation μ there are unique permutations μ_i in $\mathfrak{S}(A_i)$ such that $\mu\mu_1 \cdots \mu_r$ is increasing on each interval A_i. By a simple count of inversions, if $\sigma \in \mathfrak{T}(\mathbf{A})$ and $\nu \in \mathfrak{S}(\mathbf{A})$, then

$$\ell(\sigma\nu) = \ell(\sigma) + \ell(\nu). \tag{1.1.1}$$

Note that the simple transpositions ε_{p_i} belong to $\mathfrak{T}(\mathbf{A})$.

Clearly, the Young subgroup $\mathfrak{S}(\mathbf{A})$ has order $n_1! \cdots n_r!$, where n_i is the number of letters in the interval A_i. It follows that the subset $\mathfrak{T}(\mathbf{A})$ is of order equal to the multinomial coefficient,

$$\binom{n}{n_1, \dots, n_r} = \frac{n!}{n_1! \cdots n_r!}.$$

Denote by $\omega = \omega_A$ the order reversing permutation of A and by $\omega_i = \omega_{A_i}$ the order reversing permutation of A_i. It is obvious that a permutation σ in $\mathfrak{S}(A)$ belongs to $\mathfrak{T}(\mathbf{A})$ if and only if $\omega\sigma\omega_1 \cdots \omega_r$ belongs to $\mathfrak{T}(\mathbf{A})$. Clearly, the permutation $\omega_{\mathbf{A}}$ defined by the equation,

$$\omega = \omega_{\mathbf{A}}\omega_1 \cdots \omega_r,$$

or equivalently, $\omega_{\mathbf{A}} := \omega\omega_1 \cdots \omega_r$, belongs to $\mathfrak{T}(\mathbf{A})$. The map ω defines an order reversing bijection $A_i \to \omega A_i$ and $\omega_{\mathbf{A}}$ defines an order preserving bijection $A_i \to \omega A_i$.

In connection with difference operators we shall often use the equations,

$$\partial^{\sigma\nu} = \partial^{\sigma}\partial^{\nu}, \quad \partial^{\nu\sigma^{-1}} = \partial^{\nu}\partial^{\sigma^{-1}} \text{ for } \sigma \in \mathfrak{T}(\mathbf{A}),\ \nu \in \mathfrak{S}(\mathbf{A}), \tag{1.1.2}$$

and similar equations for ∂_μ, π^μ, and ψ^μ. The first equation follows from (1.1.1) and (DIFF.3.7), the second similarly noting that (1.1.1) remains valid for the inverses. For instance, we have

$$\partial^{\omega} = \partial^{\omega_{\mathbf{A}}}\partial^{\omega_1\cdots\omega_r} \qquad \text{and} \qquad \partial^{\omega_1\cdots\omega_r} = \partial^{\omega_1} \cdots \partial^{\omega_r}. \tag{1.1.3}$$

Note that there is a second partitioning $\omega\mathbf{A} := (\omega A_r| \dots |\omega A_1)$ of A, and that $\mathfrak{S}(\omega\mathbf{A}) = \omega\mathfrak{S}(\mathbf{A})\omega^{-1}$. Moreover, a permutation σ in $\mathfrak{S}(A)$ belongs to $\mathfrak{T}(\omega\mathbf{A})$ if and only if $\omega\sigma\omega$ belongs to $\mathfrak{T}(\mathbf{A})$.

(1.2) Definition. A rational function of $R(A)$ invariant under the Young subgroup $\mathfrak{S}(\mathbf{A})$ will be called *partially symmetric*. We denote by $\mathrm{Sym}_R(\mathbf{A})$ and $\mathrm{Sym}_R[\mathbf{A}]$ the subalgebras of $R(A)$ of functions and polynomials respectively that are partially symmetric.

Clearly, a function f is partially symmetric if and only if $\varepsilon_p(f) = f$ for all p different from the p_i.

(1.3) Lemma. *A function f of $R(A)$ is partially symmetric if and only if $\partial_p(f) = 0$ for all p different from the p_i. In addition, if f is partially symmetric, then $\partial_\mu(f) = 0$ for all permutations μ outside the subset $\mathfrak{T}(\mathbf{A})$.*

Proof. Since $(a_p - a_{p+1})\partial_p(f) = f - \varepsilon_p(f)$, we have that $\partial_p(f) = 0$ if and only if $\varepsilon_p(f) = f$. The first assertion is a consequence, because the Young group $\mathfrak{S}(\mathbf{A})$ is generated by the transpositions ε_p for p different from the p_i.

To prove the second assertion, assume that f is partially symmetric. Then clearly $\partial_\nu(f) = 0$ for every $\nu \neq 1$ in $\mathfrak{S}(\mathbf{A})$. Consider for any permutation μ outside $\mathfrak{T}(\mathbf{A})$ the factorization $\mu = \sigma\nu$ with $\sigma \in \mathfrak{T}(\mathbf{A})$, $\nu \in \mathfrak{S}(\mathbf{A})$. Then $\nu \neq 1$. Whence, using (1.1.2), $\partial_\mu(f) = \partial_\sigma\partial_\nu(f) = 0$. ∎

(1.4) Lemma. *Let σ be a permutation in $\mathfrak{S}(A)$. Then, for any of the three polynomials X_σ, $Y_{\omega\sigma}$, and $Y_{\sigma\omega_1\cdots\omega_r}$ defined in* (SCHUB.2.1), *we have that the polynomial is partially symmetric if and only if σ belongs to $\mathfrak{T}(\mathbf{A})$.*

Proof. We use Equation (SCHUB.1.6.1) of Lemma (SCHUB.1.6). Consider first the Schubert polynomial X_σ. By (SCHUB.1.6.1) we have that $\partial_\sigma(X_\sigma) = 1$. Hence it follows from Lemma (1.3) that if X_σ is partially symmetric, then σ belongs to $\mathfrak{T}(\mathbf{A})$. Assume conversely that σ belongs to $\mathfrak{T}(\mathbf{A})$. To prove that X_σ is partially symmetric, we have to show that $\partial_p X_\sigma = 0$ for p different from the p_i. The polynomial $\partial_p X_\sigma = \partial_{\varepsilon_p} X_\sigma$ is determined by Lemma (SCHUB.1.6). If p is different from the p_i, then, by (1.1.1), $\ell(\sigma\varepsilon_p) = \ell(\sigma) + \ell(\varepsilon_p)$. In particular, $\ell(\sigma\varepsilon_p^{-1}) \neq \ell(\sigma) - \ell(\varepsilon_p)$. Therefore, it follows (SCHUR.1.6.1) that $\partial_p(X_\sigma) = 0$. Hence X_σ is partially symmetric.

By (SCHUB.1.3.3), we have that $Y_{\omega\sigma} = \mathrm{sign}(\omega\sigma)\omega(X_{\omega\sigma\omega})$. Hence $Y_{\omega\sigma}$ is partially symmetric if and only if $\omega(X_{\omega\sigma\omega})$ is partially symmetric. Moreover, $\omega(X_{\omega\sigma\omega})$ is partially symmetric if and only if $X_{\omega\sigma\omega}$ is invariant under the conjugate Young group $\omega\mathfrak{S}(\mathbf{A})\omega^{-1} = \mathfrak{S}(\omega\mathbf{A})$. So, by the assertion for the Schubert polynomials X_σ, we have that $Y_{\omega\sigma}$ is partially symmetric, if and only if $\omega\sigma\omega \in \mathfrak{T}(\omega\mathbf{A})$, that is, if and only if $\sigma \in \mathfrak{T}(\mathbf{A})$.

Finally, the assertion about the polynomials $Y_{\sigma\omega_1\cdots\omega_r}$ follows from the assertion about the polynomials $Y_{\omega\sigma}$, since σ belongs to $\mathfrak{T}(\mathbf{A})$ if and only if $\omega\sigma\omega_1\cdots\omega_r$ belongs to $\mathfrak{T}(\mathbf{A})$. □

(1.5) Proposition. *The Schubert polynomials X_σ for $\sigma \in \mathfrak{T}(\mathbf{A})$ form a basis for the algebra* $\mathrm{Sym}_R(\mathbf{A})$ *of partially symmetric functions as a module of the ring* $\mathrm{Sym}_R(A)$ *of symmetric functions, and a basis for the algebra* $\mathrm{Sym}_R[\mathbf{A}]$ *of partially symmetric polynomials as a module over the ring* $\mathrm{Sym}_R[A]$ *of symmetric polynomials.*

Proof. By Theorem (SCHUB.1.7), the Schubert polynomials X_μ for μ in $\mathfrak{S}(A)$ form a $\mathrm{Sym}_R(A)$-basis for $R(A)$. Moreover, in the expansion of a function f in the basis the coefficient to X_μ is equal to the inner product,

$$\langle Y, \partial_\mu(f)\rangle\,. \tag{1.5.1}$$

The polynomials X_σ for σ in $\mathfrak{T}(\mathbf{A})$ are partially symmetric by Lemma (1.4). Therefore, to prove the assertion of the Proposition, it suffices to prove that if f is partially symmetric and μ does not belong to $\mathfrak{T}(\mathbf{A})$, then the inner product (1.5.1) vanishes. The latter assertion follows immediately from the second assertion of Lemma (1.3). □

2. Partial symmetrization.

(2.1) Definition. Keep the setup of (1.1). Consider the following special polynomial:

$$\Delta(\mathbf{A}) := \prod_{i>j} \prod_{b\in A_i,\, a\in A_j} (b-a).$$

Clearly, the polynomial $\Delta(\mathbf{A})$ is partially symmetric. Moreover, we have the following factorization of the Vandermonde determinant:

$$\Delta(A) = \Delta(\mathbf{A})\Delta(A_1)\cdots\Delta(A_r). \tag{2.1.1}$$

Define the *partial symmetrization operator* $\delta^{\mathbf{A}}$ as the following sum,

$$\delta^{\mathbf{A}} := \sum_{\sigma\in\mathfrak{T}(\mathbf{A})} \sigma\cdot\frac{1}{\Delta(\mathbf{A})}\,. \tag{2.1.2}$$

By definition, we have that $\delta^{\mathbf{A}}$ is an operator, and as such it is defined on functions f of $R(A)$, and linear with respect to symmetric functions. Most often, we will consider values of $\delta^{\mathbf{A}}$ on partially symmetric functions f. Assume that f is partially symmetric. Then, clearly, to obtain the value $\delta^{\mathbf{A}}(f)$, the summation over $\mathfrak{T}(\mathbf{A})$ in (2.1.2) can be replaced by the summation over any set of representatives for the set of left cosets $\mathfrak{S}(A)/\mathfrak{S}(\mathbf{A})$. In particular, if μ is a given permutation in $\mathfrak{S}(A)$, we may replace σ in the terms of (2.1.2) by $\mu\sigma$. It follows that the value $\delta^{\mathbf{A}}(f)$ is a symmetric function. Hence the partial symmetrization operator may be viewed as a $\mathrm{Sym}_R(A)$-linear map,

$$\delta^{\mathbf{A}}\colon\ \mathrm{Sym}_R(\mathbf{A}) \to \mathrm{Sym}_R(A).$$

(2.2) Proposition. *The total symmetrization operator $\delta^A = \partial^\omega$ satisfies the equations,*

$$\partial^\omega = \partial^{\omega_{\mathbf{A}}}\partial^{\omega_1}\cdots\partial^{\omega_r}, \qquad \delta^A = \delta^{\mathbf{A}}\delta^{A_1}\cdots\delta^{A_r}. \tag{2.2.1}$$

Proof. The equation $\delta^A = \partial^\omega$ is the fundamental equation in (DIFF.3.3.3), and the first equation in (2.2.1) follows from (1.1.3).

The second equation follows from a direct computation. Recall that δ^A is the total symmetrization operator, defined in (DIFF.2.3) as the sum $\sum_\mu \mu\cdot(1/\Delta(A))$ where the sum is over all permutations μ in $\mathfrak{S}(A)$, and δ^{A_i} is defined similarly using $\Delta(A_i)$ and permutations in $\mathfrak{S}(A_i)$.

In the summation defining δ^A, each permutation μ is uniquely a product $\mu = \sigma\mu_1\cdots\mu_r$ where σ belongs to $\mathfrak{T}(\mathbf{A})$ and μ_i belongs to $\mathfrak{S}(A_i)$. From the factorization (2.1.1) we obtain the equation of operators,

$$\mu\cdot\frac{1}{\Delta(A)} = \sigma\mu_1\cdots\mu_r\cdot\frac{1}{\Delta(\mathbf{A})}\,\frac{1}{\Delta(A_1)}\,\cdots\,\frac{1}{\Delta(A_r)} = \sigma\cdot\frac{1}{\Delta(\mathbf{A})}\mu_1\cdot\frac{1}{\Delta(A_1)}\cdots\mu_r\cdot\frac{1}{\Delta(A_r)}\,,$$

because $\Delta(\mathbf{A})$ and $\Delta(A_j)$ for $j\neq i$ are symmetric with respect to the letters of A_i. Therefore, by forming the sum over μ we obtain, using the definitions of $\delta^{\mathbf{A}}$ and δ^{A_i} the second equation in (2.2.1). ∎

(2.3) Corollary. *Assume that $g_i \in R(A_i)$ for $i = 1, \dots, r$ depends only on the letters in A_i, and let f be a partially symmetric function. Then we have the equations,*

$$\partial^{\omega_1 \cdots \omega_r}(f\, g_1 \cdots g_r) = f\, \partial^{\omega_1}(g_1) \cdots \partial^{\omega_r}(g_r)\,, \tag{2.3.1}$$

$$\delta^A(f g_1 \cdots g_r) = \delta^{\mathbf{A}}\big(f\, \delta^{A_1}(g_1) \cdots \delta^{A_r}(g_r)\big)\,. \tag{2.3.2}$$

In particular, let n_i be the cardinality of A_i, and let a^{E_i} (with an abuse of notation) be the monomial in $R[A]$ defined as the product of the letters of A_i raised to the powers $0, 1, \dots, n_i - 1$. Then, on the subalgebra of partially symmetric functions, the operator $\partial^{\omega_1 \cdots \omega_r} \cdot a^{E_1} \cdots a^{E_r}$ is the identity and the following three operators are equal:

$$\delta^A \cdot a^{E_1} \cdots a^{E_r}, \qquad \delta^{\mathbf{A}}, \qquad \partial^{\omega_{\mathbf{A}}}. \tag{2.3.3}$$

Finally, on the subalgebra, the values taken by the three operators are symmetric functions, and the values taken on partially symmetric polynomials are symmetric polynomials.

Proof. To verify (2.3.1), use that $\partial^{\omega_1 \cdots \omega_r} = \partial^{\omega_1} \cdots \partial^{\omega_r}$ and that $\partial^{\omega_i} = \delta^{A_i}$ is linear with respect to functions that are symmetric in the letters of A_i. So, when ∂^{ω_i} is applied to the product $f g_1 \cdots g_r$, all the factors except g_r are scalars. Hence the result is the product $f g_1 \cdots g_{r-1} \partial^{\omega_r}(g_r)$. Again, in the latter product, all the factors except g_{r-1} are scalars with respect to $\partial^{\omega_{r-1}}$, and so only g_{r-1} is changed when $\partial^{\omega_{r-1}}$ is applied. Thus (2.3.1) is obtained by applying successively $\partial^{\omega_r}, \dots, \partial^{\omega_1}$ to the product $f g_1 \cdots g_r$.

Next, apply $\delta^{\mathbf{A}}$ to (2.3.1). It follows from the second equation in (2.2.1) that the result is (2.3.2).

In particular, to obtain the value of the operator $\partial^{\omega_1 \cdots \omega_r} \cdot a^{E_1} \cdots a^{E_r}$ on f set $g_i = a^{E_i}$ in (2.3.1). As $\partial^{\omega_i}(a^{E_i}) = 1$ by (DIFF.3.3), the result is the equation,

$$\partial^{\omega_1 \cdots \omega_r}\big(a^{E_1} \cdots a^{E_r} f\big) = f\,.$$

From the latter equation and (2.2.1) it follows that the three operators in (2.3.3) have the same value on f.

Finally, the assertion about the values holds for the first operator in (2.3.3), since it holds for δ^A. Hence the assertion holds for all three operators. ∎

(2.4) Definition. Define for partially symmetric functions f and g their *partial inner product* as the sum:

$$\langle f, g\rangle^{\mathbf{A}} := \sum_{\sigma \in \mathfrak{T}(\mathbf{A})} \sigma\Big(\frac{fg}{\Delta(\mathbf{A})}\Big) = \delta^{\mathbf{A}}(fg). \tag{2.4.1}$$

Although the right hand side is defined for all rational functions f and g in $R(A)$, we will always assume that f and g are partially symmetric. It follows from Corollary (2.3) that

$$\langle f, g\rangle^{\mathbf{A}} = \partial^{\omega_{\mathbf{A}}}(fg). \tag{2.4.2}$$

(2.5) Lemma. *The values of the partial inner product* (2.4) *are symmetric functions. Moreover, the inner product is symmetric, and* $\mathrm{Sym}_R(A)$*-bilinear. Finally, if f and g are partially symmetric polynomials, then their inner product* $\langle f, g\rangle^{\mathbf{A}}$ *is a polynomial.*

Proof. The assertions follow immediately from the last sentence of Corollary (2.3). ∎

(2.6) Proposition. *Let f be a partially symmetric function. Then, in the expansion of f in the basis of Schubert polynomials X_σ for* $\sigma \in \mathfrak{T}(\mathbf{A})$ (see Proposition (1.5)), *the coefficient to X_σ is given by the formula,*

$$\langle Y_\sigma, f\rangle = \mathrm{sign}(\omega_1 \cdots \omega_r)\, \langle Y_{\sigma\omega_1\cdots\omega_r}, f\rangle^{\mathbf{A}}. \tag{2.6.1}$$

In particular, if σ and τ are permutations in $\mathfrak{T}(\mathbf{A})$, *then*

$$\mathrm{sign}(\omega_1 \cdots \omega_r)\langle Y_{\sigma\omega_1\cdots\omega_r}, X_\tau\rangle^{\mathbf{A}} = \begin{cases} 1 & \textit{if } \sigma = \tau, \\ 0 & \textit{otherwise.} \end{cases} \tag{2.6.2}$$

Proof. By Theorem (SCHUB.1.7), the coefficient is indeed the left side of Equation (2.6.1). To prove the equation, recall that $Y_\mu = \partial_{\mu^{-1}} Y$. For $\mu = \sigma\omega_1 \cdots \omega_r$, it follows from (1.1.2) that $\partial_{\mu^{-1}} = \partial_{\omega_1\cdots\omega_r}\partial_{\sigma^{-1}}$. Thus

$$\partial^{\omega_1\cdots\omega_r}(Y_\sigma) = \mathrm{sign}(\omega_1 \cdots \omega_r)\partial_{\omega_1\cdots\omega_r}\partial_{\sigma^{-1}}(Y) = \mathrm{sign}(\omega_1 \cdots \omega_r)Y_{\sigma\omega_1\cdots\omega_r}. \tag{2.6.3}$$

Multiply by f in (2.6.3) and apply $\delta^{\mathbf{A}}$. The operator $\partial^{\omega_1\cdots\omega_r}$ commutes with multiplication by f since f is partially symmetric. So the result is the equation,

$$\delta^{\mathbf{A}}\partial^{\omega_1\cdots\omega_r}(Y_\sigma f) = \mathrm{sign}(\omega_1 \cdots \omega_r)\delta^{\mathbf{A}}(Y_{\sigma\omega_1\cdots\omega_r} f). \tag{2.6.4}$$

On the left side we have $\delta^{\mathbf{A}}\partial^{\omega_1\cdots\omega_r} = \delta^A$ by Proposition (2.2). Hence, by definition of the inner products, Equation (2.6.4) is the required equation.

Clearly, the second assertion of the Proposition is a particular case of the first. ∎

(2.7) Note. The total symmetrization operator δ^A vanishes on the partially symmetric functions except for the trivial partitioning $\mathbf{A}$ where each interval A_i consists of a single letter. Indeed, assume that f is partially symmetric. Take $g_i = 1$ in (2.3.2) to obtain the equation,

$$\delta^A(f) = \delta^{\mathbf{A}}\big(f\,\delta^{A_1}(1)\cdots\delta^{A_r}(1)\big).$$

The operator δ^{A_i} lovers the degree by $\ell(\omega_i)$. Hence $\delta^{A_i}(1) = 0$ unless A_i consists of a single letter. Thus $\delta^A(f) = 0$ unless the partitioning is trivial.

Assume that the partitioning $\mathbf{A}$ is non-trivial. It follows in particular that the total inner product of (DIFF.4.1) vanishes identically on the module of partially symmetric functions. Note also that the polynomial Y_σ in (2.6.2) is not partially symmetric. Indeed, by Lemma (1.4), the polynomial Y_μ is partially symmetric if and only if $\omega\mu$ is increasing on each interval, and if $\sigma \in \mathfrak{T}(\mathbf{A})$, then $\omega\sigma$ is decreasing on each A_i.

(2.8) Note. *Lagrange interpolation in one variable.* Let x be an additional letter. Consider the alphabet $\{a_1, \dots, a_n, x\}$ and the partitioning $(A|x)$. Then $\Delta(A|x) = \prod_a (x - a)$ is a polynomial of degree n, and the partial symmetrization operator lowers the degree by n.

Let f be a polynomial depending only on x, and of degree strictly less than n. Since $\delta^{A|x}$ lowers the degree by n, it follows that partial symmetrization of f yields zero, that is, we obtain the following equation:

$$\sum_{\sigma \in \mathfrak{T}(A|x)} \sigma\Big(\frac{f}{\Delta(A|x)}\Big) = 0.$$

Consider the terms in the sum. The permutations σ in $\mathfrak{T}(A|x)$ are determined by the value $\sigma(x)$. If $\sigma(x) = x$, then $\sigma = 1$ and the corresponding term in the sum is the rational function $f/\Delta(A|x)$. If $\sigma(x) = b$ is a letter of A then, in the corresponding term, the numerator is $f(b)$ and the denominator is given by

$$\sigma(\Delta(A|x)) = (b - x) \prod_{a \neq b} (b - a)\,.$$

Therefore, by separating in the equation the terms corresponding to $\sigma = 1$ and $\sigma \neq 1$ and multiplying by $\Delta(A|x)$ we obtain the equation in $R(A)[x]$,

$$f = \sum_{b \in A} f(b) \prod_{a \neq b} \frac{(x - a)}{(b - a)}\,, \tag{2.8.1}$$

which is the usual form of Lagrange interpolation.

Note that (2.8.1) holds if the n letters a_i of A are replaced by any sequence of n elements α_i of R such that the differences $\alpha_p - \alpha_q$ for $p \neq q$ are invertible in R. In particular, replacing R by $R(A)$ we may apply the formula to the difference $f := \prod_a (x - ta) - \prod_a (x - a)$ for $t \in R$. Then $f(b) = (1 - t)b \prod_{a \neq b} (b - ta)$ and we obtain the formula in $R(A)(x)$,

$$\prod_a \frac{x - ta}{x - a} = 1 + (1 - t) \sum_b \frac{b}{x - b} \prod_{a \neq b} \frac{b - ta}{b - a}\,. \tag{2.8.2}$$

3. The Gysin formula.

(3.1) Setup. In the setup of (1.1), with the partitioned alphabet $\mathbf{A} = (A_1|\dots|A_r)$, the ring of partially symmetric polynomials $\mathrm{Sym}_R[\mathbf{A}] = \mathrm{Sym}_R[A_1|\dots|A_r]$ fits into a chain of subrings of $R[A]$:

$$R = R_0 \subseteq R_1 \subseteq \cdots \subseteq R_r = \mathrm{Sym}_R[\mathbf{A}],$$

where R_i is the ring $\mathrm{Sym}_R[A_1|\dots|A_i]$ of partially symmetric polynomials in the letters of $A_1 \cup \cdots \cup A_i$. Inductively, $R_i = \mathrm{Sym}_{R_{i-1}}[A_i]$ for $i = 1, \dots, r$ is the ring of symmetric polynomials in the polynomial ring $R_{i-1}[A_i]$.

Assume there is given, for each $i = 1, \dots, r$, an R-basis $(f_{i,j})_{j\in J_i}$ for $\mathrm{Sym}_R[A_i]$ (as an R-module). Then $(f_{i,j})_{j\in J_i}$ is also an R_{i-1}-basis for R_i (see Remark (SYM.8.14)). Consequently, the set of all products $f_{1,j_1}\cdots f_{r,j_r}$ is an R-basis for the module $\mathrm{Sym}_R[\mathbf{A}]$ of partially symmetric polynomials. For instance, the set of all products of monomial symmetric polynomials,

$$m^{K_1}(A_1)\cdots m^{K_r}(A_r),$$

where K_i is a weakly decreasing multi index of size n_i equal to the number of letters in A_i, is a basis. Similarly, the set of all products of Schur polynomials,

$$s^{J_1}(A_1)\cdots s^{J_r}(A_r),$$

where J_i is a strictly increasing multi index of size n_i, is a basis.

(3.2) Example. Consider a partitioning $\mathbf{A} = (A_1|A_2)$ into two subalphabets. Then we have the equation,

$$\prod_{a\in A_1, b\in A_2} (a+b) = \sum_{J\subseteq[n]} s^J(A_1)s^{\tilde{J}}(A_2),$$

where the sum is over subsets J of size n_1, and $\tilde{J} \subseteq [n]$ is the complement of J. More generally, for $d \geq 0$, we have the equation,

$$\prod_{b\in A_2} b^d \prod_{a\in A_1, b\in A_2} (a+b) = \sum_{J\subseteq[n]} s^J(A_1)s^{\hat{J}^{\sim}}(A_2), \tag{3.2.1}$$

where the sum is over subsets J of size n_1, $\hat{J}$ is the extension (see (SYM.6.13)) of J to a subset of size $n_1 + d$ of $[n+d]$, and $\hat{J}^{\sim} \subseteq [n+d]$ is the complement of $\hat{J}$.

Indeed, the first equation is the equation (SYM.8.3.3) with $A := A_1$ and $B := A_2$. To prove (3.2.1), extend A_1 to an alphabet $\hat{A}_1 = \{a_1, \dots, a_{n_1}, \hat{a}_1, \dots, \hat{a}_d\}$ with $n_1 + d$ letters. Apply the first equation to the partitioning $(\hat{A}_1|A_2)$ and substitute $\hat{a}_i := 0$ for $i = 1, \dots, d$. It follows that the left side of (3.2.1) is equal to the sum,

$$\sum_{I\subseteq[n+d]} s^I(A_1, 0, \dots, 0)s^{\tilde{I}}(A_2),$$

over subsets I of size n_1+d. Clearly, using the determinantal description of s^I from Section (SYM.7), we have $s^I(A_1,0,\dots,0)=s^I(A_1)$ since the corresponding equation holds for the complete symmetric functions s_i. Moreover, by the last sentence in Corollary (SYM.7.8)(1), if I is an extension of a size-n_1 multi index J, that is, $I=\hat{J}$, then $s^I(A_1)=s^J(A_1)$, and if I is not an extension, then $s^I(A_1)=0$. Hence Equation (3.2.1) holds.

(3.3) Bott's Formula. *For $i=1,\dots,r$, let J_i be a strictly increasing multi index of size equal to the number of letters of A_i. Then the following equation holds:*

$$\delta^{\mathbf{A}}\Big(s^{J_1}(A_1)\cdots s^{J_r}(A_r)\Big)=s^{J_1\dots J_r}(A), \tag{3.3.1}$$

where $J_1\dots J_r$ denotes the concatenated multi index.

Proof. With an abuse of notation, denote by a^{J_i} the monomial of $R[A]$ where the q'th letter of A_i appears with the exponent given by the q'th entry in J_i and all other letters appear with exponent zero.

In this notation, we have the equations,

$$s^{J_1}(A_1)\cdots s^{J_r}(A_r)=\delta^{A_1}(a^{J_1})\cdots\delta^{A_r}(a^{J_r})=\delta^{A_1}\cdots\delta^{A_r}(a^{J_1\cdots J_r}). \tag{3.3.2}$$

Indeed, by the definition in (SYM.6.9.2) or by Jacobi–Trudi's Formula (SYM.7.8.5) or (SCHUR.1.11), we have the equation $s^J(A)=\delta^A(a^J)$ for the Schur polynomial. Hence the first equation in (3.3.2) holds. The polynomials a^{J_i} and $\delta^{A_i}(a^{J_i})$ depend only on the letters of A_i. Hence they are scalars with respect to the operator δ^{A_j} for $j\neq i$. Moreover, $a^{J_1\dots J_r}=a^{J_1}\cdots a^{J_r}$. Hence the second equation in (3.3.2) holds.

By Proposition (2.2), the equation (3.3.1) follows by applying the operator $\delta^{\mathbf{A}}$ to the equation (3.3.2). ∎

(3.4) Setup. Assume that the number r of subalphabets is equal to 2. For convenience, set $B:=A_1$ and $C:=A_2$. Let m and k be the number of letters in B and C, so that $n=m+k$ is the number of letters of A. By definition of the partial inner product, the formula (3.3.1) is equivalent to the following *Gysin formula:*

$$\langle s^I(B),s^J(C)\rangle^{\mathbf{A}}=s^{IJ}(A), \tag{3.4.1}$$

where I and J are strictly increasing multi indices of sizes m and k.

(3.5) Corollary. *Consider the set of Schubert polynomials X_σ and the set of polynomials $Y_{\sigma\omega_1\omega_2}$ for $\sigma\in\mathfrak{T}(B|C)$. In addition, consider the set of Schur polynomials $s^I(B)$ for all strictly increasing multi indices I of size m with entries in the interval $[n]=[m+k]$, and the set of Schur polynomials $s^J(C)$ for all strictly increasing multi indices J of size k with entries in the same interval. Then any of the four sets is a basis for the algebra $\operatorname{Sym}_R[B|C]$ of partially symmetric polynomials as a module over its subring $\operatorname{Sym}_R[A]$ of symmetric polynomials. Moreover, for Schur polynomials of the two bases we have that*

$$\langle s^I(B),s^J(C)\rangle^{\mathbf{A}}=\operatorname{sign}(IJ), \tag{3.5.1}$$

where the right hand side is equal to the signature of IJ when the concatenated multi index IJ is a permutation of $[m+k]$ and equal to zero otherwise.

Finally, for any partially symmetric polynomial f, the expansion of f in the basis $s^I(B)$ is given by the formula,

$$f = \sum_I \operatorname{sign}(I\tilde{I})\delta^{\mathbf{A}}\big(fs^{\tilde{I}}(C)\big)\, s^I(B),$$

where $\tilde{I}$ denotes the complementary sequence of I with respect to the interval $[m+k]$. Similarly, the expansion of f in the basis $s^J(C)$ is given by the formula,

$$f = \sum_J \operatorname{sign}(\tilde{J}J)\delta^{\mathbf{A}}\big(fs^{\tilde{J}}(B)\big)\, s^J(C).$$

Proof. The equation (3.5.1) follows from (3.4.1) because the Schur polynomial $s^K(A)$ is alternating in K and equal to 1 when K is the sequence $(0, 1, \dots, n-1)$.

Clearly, the four sets have the same number d of elements, namely $d = \binom{n}{m}$. That the X_σ form a basis was proved in (1.5). The remaining assertions of the Corollary follow, by the standard argument used in the proof of Theorem (SCHUB.1.7), from the equations (2.6.2) and (3.5.1). Indeed, since the X_σ form a basis with d elements for the module $M := \operatorname{Sym}_R[B|C]$ as the module over S, it follows from (3.5.1) that the d polynomials $s^I(B)$ form an S-basis for M and the dual basis is given by the d linear forms,

$$f \mapsto \operatorname{sign}(I\tilde{I})\langle f, s^{\tilde{I}}(C)\rangle^{\mathbf{A}}.$$

By the same argument, the $s^J(C)$ form a basis, and by (2.6.2) the $Y_{\sigma\omega_1\omega_2}$ form a basis.

Since $\langle f, g\rangle^{\mathbf{A}} = \delta^{\mathbf{A}}(fg)$, the first expansion of f given in the Corollary follows from the description of the basis dual to the $s^I(B)$. The proof of the second expansion is entirely similar. ∎

4. Hall–Littlewood polynomials.

(4.1) Setup. Fix an element $t \in R$. Consider the following polynomial of $R[A]$:

$$\Delta_t(A) := \prod_{a<b}(b - ta),$$

where the product is over letters a and b of A. Clearly, $\Delta_t(A)$ has the same degree, $n(n-1)/2$, as the Vandermonde determinant $\Delta(A)$. For $t = 1$, we have that $\Delta_1(A) = \Delta(A)$. For $t = 0$, we have that $\Delta_0(A) = a^E$, where $E = (0, 1, \dots, n-1)$.

If a partitioning $\mathbf{A} = (A_1|\dots|A_r)$ is given, we write $a \ll b$ if $a < b$ and a and b belong to different intervals. In this notation, let

$$\Delta_t(\mathbf{A}) := \prod_{a \ll b}(b - ta).$$

Clearly, $\Delta_t(\mathbf{A})$ is a homogeneous partially symmetric polynomial. Moreover, we have the factorization,

$$\Delta_t(A) = \Delta_t(\mathbf{A})\Delta_t(A_1)\cdots\Delta_t(A_r). \tag{4.1.1}$$

For $t = 1$, the factorization is the factorization (2.1.1) of the Vandermonde determinant. For $t = 0$, we have in the notion of (2.3) that $\Delta_0(\mathbf{A})$ is the monomial determined by the equation,

$$a^E = \Delta_0(\mathbf{A})a^{E_1}\cdots a^{E_r}. \tag{4.1.2}$$

(4.2) Example. Consider the partitioning $(B\,|\,x)$ where $B = \{b_1, \dots, b_m\}$ is the interval consisting of the first $m = n-1$ letters of A and $x = a_n$ is the last letter of A. By definition, we have that $\Delta_t(B\,|\,x) = \prod_{q=1}^m (x - tb_q)$. Moreover, we have the equation,

$$\delta^{B|x}\Big(\Delta_t(B\,|\,x)\Big) = \sum_{j=0}^m t^j = \frac{t^n - 1}{t - 1}. \tag{4.2.1}$$

Indeed, by expanding the product $\prod_{q=1}^m (x - tb_q)$, we obtain the equation,

$$\Delta_t(B\,|\,x) = \sum_{j=0}^m (-1)^{m-j} t^{m-j} e_{m-j}(B)x^j.$$

By (SYM.6.8.1), we have that $e_{m-j}(B) = s^{0,\dots,\hat{j},\dots,m}(B)$. As $x^j = s^j(x)$, it follows from Bott's Formula (3.3) that

$$\delta^{B|x}\big(e_{m-j}(B)x^j\big) = s^{0,\dots,\hat{j},\dots,m,j}(B, x) = (-1)^{m-j}.$$

Therefore, partial symmetrization of $\Delta_t(B|x)$ yields the sum in (4.2.1).

Note that $\Delta(B\,|\,x) = \Delta_1(B\,|\,x)$ is the polynomial appearing as the denominator in the definition of $\delta^{B|x}$. As in (2.8), a permutation σ in $\mathfrak{T}(B\,|\,x)$ is determined by the value $\sigma(x)$. In terms of the original names of the letters, with $\mathbf{A} := (B\,|\,x)$, we have $\mathfrak{T}(B\,|\,x) = \mathfrak{T}(\mathbf{A})$, and σ is determined by the index $p = 1, \dots, n$ such that $\sigma(a_n) = a_p$. If $\sigma(a_n) = a_p$ we have $\sigma(\Delta_t(\mathbf{A})) = \prod_{q \neq p}(a_p - ta_q)$. Hence the formula (4.2.1) is equivalent to the following:

$$\sum_{p=1}^n \prod_{q \neq p} \frac{a_p - ta_q}{a_p - a_q} = \frac{t^n - 1}{t - 1}. \tag{4.2.2}$$

(4.3) Lemma. *The total symmetrization of the polynomial $\Delta_t(A)$ is the constant given by the equations,*

$$\delta^A\Big(\Delta_t(A)\Big) = \prod_{i=1}^{n} \frac{t^i - 1}{t-1} = \sum_{\mu \in \mathfrak{S}(A)} t^{\ell(\mu)}. \tag{4.3.1}$$

Proof. The first equation in (4.3.1) is proved by induction on the number of n of letters of A. The two sides of the equation reduce to 1 when $n = 1$. If $n > 1$, then, in the notation of Example (4.2), we may assume that the equation holds for the alphabet B. Now apply (2.3) with $r = 2$ to the partitioning $(B \mid x)$. Note that δ^x is the identity. Hence, by Equation (2.3.2) applied to the factorization $\Delta_t(A) = \Delta_t(B|x)\,\Delta_t(B)$, we obtain the first of the following two equations:

$$\delta^A\Big(\Delta_t(A)\Big) = \delta^{B|x}\Big(\Delta_t(B \mid x)\,\delta^B(\Delta_t(B))\Big) = \delta^{B|x}\Big(\Delta_t(B \mid x)\Big)\,\delta^B\big(\Delta_t(B)\big).$$

The second equation holds because $\delta^B(\Delta_t(B))$ is an element of R. In fact, by the induction hypothesis, this element is the product of the first $n-1$ factors in the product in (4.3.1). By (4.2.1), $\delta^{B|x}(\Delta_t(B|x))$ is the last of the n factors. Thus the first equation in (4.3.1) holds.

To obtain the alternative expression for $\delta^A(\Delta_t(A))$ in (4.3.1), recall that $\delta^A(\Delta_t(A))$ is the quotient obtained by dividing the following sum by the Vandermonde determinant $\Delta(A)$:

$$\sum_{\mu \in \mathfrak{S}(A)} (\operatorname{sign}\mu)\mu(\Delta_t(A)). \tag{4.3.2}$$

The sum (4.3.2) has the same degree, $n(n-1)/2$, as the Vandermonde determinant $\Delta(A)$. Therefore, to obtain the quotient, it suffices to compare, with respect to the total order "$<$" on monomials, the smallest terms of the two polynomials. For $\Delta(A)$, the smallest term is obtained as the product of the smallest terms of each factor $b - a$ for $a < b$. So the smallest term in $\Delta(A)$ is the monomial a^E obtained as the product of the letters b for all pairs (a, b) with $a < b$. Consider similarly the smallest term in $\mu(\Delta_t(A))$. The factors of $\mu(\Delta_t(A))$ are of the form,

$$\tilde b - t\tilde a \quad\text{or}\quad -t\tilde b + \tilde a, \qquad \text{for all pairs of letters } \tilde a < \tilde b,$$

where the factors of the first form correspond to factors $b - ta$ of $\Delta_t(A)$ such that (a, b) is not an inversion for μ, and the factors of the second form correspond to factors $b - ta$ such that (a, b) is an inversion for μ. It follows that the smallest term in $\mu(\Delta_t(A))$ is equal to $(-t)^{\ell(\mu)}a^E$. Hence the smallest term in the sum (4.3.2) is equal to the sum on the right hand side of (4.3.1) multiplied by a^E. Therefore, the alternative expression for $\delta^A(\Delta_t(A))$ holds. □

(4.4) Example. Consider, for a partially symmetric polynomial f, the partial symmetrization,

$$\delta^{\mathbf{A}}\big(f\,\Delta_t(\mathbf{A})\big) = \sum_{\sigma \in \mathfrak{T}(\mathbf{A})} \sigma\Big(f \prod_{a \ll b} \frac{b - ta}{b - a}\Big). \tag{4.4.1}$$

The partial symmetrization is a polynomial, symmetric in the letters of A, with coefficients depending on t. The fraction in (4.4.1) is the fraction $\Delta_t(\mathbf{A})/\Delta(\mathbf{A})$ of two homogeneous polynomials of the same degree. Hence, if f is homogeneous, then $\delta^{\mathbf{A}}\big(f\,\Delta_t(\mathbf{A})\big)$ is homogeneous of the same degree.

Denote by φ the following polynomial in t:

$$\varphi(t) := \prod_{i=1}^{r}\prod_{j=1}^{n_i}\frac{t^j-1}{t-1} = \sum_{\mu\in\mathfrak{S}(\mathbf{A})} t^{\ell(\mu)} \quad \text{where } n_i = |A_i|. \tag{4.4.2}$$

The second equation in (4.4.2) follows from the second equation in (4.3.1) since any permutation $\mu \in \mathfrak{S}(\mathbf{A})$ is a product $\mu = \mu_1\cdots\mu_r$ with uniquely determined factors $\mu_i \in \mathfrak{S}(A_i)$ and $\ell(\mu) = \ell(\mu_1) + \cdots + \ell(\mu_r)$.

Then we have the equations,

$$\delta^{A}\Big(f\,\Delta_t(A)\Big) = \sum_{\mu\in\mathfrak{S}(A)} \mu\Big(f\prod_{a<b}\frac{b-ta}{b-a}\Big) = \varphi(t)\;\delta^{\mathbf{A}}\Big(f\,\Delta_t(\mathbf{A})\Big). \tag{4.4.3}$$

Indeed, the first equation follows from the definition of the symmetrization operator δ^A. For the alternative expression, apply (2.3.2) to the factorization of (4.1.1): $f\,\Delta_t(A) = f\,\Delta_t(\mathbf{A})\,g_1\cdots g_r$ with $g_i := \Delta_t(A_i)$. The result is the equation,

$$\delta^{A}\Big(f\,\Delta_t(A)\Big) = \delta^{\mathbf{A}}\Big(f\,\Delta_t(\mathbf{A})\;\delta^{A_1}(g_1)\cdots\delta^{A_r}(g_r)\Big).$$

On the right, $\delta^{A_i}(g_i)$ is determined in Lemma (4.3). The result is a scalar, equal to the i'th factor in $\varphi(t)$. Hence (4.4.3) follows from the linearity of $\delta^{\mathbf{A}}$.

(4.5) Definition. Let J be a strictly increasing multi index of size n. Associate with J the partitioning $\mathbf{A} = \mathbf{A}_J$ defined by the sequence $p_1 < \cdots < p_{r-1}$ of indices $p < n$ such that $j_p + 1 < j_{p+1}$, cf. (1.1). Set $E := (0, 1, \dots, n-1)$. Then the sequence $J - E$ is weakly increasing. In fact, by the choice of the p_i, the sequence of exponents in the monomial a^{J-E} is constant exactly at the letters of the intervals A_i. Hence $\mathfrak{S}(\mathbf{A}) \subseteq \mathfrak{S}(A)$ is the subgroup of permutations fixing a^{J-E}, and the different permutations $\sigma(a^{J-E})$ of a^{J-E} correspond to the elements $\sigma \in \mathfrak{T}(\mathbf{A})$.

Define the *Hall–Littlewood symmetric polynomial* $P^J(A;t)$ as the partial symmetrization, for $\mathbf{A} = \mathbf{A}_J$,

$$P^J(A;t) := \delta^{\mathbf{A}}\Big(a^{J-E}\,\Delta_t(\mathbf{A})\Big) = \sum_{\sigma\in\mathfrak{T}(\mathbf{A})} \sigma\Big(a^{J-E}\,\frac{\Delta_t(\mathbf{A})}{\Delta(\mathbf{A})}\Big). \tag{4.5.1}$$

Since a^{J-E} and $\Delta_t(\mathbf{A})$ are partially symmetric polynomials, it follows that $P^J(A;t)$, as the partial symmetrization of their product, is indeed a symmetric polynomial.

By the equality of the operators in (2.3.3), the polynomial $P^J(A;t)$ is equal to the total symmetrization,

$$P^J(A;t) = \delta^A\Big(\frac{a^J a^{E_1}\cdots a^{E_r}}{a^E}\Delta_t(\mathbf{A})\Big). \tag{4.5.2}$$

By the factorization (4.1.2) we have that

$$\frac{a^{E_1}\cdots a^{E_r}}{a^E}\Delta_t(\mathbf{A}) = \frac{\Delta_t(\mathbf{A})}{\Delta_0(\mathbf{A})} = \prod_{a\ll b}\frac{b-ta}{b}.$$

Hence it follows from (4.5.2) that

$$P^J(A;t) = \delta^A\Big(a^J\prod_{a\ll b}\Big(1-t\frac{a}{b}\Big)\Big). \tag{4.5.3}$$

On the other side, by (4.4.3) we have the equations,

$$\varphi_J(t)\,P^J(A;t) = \delta^A\big(a^{J-E}\Delta_t(A)\big) = \sum_{\mu\in\mathfrak{S}(A)}\mu\Big(a^{J-E}\prod_{a<b}\frac{b-ta}{b-a}\Big), \tag{4.5.4}$$

where $\varphi_J(t)$ is the polynomial $\varphi(t)$ determined by (4.4.2) for $\mathbf{A}=\mathbf{A}_J$.

For $t=0$ we have the equation, $a^{E_1}\cdots a^{E_r}\Delta_0(\mathbf{A}) = a^E$. Hence it follows from (4.5.2) that $P^J(A;0)=\delta^A(a^J)$, that is, $P^J(A;0)$ is the Schur polynomial,

$$P^J(A;0) = s^J(A). \tag{4.5.5}$$

For $t=1$, we have that $\Delta_1(\mathbf{A}) = \Delta(\mathbf{A})$. Hence it follows from (4.5.1) that

$$P^J(A;1) = \sum_{\sigma\in\mathfrak{T}(\mathbf{A})}\sigma(a^{J-E}).$$

The monomials $\sigma(a^{J-E})$ for $\sigma\in\mathfrak{T}(\mathbf{A})$ are exactly the different permutation of the monomial a^{J-E}. Hence $P^J(A;1)$ is the monomial symmetric polynomial,

$$P^J(A;1) = m^{J-E}(A). \tag{4.5.6}$$

(4.6) Proposition. *In the basis of proper Schur polynomials, the expansion of the Hall–Littlewood polynomial $P^J(A;t)$ is of the form,*

$$P^J(A;t) = \sum_{I\gg J}\alpha_{JI}(t)s^I(A), \tag{4.6.1}$$

where the sum is over strictly increasing multi indices I of size n such that $\|I\| = \|J\|$ and $i_1+\cdots+i_s \geq j_1+\cdots+j_s$ for $s=1,\ldots,n$. Moreover, the coefficient $\alpha_{JJ}(t)$ is equal to 1.

Proof. By (4.5.3), the polynomial $P^J(A;t)$ is the total symmetrization,

$$P^J(A;t) = \delta^A\big(\Pi\big) \quad \text{where} \quad \Pi := a^J \prod_{a \ll b}\Big(1 - t\frac{a}{b}\Big).$$

The product Π is a polynomial, homogeneous of degree equal to $\|J\|$. Consider a monomial a^K occuring in Π. Fix $s \le n$ and a subset $\{l_1, \dots, l_s\}$ of s different indices. We claim that the following inequality holds:

$$k_{l_1} + \cdots + k_{l_s} \ge j_1 + \cdots + j_s. \tag{4.6.2}$$

Indeed, to prove (4.6.2), we may assume that the l_q are strictly increasing. The left side of (4.6.2) is the sum of the exponents in a^K corresponding to the indices $l_1, \dots, l_s$. This sum of exponents is defined for every monomial. For a^J it is equal to $j_{l_1} + \cdots + j_{l_s}$. Clearly, if a monomial is multiplied by a factor a/b for $a \ll b$, then the sum of the exponents is decreased by 1 if b is one of the a_{l_q} and a is not, it is increased by 1 if a is one of the a_{l_q} and b is not, and it is unchanged otherwise. Therefore, when a^J is multiplied by a number of different such factors a/b, the sum $j_{l_1} + \cdots + j_{l_s}$ is decreased at most by the number of factors a/b such that $a \ll b$ and b is one of the a_{l_q} and a is not. As the l_q are strictly increasing, the latter number of factors is at most equal to $\sum_{q=1}^{s}(l_q - q)$. Hence we have the inequality,

$$k_{l_1} + \cdots + k_{l_s} \ge j_{l_1} + \cdots + j_{l_s} - \sum_{q=1}^{s}(l_q - q). \tag{4.6.3}$$

Since the l_q were strictly increasing, we have that $l_q \ge q$. Moreover, the multi index J was strictly increasing. Hence, we have the inequalities,

$$j_{l_q} - l_q \ge j_q - q \quad \text{for } q = 1, \dots, s. \tag{4.6.4}$$

Clearly (4.6.2) follows from the inequalities (4.6.3) and (4.6.4).

It follows from (4.6.2) that the expansion of $P^J(A;t)$ has the form asserted in (4.6.1). Indeed, from the expansion of Π as an R-linear combination of monomials a^K we obtain the expansion of $P^J(A;t)$ as the corresponding R-linear combination of $\delta^A(a^K) = s^K(A)$. The polynomial $s^K(A)$ is alternating in K. Hence a monomial a^K contributes with zero if two entries in K are equal. If all entries in K are different, then $s^K(A)$ is up to sign equal to the proper Schur polynomial $s^I(A)$, where I is the strictly increasing permutation of K. Here $\|I\| = \|K\| = \|J\|$ and, obviously, the inequality

$$i_1 + \cdots + i_s \ge j_1 + \cdots + j_s, \tag{4.6.5}$$

for the strictly increasing permutation I of K, holds iff (4.6.2) holds for any s indices $l_1, \dots, l_s$. As $s = 1, \dots, n$ was arbitrary, it follows that $I \gg J$.

Obviously, the monomial a^J occurs with coefficient 1 in the expansion of Π. Therefore, to prove the last assertion of the Proposition, it suffices to prove that if the monomial a^K occurs in the expansion of Π and K is a permutation of J, then $K = J$. By an inductive argument, it suffices to prove that if $k_q = j_q$ for $q = 1, \dots, s-1$, then $k_s = j_s$. Let $l \geq s$ be the index such that $k_l = j_s$. Apply the reasoning leading to (4.6.2) with the set of l_q equal to $\{1, \dots, s-1, l\}$. As (4.6.2) is an equality, it follows in particular that the last inequality in (4.6.4) is an equality, that is, $j_l - l = j_s - s$. This equality shows that a_l and a_s belong to the same interval in the partitioning $\mathbf{A}_J$. Therefore, there are no factors a/b where $a \ll b$ and b is one of the a_{l_q} and a is not. Hence we obtain instead of (4.6.3) the inequality,

$$k_1 + \cdots + k_{s-1} + k_l \geq j_1 + \cdots + j_{s-1} + j_l.$$

Since $k_q = j_q$ for $q = 1, \dots, s-1$, it follows from the inequality that $k_l \geq j_l$. As $k_l = j_s$ and $l \geq s$, we conclude that $l = s$.

Hence both assertions of the Proposition have been proved. □

(4.7) Corollary. *For any $t \in R$, the Hall–Littlewood polynomials $P^J(A; t)$, for all strictly increasing multi indices J of size n, form an R-basis of* $\mathrm{Sym}_R[A]$.

Proof. It follows from the Proposition that the matrix expressing the $P^J(A; t)$ in terms of the basis $s^I(A)$ is a lower triangular matrix with 1 in the diagonal. In particular, the matrix is invertible. Hence the $P^J(A; t)$ form a basis. □

(4.8) Example. For $d \geq 0$ we have the equations,

$$P^{d,d+1,\dots,d+n-1}(A; t) = s^{d,d+1,\dots,d+n-1}(A) = e_n(A)^d. \tag{4.8.1}$$

Indeed, let $J := (d, d+1, \dots, d+n-1)$. Then, in the notation of (4.5), the corresponding partitioning is the trivial partitioning $\mathbf{A} = (A)$ with $r = 1$. As $a^{J-E} = (a_1 \cdots a_n)^d$, the asserted equation follows from (4.5.1).

(4.9) Example. For $d \geq 1$ we have the equations,

$$P^{0,1,\dots,n-2,n-1+d}(A; t) = \sum_b b^d \prod_{a \neq b} \frac{b - ta}{b - a} \tag{4.9.1}$$

$$= \sum_{n-d \leq j \leq n-1} (-t)^{n-1-j} s^{0,1,\dots,\hat{j},\dots,n-1,d+j}(A), \tag{4.9.2}$$

where a, b in the middle expression are letters of A. Indeed, let $J := (0, 1, \dots, n-2, n-1+d)$. Then, in the notation of (4.5), the corresponding partitioning is $\mathbf{A} = (a_1, \dots, a_{n-1}|a_n)$. Clearly, we have that $a^{J-E} = a_n^d$ and $\Delta_t(\mathbf{A}) = \prod_{q<n}(a_n - ta_q)$. In (4.5.1), the permutations σ are determined by the index p such that $\sigma(a_n) = a_p$. Hence the first expression for $P^J(A; t)$ is obtained from (4.5.1).

To obtain the second expression of $P^J(A;t)$, note that the monomials a^K occurring in the expansion of the product in (4.5.3) either have two equal exponents or appear as terms,

$$(-t)^{n-1-j}a^J(a_{j+1}\cdots a_{n-1})a_n^{-(n-1-j)} = (-t)^{n-1-j}a^{0,1,\dots,\hat{j},\dots,n-1,d+j}.$$

The second expression for $P^J(A;t)$ is a consequence, since $\delta^A(a^K) = s^K(A)$.

It is customary to extend the notation $P^J(A;t)$ to strictly increasing multi indices J of arbitrary size in exactly the same way as the extension was defined in (SYM.6.13) for the Schur polynomials $s^J(A)$. In this extended notation, the second expression is the following:

$$P^d(A;t) = \sum_{k=0}^{d-1}(-t)^k s^{1,\dots,k,d}(A). \tag{4.9.3}$$

By replacing x by $1/T$ in the Lagrange interpolation formula (2.8.2), we obtain the equation in $R(A)[[T]]$,

$$\prod_a \frac{1-taT}{1-aT} = 1 + (1-t)\sum_b \frac{bT}{1-bT}\prod_{a\neq b}\frac{b-ta}{b-a}.$$

Therefore, by (4.9.1), it follows that

$$\prod_{a\in A}\frac{1-taT}{1-aT} = 1 + (1-t)\sum_{d\geq 1} P^d(A;t)T^d. \tag{4.9.4}$$

(4.10) Example. Assume that $d \geq 1$ and $1 \leq j \leq n-1$. Consider the polynomial $P^I(A;t)$ for $I := (0,\dots,j-1,j+d,\dots,n-1+d)$. In the notation of (4.5), the corresponding partitioning is $\mathbf{A} = (A_1|A_2)$, where $A_1 := \{a_1,\dots,a_j\}$ and $A_2 := \{a_{j+1},\dots,a_n\}$. By (4.5.1) we have that $P^I(A;t)$ is the partial symmetrization of the product,

$$a^{I-E}\Delta_t(\mathbf{A}) = \prod_{b\in A_2} b^d \prod_{a\in A_1,\, b\in A_2}(b-ta).$$

The product is given by the expansion obtained from (3.2.1) after the substitution $a_i := -ta_i$ for $i = 1,\dots,j$, that is, we have the equation,

$$a^{I-E}\Delta_t(\mathbf{A}) = \sum_{J\subseteq[n]}(-t)^{\|J-E_1\|}s^J(A_1)s^{\hat{J}^{\sim}}(A_2),$$

where the sum is over subsets J of size j, $\hat{J}$ is the extension (see (SYM.6.13)) of J to a subset of size $j+d$ of $[n+d]$, and $\hat{J}^{\sim} \subseteq [n+d]$ is the complement of $\hat{J}$. The multi index E_1, corresponding to the interval A_1, is $E_1 := [j] = (0,1,\dots,j-1)$.

Therefore, by Bott's Formula, we obtain the following expression for the partial symmetrization,

$$P^I(A;t) = \sum_{J\subseteq[n]}(-t)^{\|J-E_1\|}s^{J\,\hat{J}^{\sim}}(A), \tag{4.10.1}$$

where the sum is over size-j subsets J as in the previous equation.

Consider the special case $d = 1$. If $J \neq E_1$, then there exists an index $q = 1, \dots, j$ such that $j_q \geq 1$ and $j_q - 1 \notin J$. It follows that j_q belongs to the complement of $\hat{J}$. Hence $s^{J\,\hat{J}^{\sim}}(A) = 0$. Therefore, the sum in (4.10.1) reduces to the term corresponding to $J = E_1$. Thus we obtain the first of the equations,

$$P^{0,1,\dots,\hat{j},\dots,n}(A;t) = s^{0,\dots,\hat{j},\dots,n}(A) = e_{n-j}(A). \tag{4.10.2}$$

The second equation is (SYM.6.8.1). Note that the equations hold for $j = 0$ by (4.8) and trivially for $j = n$. Note also that the first equation in (4.10.2) follows directly from (4.6): For $J = (0, 1, \dots, \hat{j}, \dots, n)$ there is only one term in the sum (4.6.1).

In the extended notation of Example (4.9), the polynomial $P^I(A;t)$ is equal to the polynomial $P^{d,d+1,\dots,d+n-1-j}(A;t)$. In particular, (4.10.2) is the following equation, for $0 \leq k \leq n$:

$$P^{1,\dots,k}(A;t) = s^{1,\dots,k}(A) = e_k(A). \tag{4.10.3}$$

(4.11) Proposition. *Let $\hat{A}$ be an alphabet obtained by adding a single letter to A. For a strictly increasing multi index I of size $n+1$, denote by $P^I(A,0;t)$ the polynomial obtained from $P^I(\hat{A};t)$ by specializing the additional letter of $\hat{A}$ to 0. If I is not an extension of a multi index of size n, then $P^I(A,0;t) = 0$. Moreover, for an extension $I = \hat{J}$ of a strictly increasing multi index J of size n, we have the equation,*

$$P^{\hat{J}}(A,0;t) = P^J(A;t). \tag{4.11.1}$$

Proof. Note that, since $P^I(\hat{A};t)$ is symmetric in the letters of $\hat{A}$, we obtain the same polynomial $P^I(A,0;t)$ by specializing any letter of $\hat{A}$ to zero and specializing the remaining letters to the letters of A. So we assume that the addtional letter is the smallest in $\hat{A}$, denoted a_0. Coordinates in size-$(n+1)$ multi indices are indexed with $0, 1, \dots, n$. We let $\hat{E} := (0, \dots, n)$.

Let $\hat{\mathbf{A}}$ be the partitioning corresponding to I. Set $\hat{K} = I - \hat{E}$. Then $\hat{K} = (k_0, K)$, where K is a weakly increasing multi index of size n and $k_0 \leq k_1$. As $\hat{a}^{\hat{K}} = a_0^{k_0} a^K$, it follows from (4.5.1) that

$$P^I(\hat{A};t) = \sum_{\sigma \in \mathfrak{T}(\hat{\mathbf{A}})} \sigma\Big(a_0^{k_0} a^K \frac{\Delta_t(\hat{\mathbf{A}})}{\Delta(\hat{\mathbf{A}})}\Big). \tag{4.11.2}$$

It follows that the polynomial $P^I(A,0;t)$ is the sum of the rational functions obtained by the substitution $a_0 := 0$ in each term of (4.11.2) (since the factors $\sigma(a_i) - \sigma(a_j)$ of the denominators $\sigma(\Delta(\hat{\mathbf{A}}))$ remain regular under the substitution).

Assume that I is not an extension of a multi index of size n. Then $k_0 \geq 1$. Hence all exponents of the monomial $\hat{a}^{\hat{K}}$ are at least equal to 1. It follows that substitution $a_0 := 0$ in $\sigma(a_0^{k_0} a^K)$, for any $\sigma \in \mathfrak{S}(\hat{A})$, yields 0. Hence $P^I(A,0;t) = 0$.

Assume that I is an extension, say $I = \hat{J}$. Then $k_0 = 0$ and $J - E = K$. Let $\hat{A}_1$ be the first interval in the partitioning $\hat{\mathbf{A}}$ corresponding to $\hat{J}$, that is, the interval of letters with exponent 0 in $\hat{a}^{\hat{K}}$. Then, in the monomial $\hat{a}^{\hat{K}}$, the letter a has exponent 0 if and only if $a \in \hat{A}_1$.

Consider a permutation $\sigma \in \mathfrak{T}(\hat{\mathbf{A}})$ and the corresponding term in (4.11.2). If $\sigma(a_0) \neq a_0$ then $a_0 = \sigma(a_i)$ for some $i > 0$. As σ is assumed to be increasing on $\hat{A}_1$, it is excluded that $a_i \in \hat{A}_1$. Hence $k_i > 0$. Thus the term contains the factor $\sigma(a_i^{k_i}) = a_0^{k_i}$, and so the term vanishes after the substitution $a_0 := 0$.

Therefore, to determine the result of the substitution, it suffices in (4.11.2) to consider terms for $\sigma \in \mathfrak{T}(\hat{\mathbf{A}})$ for which $\sigma(a_0) = a_0$. Clearly, the latter set of permutations correspond to the permutations in $\mathfrak{T}(\mathbf{A})$, where $\mathbf{A}$ is the partitioning of A corresponding to J. Moreover, since $\sigma(a_0) = a_0$, it follows that all factors of $\sigma(\Delta_t(\hat{\mathbf{A}})/\Delta(\hat{\mathbf{A}}))$ containing a_0 are of the form $(b - ta_0)/(b - a_0)$, and they yield 1 after the substitution $a_0 := 0$. It follows that the result of the substitution $a_0 := 0$ in (4.11.2) is the sum

$$\sum_{\sigma \in \mathfrak{T}(\mathbf{A})} \sigma\Big(a^K \frac{\Delta_t(\mathbf{A})}{\Delta(\mathbf{A})}\Big),$$

and hence equal to $P^J(A; t)$, as asserted. ∎

(4.12) Example. It is often convenient to consider together with $P^J(A; t)$ a second polynomial $R_K(A; t)$ defined for an arbitrary size-n multi index K as the total symmetrization,

$$R_K(A; t) := \delta^A\Big(a^K \Delta_t(A)\Big). \tag{4.12.1}$$

If K is weakly increasing, then $K + E$ is strictly increasing, and, by (4.5.4), we have the equation,

$$\varphi_{K+E}(t)\, P^{K+E}(A; t) = R_K(A; t). \tag{4.12.2}$$

Consider an arbitrary partitioning $\mathbf{A} = (A_1 \,|\, \cdots \,|\, A_r)$ of A. Write K as a concatenation $K = K_1 \cdots K_r$, where each K_i is a multi index of size $|A_i|$. For each $i = 1, \dots, r$, we let a^{K_i} denote (with an abuse of notation) the monomial in the letters of A_i whose exponents are the coordinates of K_i; thus, $a^K = a^{K_1} \cdots a^{K_r}$. Consider the factorization $a^K \Delta_t(A) = \Delta_t(\mathbf{A}) g_1 \cdots g_r$ where $g_i := a^{K_i} \Delta_t(A_i)$. Then $\delta^{A_i}(g_i) = R_{K_i}(A_i; t)$. Hence the following formula results from (2.3.3):

$$R_K(A; t) = \delta^{\mathbf{A}}\Big(\Delta_t(\mathbf{A})\, R_{K_1}(A_1; t) \,\cdots\, R_{K_r}(A_r; t)\Big). \tag{4.12.3}$$

As an example, consider the partitioning $(\bar{A} \,|\, a_n)$ where $\bar{A} = (a_1, \dots, a_{n-1})$ is the inverval of the first $n - 1$ letters. Let $\bar{K}$ denote the truncated multi index $\bar{K} = (k_1, \dots, k_{n-1})$, and use a similar notation for all size-n multi indices. In the notation of (4.12.3) we have $K_1 = \bar{K}$ and $K_2 = (k_n)$. In particular, $R_{K_2}(A_2; t) = a_n^{k_n}$. Then (4.12.3) is the following formula:

$$R_K(A; t) = \sum_{\sigma \in \mathfrak{T}(\bar{A} \,|\, a_n)} \sigma\Big(\frac{\Delta_t(\bar{A} \,|\, a_n)}{\Delta(\bar{A} \,|\, a_n)}\, R_{\bar{K}}(\bar{A}; t)\, a_n^{k_n} \Big). \tag{4.12.4}$$

The notation in (4.12.4) is sometimes simplified by introducing the following rational function $\Pi_b = \Pi_b(A)$, for any letter $b \in A$:

$$\Pi_b = \prod_{a \neq b} \frac{b - ta}{b - a}.$$

The sum in (4.12.4) is essentially a sum over $p = 1, \ldots, n$ since there is for every p a unique permutation $\sigma \in \mathfrak{T}(\bar{A} \mid a_n)$ such that $\sigma(a_n) = a_p$. Moreover, we have $\Delta_t(\bar{A}|a_n)/\Delta(\bar{A}|a_n) = \Pi_{a_n}$. Thus,

$$R_K(A;t) = \sum_{p=1}^{n} \Big(\Pi_{a_p}\, R_{\bar{K}}(a_1, \ldots, \widehat{a_p}, \ldots, a_n\,;\, t)\, a_p^{k_n} \Big). \tag{4.12.5}$$

For a strictly increasing multi index J, Equation (4.12.5) for $K := J - E$ implies the following equation for $P^J(A;t)$ (connected to $R_{J-E}(A;t)$ via (4.12.2)):

$$\frac{t^{n_r} - 1}{t - 1} P^J(A;t) = \sum_{p=1}^{n} \Big(\Pi_{a_p}\, P^{\bar{J}}(a_1, \ldots, \widehat{a_p}, \ldots, a_n; t)\, a_p^{j_n - (n-1)} \Big). \tag{4.12.6}$$

Indeed, according to the definition in (4.4) and (4.5), we have $\varphi_{\bar{J}}(t)(t^{n_r} - 1)/(t - 1) = \varphi_J(t)$, as is seen by checking the two cases $n_r > 1$ and $n_r = 1$. Hence, when $\varphi_{\bar{J}}(t)$ is regular in R, (4.12.6) follows from (4.12.5) after division by $\varphi_{\bar{J}}(t)$. In particular, Equation (4.12.6) holds when the ground ring R is replaced by the polynomial ring $R[T]$ and t by T. Then the case of an arbitrary element $t \in R$ follows after the substitution $T := t$.

Schur functions

1. Schur functions.

(1.1) Setup. Fix a finite alphabet $A = \{a_1, \dots, a_n\}$ of $n \geq 0$ letters. In addition to the ring $R[A]$ of polynomials and the ring $R(A)$ of rational functions we consider the ring of formal power series $R[[A]]$. Clearly, the action of $\mathfrak{S}(A)$ on the polynomials extends canonically to an action on the power series. More generally, as explained in (DIFF.3.13), the action of the ring $\mathcal{E}_R[A]$ on $R[A]$ extends canonically to an action on $R[[A]]$.

For a power series p in $R[[A]]$, we denote by p_j its homogeneous term of degree j. By convention, $p_j = 0$ for $j < 0$.

(1.2) Definition. In the sequel we will consider matrices with coefficients in $R[A]$, possibly with an infinite number of rows or columns. Let s be a power series in $R[[A]]$. For any integer i we denote by $s[i]$ the infinite row of homogeneous terms of s shifted i places to the right, that is, $s[i]$ is the row whose j'th entry, for $j = 0, 1, \dots$, is equal to s_{j-i}. For any integer j we denote by $s^{j\geq}$ the infinite column whose i'th entry, for $i = 0, 1, \dots$, is equal to s_{j-i}. The infinite matrix S whose rows are $s[i]$ for $i = 0, 1, \dots$, or equivalently, whose columns are $s^{j\geq}$ for $j = 0, 1, \dots$, is an upper triangular matrix,

$$S = \begin{pmatrix} s[0] \\ s[1] \\ \vdots \end{pmatrix} = \left(s^{0\geq}, s^{1\geq}, \dots\right) = \begin{pmatrix} s_0 & s_1 & s_2 & \dots \\ 0 & s_0 & s_1 & \dots \\ 0 & 0 & s_0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

If $I = (i_1, \dots, i_r)$ and $J = (j_1, \dots, j_r)$ are sequences with the same number r of non-negative integers, we denote by $S^{I/J}$ the determinant of the $r \times r$ matrix obtained from S by selecting the rows with indices from I and the columns with indices from J, that is,

$$S^{I/J} = \begin{vmatrix} s_{j_1-i_1} & s_{j_2-i_1} & \dots & s_{j_r-i_1} \\ s_{j_1-i_2} & s_{j_2-i_2} & \dots & s_{j_r-i_2} \\ \vdots & \vdots & & \vdots \\ s_{j_1-i_r} & s_{j_2-i_r} & \dots & s_{j_r-i_r} \end{vmatrix}. \tag{1.2.1}$$

Clearly, if p is a second power series, then, for $i \geq 0$,

$$p[i]\, s^{j\geq} = (ps)_{j-i}, \tag{1.2.2}$$

where the left hand side is the product of a row and a column. In particular,

$$p[i]S = (ps)[i] \quad \text{and} \quad Sp^{j\geq} = (sp)^{j\geq}. \tag{1.2.3}$$

(1.3) Definition. Let W be a (commutative) word in the letters of A, that is, W is an element in the free abelian group generated by the letters of A. In more concrete terms, W is a formal sum,

$$W = z_1a_1 + \cdots + z_na_n,$$

where the z_k's are integers. The *degree* of the word W is the sum of the coefficients, $|W| = z_1 + \cdots + z_n$. Associate with W the following power series in $R[[A]]$:

$$S(W) = \prod_{k=1}^{n}\Big(\frac{1}{1-a_k}\Big)^{z_k},$$

and denote by $S_j(W) = S(W)_j$ the homogeneous term of degree j in $S(W)$. Obviously, the map $W \mapsto S(W)$ is a homomorphism from the group of words to the multiplicative group of series in $R[[A]]$ with constant term 1. If the word W is *positive*, that is, if the coefficients z_k are non-negative, then $S(-W) = \prod(1-a_k)^{z_k}$ is a polynomial of degree equal to the degree of W. In particular, if W is positive, then $S_j(-W) = 0$ for $j > |W|$.

The special word $a_1 + \cdots + a_n$ will be denoted A. As a word, the degree of A is the cardinality of the alphabet A. Clearly $S_j(A)$ is the j'th *complete symmetric function* in the letters of A. The power series $S(-A)$ is the polynomial $\prod_{k=1}^{n}(1-a_k)$, and $(-1)^jS_j(-A)$ is the j'th *elementary symmetric function* in the letters of A.

(1.4) Remark. Enlarge the alphabet A with a single letter x, and consider the word $x - A$ in the letters of the enlarged alphabet. The power series $S(x - A)$ is the product of $(1 - x)^{-1}$ and $\prod_{k=1}^{n}(1-a_k)$. For $j \geq n$ we thus obtain the equation $S_j(x - A) = x^{j-n}\prod_{k=1}^{n}(x - a_k)$ for the j'th homogeneous term. The latter equation, with a zero on both sides, holds also if x is one of the a_k.

(1.5) Definition. Let W be a commutative word in the letters of A. For any two multi indices I and J of size r, we define the associated *skew Schur function* $S^{I/J}(W)$ as the determinant $S^{I/J}$ of (1.2.1) associated with the power series $s := S(W)$. In other words, if we consider the matrices,

$$\begin{pmatrix} S(W)[i_1] \\ \vdots \\ S(W)[i_r] \end{pmatrix} \quad \text{and} \quad \Big(S(W)^{j_1\geq}, \dots, S(W)^{j_r\geq}\Big), \tag{1.5.1}$$

then the skew Schur function $S^{I/J}(W)$ is equal to the determinant of the matrix that is obtained either from the first matrix in (1.5.1) by selecting its columns from J (assuming that all entries in J are nonnegative), or from the second matrix in (1.5.1) by selecting its rows from I (assuming that all entries in I are nonnegative).

The Schur function is a polynomial in $R[A]$. When I is the sequence $0, 1, \dots, r-1$ we write $S^J(W) := S^{I/J}(W)$. The polynomial $S^J(W)$ is the *ordinary Schur function* $S^J(W)$.

(1.6) Observation. Several properties of the skew Schur functions $S^{I/J}(W)$ are obvious from their definition as determinants. Clearly, the function is alternating in I, that is, if the entries in I are permuted, then the function is changed by the signature of the permutation and the function vanishes if two entries of I are equal. Similarly, the function is alternating in J.

The Schur function $S^{I/J}(W)$ is the determinant of an $r \times r$ matrix whose (p,q)'th entry is the homogeneous part of degree $j_q - i_p$ of the power series $S(W)$. In particular, the entry vanishes if $j_q < i_p$. As a consequence, if both sequences I and J are strictly increasing, then $S^{I/J}(W)$ vanishes unless $i_k \leq j_k$ for $k = 1, \dots, r$. Note that under the latter condition, each product in the expansion of the determinant $S^{I/J}(W)$ is homogeneous of degree $\sum j_k - \sum i_k = \|J\| - \|I\|$ in the letters of A. In particular, the ordinary Schur function $S^J(W)$ is homogeneous of degree $\|J\| - r(r-1)/2$.

(1.7) Remark.

Clearly, the Schur function $S^{I/J}(W)$, defined as the determinant of a matrix of the form $\{s_{j_p - i_q}\}$, is unchanged if the same integer is added to all entries of I and J. As a consequence, we may often assume that all entries in I and J are nonnegative. If all entries in I are nonnegative and a single entry in J is negative then $S^{I/J}(W) = 0$ since the column $S(W)^{j\geq}$ vanishes when $j < 0$.

Consider the Schur function $S^{I/J}(W)$ corresponding to two sequences $I = (i_1, \dots, i_r)$ and $J = (j_1, \dots, j_r)$ where I is strictly increasing. Assume that $i_1 = j_1$. Then the first column in the matrix defining the Schur function has 1 as its first entry and 0 as the remaining entries. Therefore, the Schur function is unchanged if the sequences I and J are replaced by $(i_2, \dots, i_r)$ and $(j_2, \dots, j_r)$. Clearly, the argument can be repeated if the first p entries in I and J agree. In particular, $S^{J/J}(W) = S^{0/0}(W) = 1$ for any strictly increasing J.

These observations show for the ordinary Schur function $S^J(W)$ that $S^J(W) = 0$ if an entry in J is negative. Moreover, $S^J(W)$ is unchanged if J is replaced for some $p \geq 1$ by the *extended* sequence, of length $p + r$,

$$\hat{J} = (0, 1, \dots, p-1, p + j_1, \dots, p + j_r).$$

(1.8) Additivity Formula. *In the setup of* (1.5), *with sequences I, J of nonnegative integers, let W' be a second word in the letters of A. Then,*

$$S^{I/J}(W' + W) = \sum_K S^{I/K}(W') S^{K/J}(W),$$

where the sum is over all strictly increasing sequences $K = (k_1, \dots, k_r)$.

Proof. It is clear from the definition of power series associated to words that $S(W' + W) = S(W')S(W)$. Therefore, the first formula of (1.2.3) yields the following matrix equation:

$$\begin{pmatrix} S(W'+W)[i_1] \\ \vdots \\ S(W'+W)[i_r] \end{pmatrix} = \begin{pmatrix} S(W')[i_1] \\ \vdots \\ S(W')[i_r] \end{pmatrix} \Big(S(W)^{0\geq}, S(W)^{1\geq}, \dots \Big).$$

From the matrix equation, extract the equation corresponding to the columns in J. The asserted additivity formula follows from the formula for the determinant of a product of matrices (the Cauchy–Binet Formula, see Appendix (1.2)). □

(1.9) Duality Formula. *Let W be a word in the letters of A, and let I and J be strictly increasing sequences of r nonnegative integers. Then*

$$S^{I/J}(W) = (-1)^{\|J\|-\|I\|} S^{I'/J'}(-W), \tag{1.9.1}$$

where the primes indicate the dual sequences, defined as follows: Choose an integer N greater than all entries of I and J. Identify I with a subset of $\{0, 1, \ldots, N-1\}$, and denote by I^c the complement of I in $\{0, 1, \ldots, N-1\}$. Then I', as a strictly increasing sequence, is the image of the complement under the reflection $i \mapsto N-1-i$.

Proof. Consider the power series $s := S(W)$ and $t := S(-W)$. As $ts = 1$, the following matrix equation results from the Equations (1.2.2) for $i, j = 0, \ldots, N-1$:

$$\begin{pmatrix} t[0] \\ \vdots \\ t[N-1] \end{pmatrix} \left(s^{0\geq}, \ldots, s^{N-1\geq} \right) = 1,$$

where the right hand side is the $N \times N$ unit matrix. In the second factor on the left hand side, only the first N rows are nonzero. Therefore, if S denotes the $N \times N$ matrix consisting of the first N rows of the second factor, and T denotes the $N \times N$ matrix consisting of the first N columns of the first factor, then we obtain the matrix equation $TS = 1$. Moreover, the matrix T has determinant 1, since it is an upper triangular matrix with $t_0 = 1$ in the diagonal.

Clearly, the two Schur functions $S^{I/J}(W)$ and $S^{I'/J'}(-W)$ of the Duality Formula are minors in the matrices S and T. Now, since $\det T = 1$, it is well known that the matrix equation $TS = 1$ implies that the (I, J)'th minor $S^{I/J}$ in S is equal to the algebraic complement of the (J, I)'th minor in T, that is,

$$S^{I/J} = \operatorname{sign}(II^c)\operatorname{sign}(JJ^c) T^{J^c/I^c}, \tag{1.9.2}$$

where II^c and JJ^c denote the concatenated sequences viewed as permutations of the sequence $(0, \ldots, N-1)$. Indeed, from $TS = 1$ and $\det T = 1$ it follows, see (APP.1.5.1), that S is the adjugate of T. Hence (1.9.2) follows from (APP.1.5.2).

Clearly, the left side of Equation (1.9.2) is the Schur function on the left side of the Duality formula. The right hand side of Equation (1.9.2) is easily transformed to the right hand side of the Duality Formula. Indeed, if $I = (i_1, \ldots, i_r)$, then the permutation II^c of the integers $0, \ldots, N-1$ has length equal to $i_1 + i_2 - 1 + \cdots + i_r - (r-1)$. Hence the product of the two signs on the right hand side of (1.9.2) is equal to $(-1)^{\|J\|-\|I\|}$. Moreover, the minor T^{J^c/I^c} on the right hand side of (1.9.2) is a minor in the matrix $\{t_{j-i}\}$. Therefore, T^{J^c/I^c} is equal T^{I^*/J^*} where I^* and J^* denote the images of the sequences I^c and J^c under the reflection $i \mapsto N-1-i$. Finally, the sequences I' and J' are obtained from I^* and J^* by reversing the order of the elements. The reversion of rows and columns does not change the determinant. Hence T^{I^*/J^*} is equal to the minor $T^{I'/J'}$ of T. Finally, since T was defined from the power series $t = S(-W)$, the minor $T^{I'/J'}$ is the Schur function $S^{I'/J'}(-W)$. Thus the Duality Formula has been proved. □

(1.10) Example. Take $I = (0, \dots, r-1)$ and $J = (1, \dots, r)$ (and $N := r+1$). Then $I' = (0)$ and $J' = (r)$. Hence $S^{I/J}(W) = (-1)^r S^r(-W)$ by the Duality Formula, and clearly $S^r(-W) = S_r(-W)$. In particular, when $W = A$, we obtain that the Schur function $S^{1,\dots,r}(A)$ is equal to the r'th elementary symmetric function in the letters of A, see (1.3).

(1.11) Jacobi–Trudi's Formula. *Let $a_1, \dots, a_n$ be the n letters of A, and let J be a sequence of n non-negative integers. Then,*

$$\Delta^J(a_1, \dots, a_n) = \Delta(a_1, \dots, a_n) S^J(A). \tag{1.11.1}$$

Moreover, if D is the $n \times n$ matrix whose k'th row is the ordered set of the n first homogeneous terms in the polynomial $\prod_{i \neq k}(1 - a_i)$, then $\det D = \Delta(a_1, \dots, a_n)$.

Proof. For $k = 1, \dots, n$, we have that $a_k = (a_k - A) + A$, and hence $S(a_k) = S(a_k - A)S(A)$. Consequently, the first equation in (1.2.3), applied with $p = S(a_k - A)$ for $k = 1, \dots, n$, yields the following matrix equation:

$$\begin{pmatrix} S(a_1)[0] \\ \vdots \\ S(a_n)[0] \end{pmatrix} = \begin{pmatrix} S(a_1 - A)[0] \\ \vdots \\ S(a_n - A)[0] \end{pmatrix} \Big(S(A)^{0\geq}, S(A)^{1\geq}, \dots \Big).$$

In the first matrix on the right, the power series $S(a_k - A)$ is the polynomial $\prod_{i \neq k}(1 - a_i)$ of degree $n - 1$. Hence, in the first matrix on the right, the first n columns form the matrix D and the remaining columns are equal to zero. Therefore, the product on the right side is unchanged if the first matrix is replaced by D and the second matrix is replaced by its first n rows. Now, from the replaced equation select the columns corresponding to the elements of J, and take determinants. On the left side we obtain the determinant $\Delta^J = \Delta^J(a_1, \dots, a_n)$, and on the right side we obtain the product $(\det D) S^J(A)$. Hence we obtain the equation,

$$\Delta^J = (\det D) S^J(A). \tag{1.11.2}$$

Take $J = (0, 1, \dots, n-1)$ in (1.11.2). The left side becomes the Vandermonde determinant Δ. On the right side we have that $S^J(A) = 1$. Therefore, Equation (1.11.2) implies first that $\Delta = \det D$ and next that the Jacobi–Trudi Formula (1.11.1) holds. ∎

(1.12) Corollary. *The Schur functions $S^J(A)$, for all strictly increasing sequences J of n non-negative integers, form an R-basis for the algebra* $\mathrm{Sym}_R[A]$ *of symmetric polynomials.*

Proof. By Jacobi–Trudi's formula, $S^J = \Delta^J/\Delta$, and the polynomials Δ^J/Δ form an R-basis by Corollary (SYM.6.7). ∎

(1.13) Remark. Every ordinary Schur function $S^K(A)$, where $K = (k_1, \dots, k_r)$ is a sequence with an arbitrary number r of entries, is either equal to zero or up to a sign equal to one of the Schur functions $S^J(A)$ of Corollary (1.12), that is, a Schur function defined by a strictly increasing sequence J with n entries. Indeed, since $S^K(A)$ is alternating in K, we may assume that the sequence K is strictly increasing. If $r < n$, we have that $S^K(W) = S^J(W)$, where $J = \tilde{K}$ is the sequence obtained by extending K as in Remark (1.7). Assume that $r > n$. If K is an extension in the sense of Remark (1.7) of a sequence J with n entries, then $S^K(A) = S^J(A)$.

If K is not an extension of a sequence with n entries, then $S^K(A) = 0$. Indeed, proceeding by induction on r, we may assume that K is not an extension of a sequence with $r-1$ entries, that is, we may assume that $k_1 > 0$. Consider an alphabet $\tilde{A}$ with r letters obtained by adding $r - n$ letters $a_{n+1}, \dots, a_r$ to A. Clearly, the Schur function $S^K(A)$ is obtained from the Schur function $S^K(\tilde{A})$ by specializing the additional variables $a_{n+1}, \dots, a_r$ to 0. Now, if the last variable a_r is specialized to 0, then the Vandermonde determinant $\Delta(\tilde{A})$ specializes to a regular element in the polynomial ring $R[a_1, \dots, a_{r-1}]$, and, since $k_1 > 0$, the determinant $\Delta^K(\tilde{A})$ specializes to 0. Therefore, by Jacobi–Trudi's Formula, the Schur function $S^K(\tilde{A})$ specializes to 0. Consequently $S^K(A) = 0$.

(1.14) Remark. Consider the symmetrization operator δ^A of Section (DIFF.2.3),

$$\delta^A(f) = \sum_{\sigma \in \mathfrak{S}(A)} \sigma\Big(\frac{f}{\Delta(A)}\Big) = \frac{1}{\Delta(A)} \sum_{\sigma \in \mathfrak{S}(A)} (\operatorname{sign} \sigma)\sigma(f).$$

The operator δ^A is $\mathrm{Sym}_R[A]$-linear. Let a^J be a monomial. It is elementary, see Theorem (DIFF.3.3), to prove that $\delta^A(a^J) = \Delta^J/\Delta$. Hence it follows from Jacobi–Trudi's Formula that $\delta^A(a^J) = S^J(A)$. As a consequence, if f is a symmetric polynomial, then $fS^J(A) = \delta^A(fa^J)$. In particular, if a symmetric polynomial f is given as a sum of monomials, $f = \sum_K f_K a^K$, then we obtain the formula

$$fS^J(A) = \sum_K f_K S^{K+J}(A).$$

(1.15) Example. *Pieri's Formula.* (1) Let I be a sequence of n non-negative integers, and denote by m^I the *monomial symmetric polynomial* corresponding to I, that is,

$$m^I := \sum_\sigma{}' a^{\sigma I},$$

where the sum is over all *different* permutations σI of the sequence I of exponents. Then, for every sequence J of n non-negative integers, we obtain from (1.14) the formula,

$$m^I S^J(A) = \sum_\sigma{}' S^{\sigma I + J}(A), \tag{1.15.1}$$

where the sum is over all different permutations σI of I. In particular, when J is the sequence $(0, 1, \dots, n-1)$ we have that $S^J(A) = 1$, and we obtain an explicit formula for m^I as a linear combination of Schur functions.

For the sequence $I = (1, \dots, 1, 0, \dots, 0)$ with $r \le n$ entries equal to 1, the polynomial m^I is the r'th *elementary symmetric function* $(-1)^r S_r(-A)$, and we obtain *Pieri's formula*,

$$(-1)^r S_r(-A) S^J(A) = \sum_K S^{K+J}(A), \tag{1.15.2}$$

where the sum is over all sequences K of n integers where r entries are equal to 1 and the remaining entries are equal to 0. For $J = (0, 1, \dots, n-1)$, we recover the formula of Example (1.10),

$$(-1)^r S_r(-A) = S^{1,\dots,r}(A). \tag{1.15.3}$$

Indeed, let $J = (0, 1, \dots, n-1)$. Then the left hand side of (1.15.2) is equal to the r'th elementary symmetric function since $S^J(A) = 1$. On the right hand side of (1.15.2), the sequence $K + J$ has two equal entries unless $K = (0, \dots, 0, 1, \dots, 1)$. Hence the only non-vanishing term on the right hand is the Schur function $S^{K+J}(A)$ where $K = (0, \dots, 0, 1, \dots, 1)$. Moreover, for $K = (0, \dots, 0, 1, \dots, 1)$ the sequence $K + J$ is an extension in the sense of (1.7) of the sequence $(1, \dots, r)$. Therefore, the right hand side of (1.15.2) is the Schur function $S^{1,\dots,r}(A)$.

(2) Similarly, the j'th *complete symmetric function* is the sum,

$$S_j(A) = \sum_{\|K\|=j} a^K,$$

where the sum is over size-n multi indices K. Hence we obtain from (1.14) the formula,

$$S_j(A)S^J(A) = \sum_{\|K\|=j} S^{K+J}(A). \tag{1.15.3}$$

(1.16) Notation. Most of the results in this section may also be found in (SYM.7) with a slightly different notation and a slightly different point of view.

Objects associated in this section to a positive word,

$$W = z_1a_1 + \dots + z_na_n,$$

like the series $S(W)$, the skew Schur function $S^{I/J}(W)$, the ordinary Schur function $S^I(W)$, etc., correspond to objects associated in (SYM.7) to the set of $|W|$ coordinates,

$$\alpha = (\underbrace{a_1, \dots, a_1}_{z_1 \text{ times}}, \dots, \underbrace{a_n, \dots, a_n}_{z_n \text{ times}}).$$

Via the natural inclusion $R[[A]] \subseteq R[A][[T]]$ (identifying the series $\sum f_d$ where f_d is homogeneous of degree d with the series $\sum f_d T^d$) the series $S(W) \in R[[A]]$ of (1.3) is equal to the series $s(\alpha) \in R[A][[T]]$ of (SYM.7.7). The skew Schur function $S^{I/J}(W)$ equals the skew Schur polynomial $s_I^J(\alpha)$ (the switch from I, J to J, I is due to the fact that the underlying matrix S of (1.2) is the transpose of the underlying matrix $M(s)$ of (SYM.7.2)). Finally, for the ordinary Schur functions/polynomials we have $S^J(W) = s^J(\alpha)$.

2. Multi Schur functions.

(2.1) Setup. In this section we consider simultaneously several alphabets. All alphabets are assumed to be subalphabets of a fixed (universal) alphabet. Letters and (commutative) words will be taken from the universal alphabet. As in (1.3), if A is an alphabet, the word defined as the sum of the letters of A will also be denoted by A. If $\{a_1, \dots, a_n\}$ are the letters of A, we write $A_{\le p} := a_1 + \dots + a_p$, and we define $A_{<p}$, $A_{\ge p}$, and $A_{>p}$ similarly.

(2.2) Definition. Let $\mathcal{W} = \{W_{pq}\}$ be an $r \times r$ matrix of words W_{pq}. Let I and J be sequences of r integers. Define the associated *multi skew Schur function* $S^{I/J}(\mathcal{W})$ as the determinant of the $r \times r$ matrix whose (p, q)'th entry is $S_{j_q - i_p}(W_{pq})$, that is,

$$S^{I/J}\begin{pmatrix} W_{11} & \dots & W_{1r} \\ \vdots & & \vdots \\ W_{r1} & \dots & W_{rr} \end{pmatrix} = \begin{vmatrix} S_{j_1-i_1}(W_{11}) & \dots & S_{j_r-i_1}(W_{1r}) \\ \vdots & & \vdots \\ S_{j_1-i_r}(W_{r1}) & \dots & S_{j_r-i_r}(W_{rr}) \end{vmatrix}. \tag{2.2.1}$$

As in Definition (1.5), when $I = (0, 1, \dots, r-1)$ we write $S^J(\mathcal{W}) := S^{I/J}(\mathcal{W})$ and obtain the ordinary *multi Schur function* $S^J(\mathcal{W})$. When all the words W_{pq} are equal to the same word W, we recover the Schur function $S^{I/J}(W)$ of (1.5).

Clearly, the multi skew Schur function is unchanged if the same integer is added to all entries of the two sequences I and J. In particular we will usually assume that the entries of the two sequences I and J are non-negative. Note also that the function $S^{I/J}(\mathcal{W})$ is alternating with respect to J and the columns of the matrix $\mathcal{W}$, and with respect to I and the rows of $\mathcal{W}$. Note finally, that the function is symmetric in the following sense: choose an integer N (say, greater than all elements in the two sequences I and J). Consider the sequences I^* and J^* obtained from I and J by the reflection $x \mapsto N - 1 - x$. Then $S^{I/J}(\mathcal{W}) = S^{J^*/I^*}(\mathcal{W}^{\mathrm{tr}})$, where $\mathcal{W}^{\mathrm{tr}}$ denotes the transposed matrix of $\mathcal{W}$.

Let $(W_1, \dots, W_r)$ be a sequence of r words in the letters of A. Then we denote by

$$S^{I/J}\begin{pmatrix} W_1 \\ \vdots \\ W_r \end{pmatrix} \quad \text{and} \quad S^{I/J}(W_1, \dots, W_r) \tag{2.2.2}$$

the two special functions defined as follows: The first function is the Schur function $S^{I/J}(\mathcal{W})$ obtained from the matrix $\{W_{pq}\}$ where $W_{pq} = W_p$ and the second function is defined similarly by the matrix where $W_{pq} = W_q$. Note that the two special Schur functions are maximal minors of the two matrices,

$$\begin{pmatrix} S(W_1)[i_1] \\ \vdots \\ S(W_r)[i_r] \end{pmatrix} \quad \text{and} \quad \big(S(W_1)^{j_1\ge}, \dots, S(W_r)^{j_r\ge}\big). \tag{2.2.3}$$

The first function is the minor of the first matrix corresponding to the column indices in J (assuming that all entries in J are nonnegative), the second function is the minor of the second matrix corresponding to the row indices in I (assuming that all entries in I are nonnegative).

When $I = (0, 1, \dots, r-1)$ we write $S^J(W_1, \dots, W_r)$ and $S^J(W_1, \dots, W_r)^{\mathrm{tr}}$ for the functions $S^{I/J}(W_1, \dots, W_r)$ and $S^{I/J}(W_1, \dots, W_r)^{\mathrm{tr}}$.

(2.3) Additivity Formula. *Let $(W_1', \dots, W_r')$ and $(W_1, \dots, W_r)$ be two sequences of r words, and let I and J be two sequences of r non-negative integers. Then,*

$$S^{I/J}\begin{pmatrix} W_1' + W_1 & \dots & W_1' + W_r \\ \vdots & & \vdots \\ W_r' + W_1 & \dots & W_r' + W_r \end{pmatrix} = \sum_K S^{I/K}\begin{pmatrix} W_1' \\ \vdots \\ W_r' \end{pmatrix} S^{K/J}(W_1, \dots, W_r),$$

where the summation is over all strictly increasing sequences K of r non-negative integers.

Proof. The argument is analogous to the proof of (1.8), once it is noted that the left side of the formula is the determinant of the matrix obtained as the product of the two matrices in (2.2.3), the first taken with $W_p := W_p'$. ∎

(2.4) Duality Formula. *Let $W_0, W_1, \dots$ be a sequence of words such that, for $j > 0$, the word $W_{j-1} - W_j$ is either equal to 0 or equal to a word formed by a single letter. Let I and J be strictly increasing sequences of r nonnegative integers. Finally, let N be an integer greater than all entries in the two sequences I and J. Then*

$$S^{I/J}(W_{j_1}, \dots, W_{j_r}) = (-1)^{\|J\|-\|I\|} S^{I'/J'}(-W_{j_t^c+1}, \dots, -W_{j_1^c+1}),$$

where the primes indicate the dual sequences with respect to N as in (1.9) and $J^c = (j_1^c, \dots, j_t^c)$ is the complement of J in $\{0, 1, \dots, N-1\}$.

Proof. The proof is similar to the proof of Duality (1.9). Consider the following product of matrices:

$$\begin{pmatrix} S(-W_1)[0] \\ \vdots \\ S(-W_N)[N-1] \end{pmatrix} \Big(S(W_0)^{0\geq}, \dots, S(W_{N-1})^{N-1\geq} \Big). \tag{1}$$

It follows from Formula (1.2.2) that the (i, j)'th entry in the product, for $i, j = 0, 1, \dots, N-1$, is equal to $S_{j-i}(W_j - W_{i+1})$. Clearly, the entry $S_{j-i}(W_j - W_{i+1})$ is equal to 0 for $j < i$ and equal to 1 for $j = i$. Moreover, the entry is equal to 0 for $j > i$, because, by hypothesis, the word $-(W_j - W_{i+1})$ is a positive word of degree at most $j - i - 1$. Hence the product of matrices (1) is the $N \times N$ unit matrix 1.

From the two matrices in the product (1), let T be the submatrix of the first factor formed by the first N columns, and let S be the submatrix of the second factor formed by the first N rows. As in the proof of the Duality Formula (1.9), it follows that $TS = 1$ and $\det T = 1$. It is a consequence, as we saw in the proof of the Duality Formula (1.9), that the (I, J)'th minor $S^{I/J}$ in S is equal to the algebraic complement of the (J, I)'th minor in T, that is,

$$S^{I/J} = (-1)^{\|J\|-\|I\|} T^{J^c/I^c}. \tag{2}$$

Clearly, the left side of Equation (2) is the Schur function on the left side of the asserted Duality Formula. The right hand side of Equation (2) is easily transformed to the right hand side of the Duality Formula. Indeed, the minor T^{J^c/I^c} on the right hand side of (2) is the Schur function,

$$S^{J^c/I^c}\begin{pmatrix} -W_{j_1^c+1} \\ \vdots \\ -W_{j_t^c+1} \end{pmatrix}.$$

To obtain the Schur function on the right side of the Duality Formula, apply the reflection $i \mapsto N-1-i$ to I^c and J^c and transpose the matrix, and then reverse the order of rows and columns.

Hence the Duality Formula is a consequence of Equation (2). ∎

(2.5) Corollary. *Let $A = \{a_1, \ldots, a_n\}$ be an alphabet with n letters, and denote by $A_{\le j}$ the sum of the letters a_p for $p \le j$ (in particular, $A_{\le j} = A$ for $j \ge n$). Moreover, let I and J be strictly increasing sequences of r nonnegative integers, and set $\lambda_p := j_p - (p-1)$ for $p = 1, \ldots, r$. Finally, let N be an integer greater than all entries of the two sequences I and J. Then, for any word W,*

$$S^{I/J}(W - A_{\le\lambda_1}, \ldots, W - A_{\le\lambda_r}) = (-1)^{\|J\|-\|I\|} S^{I'/J'}(A_{\le t} - W, \ldots, A_{\le 1} - W),$$

where the primes indicate the dual sequences with respect to N, and where $t = N - r$.

Proof. For any integer j, denote by $\mu(j)$ the number of elements in the complement J^c that are strictly less than j. Then the following equations hold:

$$\mu(j_k) = \lambda_k \quad\text{and}\quad \mu(j_k^c + 1) = k. \tag{1}$$

Indeed, the first equation holds because there are j_k nonnegative integers strictly less than j_k and of these exactly $k-1$ belong to J. The second equation holds, because the nonnegative integers in J^c that are strictly less than $j_k^c + 1$ are the k integers $j_1^c, \ldots, j_k^c$.

Now apply Duality (2.4) with $W_j := W - A_{\le\mu(j)}$. It follows from the equations (1) that $W_{j_k} = W - A_{\le\lambda_k}$ and $-W_{j_k^c+1} = A_{\le k} - W$. Hence the asserted formula follows from the Duality Formula. ∎

(2.6) Jacobi's Lemma. *Let $(W_1, \ldots, W_r)$ be a sequence of r words, and let $(C_1, \ldots, C_r)$ be a sequence of r positive words. Moreover, let I and J be sequences of r non-negative integers. If $|C_p| + i_p < r$ for all $p = 1, \ldots, r$, then the following formula holds:*

$$S^{I/J}\begin{pmatrix} W_1 - C_1 & \ldots & W_r - C_1 \\ \vdots & & \vdots \\ W_1 - C_r & \ldots & W_r - C_r \end{pmatrix} = S^{I/0,1,\ldots,r-1}\begin{pmatrix} -C_1 \\ \vdots \\ -C_r \end{pmatrix} S^J(W_1, \ldots, W_r).$$

In particular, if $|C_p| \le r - p$ for all $p = 1, \ldots, r$, then

$$S^{J}\begin{pmatrix} W_1 - C_1 & \ldots & W_r - C_1 \\ \vdots & & \vdots \\ W_1 - C_r & \ldots & W_r - C_r \end{pmatrix} = S^J(W_1, \ldots, W_r).$$

Proof. To prove the first assertion, assume that $|C_p| + i_p < r$ for all p. Apply the Additivity Formula (2.3) with $W'_p := -C_p$. The Schur functions $S^{I/K}$ in the sum on the right hand side of the Additivity Formula are the maximal minors of the matrix whose p'th row is $S(-C_p)[i_p]$ for $p = 1, \dots, r$. It follows from the hypothesis on the degree of C_p that only the first r entries in the p'th row can be nonzero. Hence the sum on the right hand side of (2.3) reduces to its single term corresponding to $K = (0, 1, \dots, r-1)$. Clearly, the latter term is the product on the right hand side of the asserted formula. Thus the first assertion holds.

Assume in particular that $I = (0, \dots, r-1)$. Then the first factor on the right side of the first formula is the determinant of an upper triangular matrix with 1 in the diagonal. Hence the first assertion implies the second. ∎

(2.7) Example. *The Jacobi–Trudi Formula.* When $I = (0, \dots, 0)$, the condition for Lemma (2.6) is that $|C_p| < r$ for all p. For example, let $A = \{a_1, \dots, a_n\}$ be an alphabet with n letters, and let $r = n$. Take $W_p := A$ and $C_p := A - a_p$ in Jacobi's Lemma. Then we obtain the formula,

$$S^{0,\dots,0/J}\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = S^{0,\dots,0/0,1,\dots,n-1}\begin{pmatrix} a_1 - A \\ \vdots \\ a_n - A \end{pmatrix} S^J(A, \dots, A).$$

The Schur function on the left side is the determinant $\Delta^J(a_1, \dots, a_n)$. On the right side, the second Schur function is the simple Schur function $S^J(A)$, and the first Schur function is the determinant $\det D$ from Jacobi–Trudi's Formula (1.11). Hence we recover Jacobi–Trudi's Formula (or, more precisely, (1.11.2)) from (2.6) (and we recover the proof from the proof of (2.6)).

(2.8) Factorization Formula. *Let $(a_1, \dots, a_r)$ be a sequence of r letters, and set $A_{\ge p} := a_p + \cdots + a_r$ for $p = 1, \dots, r$. Moreover, let J be a sequence of r non-negative integers, and set $\lambda_p := j_p - (p-1)$ for $p = 1, \dots, r$. Finally, let $B_1, \dots, B_r$ be a sequence of r positive words. Assume that $|B_p| \le \lambda_p$ for $p = 1, \dots, r$. Then*

$$S^J(A_{\ge 1} - B_1, \dots, A_{\ge r} - B_r) = \prod_{p=1}^{r} S_{\lambda_p}(a_p - B_p). \tag{2.8.1}$$

Proof. Apply the second assertion in Jacobi's Lemma (2.6) with $W_p := A_{\ge p} - B_p$ and $C_p := A_{>p}$. It follows that the Schur function on the left hand side of (2.8.1) is equal to the following Schur function:

$$S^J\begin{pmatrix} A_{\ge 1} - A_{>1} - B_1 & \dots & A_{\ge r} - A_{>1} - B_r \\ \vdots & & \vdots \\ A_{\ge 1} - A_{>r} - B_1 & \dots & A_{\ge r} - A_{>r} - B_r \end{pmatrix}. \tag{1}$$

The latter Schur function is the determinant of a matrix whose (p, q)'th entry is equal to

$$S_{j_q - (p-1)}(A_{\ge q} - A_{>p} - B_q). \tag{2}$$

Consider an entry above the diagonal, that is, for $p < q$. Clearly, the word $A_{\geq q} - A_{>p} - B_q$ is equal to $-B_{pq}$, where B_{pq} the sum of the $q - p - 1$ letters a_k for $p < k < q$ and the word B_q. Thus B_{pq} is a positive word, and since $|B_q| \leq \lambda_q$, the degree of B_{pq} is strictly less than $\lambda_q + (q - p) = j_q - (p - 1)$. Hence the entry (2) vanishes above the diagonal.

Therefore the determinant (1) is the product of its diagonal entries. Clearly the p'th diagonal word is $a_p - B_p$, and so the diagonal entry is $S_{j_p-(p-1)}(a_p - B_p) = S_{\lambda_p}(a_p - B_p)$. Thus the asserted formula has been proved. □

(2.9) Remark. In the setup of Lemma (2.8), assume that B_p is the word of an alphabet with β_p letters. Then $\beta_p \leq \lambda_p$ by assumption. Therefore, by Remark (1.4), the p'th factor on the right hand side of (2.8.1) is the product,

$$a_p^{\lambda_p - \beta_p} \prod_{b \in B_p} (a_p - b).$$

In particular, if all the alphabets B_p are empty, then the right hand side of (2.8.1) is the product, $a^\lambda = a_1^{\lambda_1} \cdots a_r^{\lambda_r}$, and we obtain the formula,

$$S^J(A_{\geq 1}, \dots, A_{\geq r}) = a_1^{j_1} a_2^{j_2 - 1} \cdots a_r^{j_r - (r-1)}.$$

(2.10) Example. Let $A = \{a_1, \dots, a_n\}$ be an alphabet. Then the determinant $\Delta^J(a_1, \dots, a_n)$ defined in (SYM.6.3) is the multi Schur function,

$$\Delta^J(a_1, \dots, a_n) = S^{0,\dots,0/J} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}. \tag{1}$$

Assume that $J = (0, 1, \dots, n-1)$. Then the determinant Δ^J is the Vandermonde determinant $\Delta(a_1, \dots, a_n)$. It follows by the symmetry of (2.2), or directly, that the Schur function in (1) is equal to $S^{n-1,\dots,n-1}(a_n, \dots, a_1)$, and hence,

$$\Delta(a_1, \dots, a_n) = S^{n-1,\dots,n-1}(A_{\leq n} - A_{<n}, \dots, A_{\leq 1} - A_{<1}). \tag{2}$$

Thus, from the Factorization Formula (2.8), applied to the reversed sequence $(a_n, \dots, a_1)$ and $B_i := A_{<n-i+1}$, we obtain the factorization,

$$\Delta(a_1, \dots, a_n) = S_{n-1}(a_n - A_{<n}) \cdots S_1(a_2 - A_{<2}) S_0(a_1).$$

By Remark (1.4), the latter factorization is the equation of (SYM.6.7),

$$\Delta(a_1, \dots, a_n) = \prod_{p>q} (a_p - a_q).$$

(2.11) Lemma. *Let $(W_1, \dots, W_r)$ be a sequence of r words, and let I and J be sequences of r non-negative integers. Assume for integers $0 \le d < k \le r$ that the following holds: $W_{k-q} = W_k$ and $j_{k-q} = j_k - q$ for $q = 1, \dots, d$. Then the Schur function $S^{I/J}(W_1, \dots, W_r)$ is unchanged if W_k is replaced by $W_k - B$, where B is any positive word of degree at most d.*

Proof. Set $j := j_{k-d}$. Set $W := W_k$ and $s := S(W)$. By hypothesis, $W_q = W$ and $j_q = j + (q - k + d)$ for $k - d \le q \le k$.

By definition, the Schur function $S^{I/J}(W_1, \dots, W_r)$ is the maximal minor with the row numbers in I in the matrix,

$$\big(S(W_1)^{j_1 \ge}, \dots, S(W_r)^{j_r \ge}\big). \tag{1}$$

The following submatrix of (1), formed by the $d+1$ columns with indices $q = k-d, \dots, k$:

$$\big(s^{j\ge}, s^{j+1\ge}, \dots, s^{j+d\ge}\big), \tag{2}$$

is also a submatrix with consecutive columns of the matrix S in (1.2). If the k'th word W is replaced by $W - B$, then the series $s = S(W)$ in the last column in (2) is replaced by $S(W - B) = sp$, where $p := S(-B)$. By the second equation of (1.2.3) we have $(sp)^{j+d\ge} = S\,p^{j+d\ge}$, and since $p = S(-B)$ is a polynomial of degree at most d (with $p_0 = 1$), it follows that

$$(sp)^{j+d\ge} = p_d s^{j\ge} + p_{d-1} s^{j+1\ge} + \dots + s^{j+d\ge}.$$

Therefore, the replacement results in the following change in (2): to the last column is added a linear combination of the first d columns. For (1), the result is that to the k'th column is added a linear combination of the previous d columns. Consequently, the replacement does not change the maximal minors of (1). ∎

(2.12) Example. Let $A = \{a_1, \dots, a_n\}$ and $B = \{b_1, \dots, b_m\}$ be alphabets with n and m letters. As an application of Lemma (2.11) we will prove the formula,

$$S^{m,m+1,\dots,m+n-1}(A - B) = \prod_{p,q}(a_p - b_q). \tag{2.12.1}$$

Set $A_i := a_1 + \dots + a_i$, and let W be any word. Consider for $r = n$ and $J = (m, m+1, \dots, m+n-1)$ the Schur function $S^J(W) = S^J(W, \dots, W)$. Then, by repeated application of Lemma (2.11), it follows that we can replace in the Schur function $S^J(W, \dots, W)$, for $k = n, \dots, 2$, the k'th word W by $W - A_{k-1}$. Hence we obtain the equation (where $A_0 = 0$),

$$S^{m,m+1,\dots,m+n-1}(W) = S^{m,m+1,\dots,m+n-1}(W - A_0, W - A_1, \dots, W - A_{n-1}).$$

In particular, for $W := A - B$ we obtain the equation,

$$S^{m,m+1,\dots,m+n-1}(A - B) = S^{m,m+1,\dots,m+n-1}(A_{\ge 1} - B, \dots, A_{\ge n} - B).$$

Finally, by applying the Factorization Formula to the Schur function on the right hand side, cf. Remark (2.9), we obtain the formula (2.12.1).

3. Differenciation of Schur functions.

(3.1) Setup. Keep the setup of (2.1). In particular, the alphabet A with n letters $a_1, \dots, a_n$ and the words in the following are taken from the fixed (universal) alphabet. The series $S(W)$ associated to a word is considered in the power series ring over the universal alphabet. Consider, with respect to the alphabet A, the simple operators ∂^p, π^p and the general operators $\partial^\omega, \pi^\omega, \partial_\omega$, and π_ω of Section (DIFF.3) where $\omega = \omega_A$ is the maximal permutation of A. By (DIFF.3.6), $\partial_\omega = \omega\partial^\omega\omega$ and $\pi_\omega = \omega\pi^\omega\omega$.

(3.2) Note. We proved in (DIFF.3.4) that the simple operators satisfy the Coxeter-Moore relations, and in (DIFF.3.5) that the general operators can be defined by any reduced presentation of ω. The two following equations are consequences since $\omega = \varepsilon_1 \cdots \varepsilon_{n-1}\bar{\omega}$ and $\ell(\omega) = n-1 + \ell(\bar{\omega})$ ($\bar{\omega}$ is the maximal permutation of $\bar{A} := \{a_1, \dots, a_{n-1}\}$). Here we may take the two equations as inductive definitions:

$$\partial^\omega = \partial^1 \cdots \partial^{n-1}\partial^{\bar{\omega}}, \quad \pi^\omega = \pi^1 \cdots \pi^{n-1}\pi^{\bar{\omega}}.$$

(3.3) Lemma. *Fix an integer p such that $1 \le p < n$. Let W be a word which is symmetric with respect to the letters a_p and a_{p+1}. Then we have the following two identities of power series:*

$$\partial^p S(W - a_p) = S(W),$$
$$\pi^p S(W - a_p) = S(W).$$

Proof. Clearly, $S(W - a_p) = S(W)(1 - a_p)$, and the power series $S(W)$ is symmetric in the letters a_p and a_{p+1}. The operators ∂^p and π^p are linear with respect to polynomials and series symmetric in a_p and a_{p+1}. Therefore the two equations of the Lemma result from the following equations,

$$\partial^p(1 - a_p) = 1, \quad \pi^p(1 - a_p) = 1.$$

The latter equations result immediately from the definitions of ∂^p and π^p. □

(3.4) Lemma. *Consider a Schur function $S^J(W_1, \dots, W_r)$, where J is a sequence of r integers. Fix positive integers $k \le r$ and $p < n$. Assume that all the words W_q are symmetric with respect to the letters a_p and a_{p+1}. Then,*

$$\partial^p S^J(W_1, \dots, W_k - a_p, \dots, W_r) = S^{j_1, \dots, j_k - 1, \dots, j_r}(W_1, \dots, W_k, \dots, W_r), \tag{3.4.1}$$
$$\pi^p S^J(W_1, \dots, W_k - a_p, \dots, W_r) = S^{j_1, \dots, j_k, \dots, j_r}(W_1, \dots, W_k, \dots, W_r). \tag{3.4.2}$$

Proof. The Schur function on the left hand sides of the equations is the determinant of a matrix whose k'th column is the sequence of homogeneous terms $S_{j_k - i}(W_k - a_p)$ for $i = 0, \dots, r-1$ and where the entries of the remaining columns are symmetric in a_p and a_{p+1}. Since the operators ∂^p and π^p are linear with respect to polynomials symmetric in a_p and a_{p+1}, the left sides are therefore equal to the determinants obtained by applying the operators to the entries of the k'th column.

Consider the first equation. Since ∂^p lowers the degree of polynomials by 1, it follows from Lemma (3.3) that $\partial^p S_l(W_k - a_p) = S_{l-1}(W_k)$. Clearly, the first equation of the Lemma is a consequence.

Similarly, it follows from Lemma (3.3) that $\pi^p S_l(W_k - a_p) = S_l(W_k)$, and consequently the second equation of the Lemma holds. □

(3.5) Differenciation Lemma. *Consider for $n \leq r$ the Schur function $S^J(W_1, \dots, W_r)$, where J is a sequence of r integers. Assume that all the words W_q are symmetric in the letters of A. Then the following two formulas hold:*

$$\partial^\omega S^J(W_1 - A_{<1}, \dots, W_n - A_{<n}, W_{n+1}, \dots, W_r) = S^{J-E}(W_1, \dots, W_r),$$
$$\pi^\omega S^J(W_1 - A_{<1}, \dots, W_n - A_{<n}, W_{n+1}, \dots, W_r) = S^J(W_1, \dots, W_r),$$

where $J - E$ is the sequence obtained from J by subtracting $k-1$ from j_k for $k = 1, \dots, n$.

Proof. The assertion will be proved by induction on n. The formulas have no content when $n = 1$. So we may assume that $n > 1$ and that the assertion holds for the alphabet $\bar{A} := \{a_1, \dots, a_{n-1}\}$.

In the Schur functions on the left sides of the equations, the n'th word $W_n - A_{<n}$ is symmetric with respect to the letters of $\bar{A}$. Therefore, by the inductive hypothesis and the inductive definition of ∂^ω in (3.2), the left hand side of the first equation is equal to the expression,

$$\partial^1 \cdots \partial^{n-1} S^{J-\bar{E}}(W_1, \dots, W_{n-1}, W_n - \bar{A}, W_{n+1}, \dots, W_r), \tag{1}$$

where $J - \bar{E}$ is the sequence obtained from J by subtracting $k-1$ from j_k for $k = 1, \dots, n-1$. Next, apply to the expression (1) $n-1$ times the equation (3.4.1) for $k = n$ for $p = n-1, \dots, 1$ and, successively, $W_n - a_1 - \cdots - a_p$ as the n'th word on the left side of (3.4.1). As a consequence, the expression (1) is equal to the right hand side of the first formula. Hence the first formula holds.

The proof of the second formula is entirely similar. □

(3.6) Remark. Conjugation by ω yields $\partial_\omega = \omega\partial^\omega\omega$ and $\pi_\omega = \omega\pi^\omega\omega$, see (3.1). As a consequence, we obtain from (3.5):

$$\partial_\omega S^J(W_1 - A_{>n}, \dots, W_n - A_{>1}, W_{n+1}, \dots, W_r) = S^{J-E}(W_1, \dots, W_r),$$
$$\pi_\omega S^J(W_1 - A_{>n}, \dots, W_n - A_{>1}, W_{n+1}, \dots, W_r) = S^J(W_1, \dots, W_r).$$

Indeed, apply the permutation ω to the first equation of (3.5). The right side is unchanged, because it is symmetric in the letters of A. On the left side, the result is the operator $\omega\partial^\omega = \partial_\omega\omega$ applied to the Schur function $S^J(W_1 - A_{<1}, \dots, W_n - A_{<n}, W_{n+1}, \dots, W_r)$. As ω changes the word $W_k - A_{<k}$ into the word $W_k - A_{>n-k+1}$, we obtain the first formula asserted above. The verification of the second formula is completely analogous.

Using the fact that the Schur function $S^J(W_1, \dots, W_r)$ is alternating in J and the words W_q we get analogous formulas when the n words $A_{<1}, \dots, A_{<n}$ are subtracted from any n different words in the sequence $(W_1, \dots, W_r)$. For example, from the formulas for π^ω in (3.5) and π_ω above, we obtain the formulas (where $t := r - n$),

$$\pi^\omega S^J(W_1, \dots, W_t, W_{t+1}-A_{<1}, \dots, W_r - A_{<n}) = S^J(W_1, \dots, W_r), \tag{3.6.1}$$

$$\pi_\omega S^J(W_1, \dots, W_t, W_{t+1}-A_{>n}, \dots, W_r - A_{>1}) = S^J(W_1, \dots, W_r). \tag{3.6.2}$$

(3.7) Note. Let J be a sequence of n non-negative integers and set $E := (0, 1, \dots, n-1)$. Then the following two formulas hold:

$$\partial^\omega(a^J) = S^J(A), \tag{3.7.1}$$

$$\pi^\omega(a^J) = S^{J+E}(A). \tag{3.7.2}$$

Indeed, we have that $A_{\geq q} = A - A_{<q}$. Therefore, it follows from the Factorization Formula (2.8), applied to the sequence $J + E$ and the words $B_q := 0$, that

$$S^{J+E}(A - A_{<1}, \dots, A - A_{<n}) = a^J.$$

Hence the two asserted formulas follow from the Differenciation Lemma (3.5).

Note that, by the formulas of Theorem (DIFF.3.3), the left side of (3.7.1) is equal to Δ^J/Δ. Hence the formula (3.7.1) also follows from Theorem (DIFF.3.3) and Jacobi–Trudi's Lemma (1.11). The formula (3.7.2) also follows from the (3.7.1) and the second operator equation of (DIFF.3.3.3).

(3.8) Sergeev–Pragacz's Formula. *Let $B = \{b_1, \dots, b_m\}$ be a second alphabet, disjoint from A. Let J be a strictly increasing sequence of r non-negative integers. Set $\lambda_k := j_k - (k-1)$ for $k = 1, \dots, r$. Assume that $r \geq n$, where n is the number of letters of A, and set $t := r - n$. Then the following formula holds,*

$$S^J(A - B) = \pi^{\omega_A}\pi_{\omega_B}\Big(\prod_{k=1}^{t} S_{\lambda_k}(-B_{\leq\lambda_k}) \prod_{k=t+1}^{t+n} S_{\lambda_k}(a_{k-t} - B_{\leq\lambda_k})\Big).$$

Proof. Form the dual sequence J' with respect to an integer N, see (1.9). By choosing N large we may assume that the dual sequence J' has at least m elements. Consider the Schur function $S^J(A-B)$ on the left side of the asserted formula. By Duality (1.9),

$$S^J(A-B) = \varepsilon S^{J'}(B-A) = \varepsilon S^{J'}(B-A, \dots, B-A), \tag{1}$$

where $\varepsilon = (-1)^{\|J\|-r(r-1)/2}$. Now apply the Differenciation Lemma, in the form of Equation (3.6.2), to the alphabet B and the sequence of equal words $W_q := B-A$. We obtain the formula,

$$S^{J'}(B-A, \dots, B-A) = \pi_{\omega_B} S^{J'}(B-A, \dots, B-A, B_{\leq m}-A, \dots, B_{\leq 1}-A). \tag{2}$$

By Duality (2.5), applied to the letters of B and $W := A$ (note that $N - r \geq m$),

$$\varepsilon S^{J'}(B-A, \ldots, B-A, B_{\leq m}-A, \ldots, B_{\leq 1}-A) = S^{J}(A-B_{\leq \lambda_1}, \ldots, A-B_{\leq \lambda_r}), \quad (3)$$

with ε as before. Apply again the Differenciation Lemma, in the form of (3.6.1), to the alphabet A and the words $W_q := A-B_{\leq \lambda_q}$. We obtain the equation,

$$\begin{aligned} &S^{J}(A-B_{\leq \lambda_1}, \ldots, A-B_{\leq \lambda_r}) \\ &\qquad = \pi^{\omega_A} S^{J}(A-B_{\leq \lambda_1}, \ldots, A-B_{\leq \lambda_t}, A_{\geq 1}-B_{\leq \lambda_{t+1}}, \ldots, A_{\geq n}-B_{\leq \lambda_r}). \quad (4) \end{aligned}$$

Finally, the Schur function on the right hand side of (4) is equal to the product on the right hand side of Sergeev–Pragacz's Formula, as it follows by applying the Factorization Formula, see Remark (2.9), to a sequence of letters $(a'_1, \ldots, a'_t, a_1, \ldots, a_n)$ and then specializing the additional letters a'_q to zero. Therefore, the asserted formula follows from Equations (1)–(4). ∎

(3.9) Note. The condition for the Schur function $S^{J}(A-B)$ that the sequence J has at least n elements can always be obtained by *extending* the sequence J, cf. (1.7).

Clearly, the factor $S_{\lambda_k}(-B_{\leq \lambda_k})$ in the first product on the right hand side is only non-zero if the word $B_{\leq \lambda_k}$ has λ_k letters, that is, if λ_k is less than or equal to the number m of letters of B. Hence the Schur function $S^{J}(A-B)$ is only non-zero if $\lambda_t \leq m$, that is, if $j_t \leq m + t - 1$. Sequences J satisfying the latter condition correspond to partitions $\mu_1 \geq \cdots \geq \mu_r$ ($\mu_i = \lambda_{r-i+1}$) "contained in the (n, m)-*hook*": $\mu_i \leq m$ for $i > n$.

Assume that $\lambda_{t+1} \geq m$, that is, $j_{t+1} \geq t + m$. Then the last product in Sergeev–Pragacz's formula is equal to the following product,

$$\prod_{p=1}^{n} a_p^{\lambda_{t+p}-m} \prod_{a \in A,\, b \in B} (a-b).$$

In particular, when $J = (m, m+1, \ldots, m+n-1)$, we recover the formula of Example (2.12),

$$S^{m,m+1,\ldots,m+n-1}(A-B) = \prod_{a \in A,\, b \in B} (a-b),$$

since $\pi^{\omega_A}(1) = 1$ and $\pi_{\omega_B}(1) = 1$.

Appendix

1. On determinants.

(1.1) Setup. Work with square matrices, of size $n \geq 1$ if not otherwise specified. The unit matrix is denoted 1_n. Fix an $n \times n$ matrix $U = (u^i_j)$. Its determinant is the signed sum,

$$\det U = \sum_{\mu \in \mathfrak{S}_n} (-1)^{\ell(\mu)} u^1_{\mu_1} \cdots u^n_{\mu_n}. \tag{1.1.1}$$

In the products, the superscripts refer to the n row indices, assumed in (1.1.1) to be $1, 2, \dots, n$. The sum is over the set of bijective maps from the set of the n row indices to the set of the n column indices, indicated by the appearence of the symmetric group $\mathfrak{S}_n$.

The i'th row of U is denoted U^i, the j'th column is denoted U_j. If $J = (j_1, \dots, j_r)$ is any sequence of column indices, of length r, we denote by U_J the $n \times r$ matrix with the columns $U_{j_1}, \dots, U_{j_r}$. A subset J of r of the column indices will be identified with the corresponding strictly increasing sequence, but in general the sequence J is not assumed to be increasing and its entries are not assumed to be different. We use a parallel notation U^I for a sequence I of row indices. In particular, for two sequences I and J of the same length r the matrix U^I_J is a square matrix of size r. An *r-minor* of U is a determinant $\det U^I_J$ given by strictly increasing sequences (or subsets) I and J of size r of, respectively, row and column indices. The *complement* of the minor is the $(n-r)$-minor $\det U^{\tilde I}_{\tilde J}$ where $\tilde I$ and $\tilde J$, respectively, denote the complements of I and J in the sets of row indices and column indices.

Usually, row and column indices are taken from the same finite totally ordered set of indices, of cardinality n, in fact from the set $\{1, \dots, n\}$. For any subset I of indices we denote by $\tilde I$ the complementary set of indices. For $0 \leq r \leq n$ we denote by $[r]$ the subset of the first r indices, and thus $[r]^\sim$ is the subset of the last $n - r$ indices.

(1.2) The Cauchy–Binet Formula. *Let U and V be matrices of size, respectively, $r \times n$ and $n \times r$, where $r \leq n$. Then we have the formula,*

$$\det(UV) = \textstyle\sum_{|J|=r} \det U_J \, \det V^J \,, \tag{1.2.1}$$

where the sum is over subsets $J \subseteq [n]$.

Proof. We use two well known facts about the characteristic polynomial. For an $n \times n$ matrix A, set

$$\chi_A(t) := \det(A + t1_n) \in R[t].$$

Then $\chi_A(t)$ is, up to the substitution $t := -t$, the characteristic polynomial of A.

Fact 1. *If $c_r(A)$ denotes the coefficient to t^{n-r} in $\chi_A(t)$, then $c_r(A)$ is the sum of the principal r-minors of A:*

$$c_r(A) = \sum\nolimits_{|J|=r} \det A_J^J.$$

Fact 2. *If A and B are square matrices of size n, then $\chi_{AB}(t) = \chi_{BA}(t)$.*

Fact 1 follows easily from the definition of the determinant. The equation in Fact 2 follows easily when B is invertible, since $AB + t1_n$ and $BA + t1_n = B(AB + t1_n)B^{-1}$ have the same determinant. Next, when the determinant $b := \det B$ is regular, the equation holds since it holds for the ring $R[b^{-1}]$. Finally, to verify the equation in general, consider the polynomial ring $R[X]$ as an extension of R. The determinant of $B + X1_n$ is the monic polynomial $\chi_B(X)$ which is regular in $R[X]$. Hence the equation holds for the matrices A and $B + X1_n$ over $R[X]$. The substitution $X := 0$ in the latter equation yields the equation for A and B.

To prove the Cauchy–Binet formula, add $n - r$ zero columns to the matrix V to obtain an $n \times n$ matrix denoted $(V\ 0)$. Use a similar notation for the $n \times n$ matrix obtained from U by adding $n - r$ zero rows. Then we have the two equations of $n \times n$ matrices:

$$\begin{pmatrix} U \\ 0 \end{pmatrix} (V\ 0) = \begin{pmatrix} UV & 0 \\ 0 & 0 \end{pmatrix}, \qquad (V\ 0)\begin{pmatrix} U \\ 0 \end{pmatrix} = VU.$$

By Fact 2, the two right sides have the same "characteristic" polynomial. In particular, by Fact 1, the two sums of principal r-minors are the same. For the right side of the first equation, the only non-trivial r-minor is $\det(UV)$. Hence $\det(UV)$ is equal to the sum of the principal r-minors of the right side of the second equation, that is,

$$\det(UV) = \sum\nolimits_{|J|=r} \det(VU)_J^J.$$

As $\det(VU)_J^J = \det(V^J\, U_J) = \det V^J\, \det U_J = \det U_J\, \det V^J$, the asserted formula (1.2.1) is a consequence. □

(1.3) Laplace development. *For the given $n \times n$ matrix U and $0 \le r \le n$ we have the equation,*

$$\det U = \sum_{|J|=r} (-1)^{\ell(J\tilde{J})} \det U_J^{[r]} \det U_{\tilde{J}}^{[r]\tilde{}}, \tag{1.3.1}$$

where the sum is over all size-r subsets $J \subseteq [n]$.

Proof. For any subset $J \subseteq [n]$ of r indices denote by $\mathfrak{S}_J$ the subset of permutations $\mu \in \mathfrak{S}_n$ for which $\mu([r]) = J$. Then $\mathfrak{S}_n$ is the disjoint union of the subsets $\mathfrak{S}_J$. Accordingly, the expression for $\det U$ is a sum of partial sums,

$$\det U = \sum_{|J|=r} P_J, \qquad \text{where } P_J := \sum\nolimits_{\mu \in \mathfrak{S}_J} (-1)^{\ell(\mu)} u_{\mu_1}^1 \cdots u_{\mu_n}^n. \tag{1.3.2}$$

Set $s := n-r$, and let $\tilde{J}$ be the complement of J, say $J = (j_1, \dots, j_r)$ and $\tilde{J} = (k_1, \dots, k_s)$. Clearly, every permutation $\mu \in \mathfrak{S}_J$ is of the form,

$$\mu = (\mu_1, \dots, \mu_n) = (j_{\sigma_1} \dots, j_{\sigma_r}, k_{\tau_1}, \dots, k_{\tau_s}),$$

with uniquely determined permutations $\sigma \in \mathfrak{S}_r$ and $\tau \in \mathfrak{S}_s$. By a simple count of inversions,

$$\ell(\mu) = \ell(J\tilde{J}) + \ell(\sigma) + \ell(\tau).$$

Consequently,

$$P_J = (-1)^{\ell(J\tilde{J})} \sum_{\sigma,\tau} (-1)^{\ell(\sigma)} (-1)^{\ell(\tau)} u^1_{j_{\sigma_1}} \cdots u^r_{j_{\sigma_r}} u^{r+1}_{k_{\tau_1}} \cdots u^n_{k_{\tau_s}}. \tag{1.3.3}$$

Clearly, the sum over σ, τ is the product of two determinants,

$$\sum_{\sigma,\tau} (-1)^{\ell(\sigma)} (-1)^{\ell(\tau)} u^1_{j_{\sigma_1}} \cdots u^r_{j_{\sigma_r}} u^{r+1}_{k_{\tau_1}} \cdots u^n_{k_{\tau_s}} = \det U^{[r]}_J \, \det U^{[r]\tilde{}}_{\tilde{J}}. \tag{1.3.4}$$

Obviously, the asserted formula (1.3.1) follows from (1.3.2), (1.3.3), and (1.3.4). ∎

(1.4) Corollary. *For any size-r subset $I \subseteq [n]$ we have the development,*

$$\det U = \sum_{|J|=r} (-1)^{\ell(I\tilde{I})} (-1)^{\ell(J\tilde{J})} \det U^I_J \, \det U^{\tilde{I}}_{\tilde{J}}. \tag{1.4.1}$$

More generally, if $K \subseteq [n]$ is a second size-r subset, then (with Kronecker's delta),

$$(\det U)\, \delta^I_K = \sum_{|J|=r} (-1)^{\ell(K\tilde{K})} (-1)^{\ell(J\tilde{J})} \det U^I_J \, \det U^{\tilde{K}}_{\tilde{J}}. \tag{1.4.2}$$

Proof. For any two size-r subsets I, $K \subseteq [n]$ consider the $n \times n$ matrix $U^{I\tilde{K}}$ whose rows are the rows of U with the indices of the concatenated sequence $I\tilde{K}$. The concatenated sequence has two equal entries unless $\tilde{K}$ is the complement of I, that is, unless $I = K$.

Consider Laplace development of the matrix $U^{I\tilde{K}}$, and multiply by $(-1)^{\ell(K\tilde{K})}$ to obtain the following equation:

$$(-1)^{\ell(K\tilde{K})} \det U^{I\tilde{K}} = \sum_{|J|=r} (-1)^{\ell(K\tilde{K})} (-1)^{\ell(J\tilde{J})} \det U^I_J \det U^{\tilde{K}}_{\tilde{J}}. \tag{1.4.3}$$

Consider the left side of (1.4.3). If $I \neq K$ then the matrix $U^{I\tilde{K}}$ has two equal rows, and hence $\det U^{I\tilde{K}} = 0$. If $I = K$, then the rows of the matrix $U^{I\tilde{K}} = U^{K\tilde{K}}$ form a permutation of the rows of U, and $(-1)^{\ell(K\tilde{K})} \det U^{K\tilde{K}} = \det U$. Hence the left sides of (1.4.3) and (1.4.2) are equal. As the right sides are the same, the equality in (1.4.2) is a consequence.

Equation (1.4.1) is the special case $K = I$ of (1.4.2). ∎

(1.5) Jacobi's Theorem on the adjugate. *For the given $n \times n$ matrix U let ${}^{\dagger}U$ be the adjugate of U (the transposed cofactor matrix), with the entries,*

$$ {}^{\dagger}u^i_j := (-1)^{i+j} \det U^{j\tilde{}}_{i\tilde{}} . $$

Then the following equations hold:

$$ {}^{\dagger}U\, U = U\, {}^{\dagger}U = (\det U) 1_n. \tag{1.5.1} $$

Moreover, for $1 \le r \le n$, we have the formula for the r-minors of the adjugate matrix, for any two size-r subsets I, $J \subseteq [n]$,

$$ \det\bigl({}^{\dagger}U\bigr)^I_J = (-1)^{\ell(I\tilde{I})}(-1)^{\ell(J\tilde{J})}(\det U)^{r-1} \det U^{\tilde{J}}_{\tilde{I}}. \tag{1.5.2} $$

Proof. Define, for any two size-r subsets I, $J \subseteq [n]$:

$$ \tilde{u}^I_J := (-1)^{\ell(I\tilde{I})}(-1)^{\ell(J\tilde{J})} \det U^{\tilde{J}}_{\tilde{I}}. $$

With this notation, Equation (1.4.2) is the following, for size-r subsets I, $K \subseteq [n]$:

$$ (\det U)\, \delta^I_K = \sum\nolimits_{|J|=r} \det U^I_J\ \tilde{u}^J_K. \tag{1.5.3} $$

The size-r subsets of $[n]$ form a finite, totally ordered set with $N = \binom{n}{r}$ elements. Hence the determinants $\det U^I_J$ and the family $\tilde{u}^I_J$ may be viewed as $N \times N$ matrices, denoted, respectively, ${}^rU$ and ${}^r\widetilde{U}$. As such the set of Equations (1.5.3) for all size-r subsets I, K is equivalent to the following equation of $N \times N$ matrices:

$$ (\det U) 1_N = {}^rU\ {}^r\widetilde{U}, \tag{1.5.4} $$

where 1_N is the identity $N \times N$ matrix.

Applied with $r = 1$, the construction results in $n \times n$ matrices, and ${}^1U = U$; moreover, an easy count of inversions in the permutations $(i\ i\tilde{})$ and $(j\ j\tilde{})$ shows that that $\tilde{u}^i_j = {}^{\dagger} u^i_j$; hence ${}^1\tilde{U} = {}^{\dagger}U$ is the adjugate matrix. Thus (1.5.4) for $r = 1$ is the second equation of (1.5.1). The similar equation for ${}^{\dagger}U\, U$ may be proved similarly, or by applying the first equation to the transpose of U.

To prove Equation (1.5.2), consider the r-minors of the equation of $n \times n$ matrices $U\, {}^{\dagger}U = (\det U)1_n$ just proved. Clearly, we have $\det\, (1_n)^I_J = \delta^I_J$ and hence for the scalar matrix $(\det U)\, 1_n$,

$$ \det\bigl((\det U)(1_n)^I_J\bigr) = (\det U)^r \delta^I_J. $$

The corresponding r-minor of $U\, {}^{\dagger}U$ is the determinant of $(U\, {}^{\dagger}U)^I_J = U^I\, {}^{\dagger}U_J$. Hence, by the Cauchy–Binet Formula,

$$ (\det U)^r \delta^I_J = \det\bigl(U^I\, {}^{\dagger}U_J\bigr) = \sum\nolimits_{|K|=r} \det U^I_K \det {}^{\dagger}U^K_J\ . \tag{1.5.5} $$

Therefore, if ${}^{r\dagger}U$ is the $N \times N$ matrix with entries $\det {}^{\dagger}U^I_J$, we have, by (1.5.5) and (1.5.4), the matrix equations,

$$ {}^{r}U\, {}^{r\,\dagger}U = (\det U)^r 1_N, \qquad {}^{r}U\, {}^{r}\widetilde{U} = (\det U) 1_N. \tag{1.5.6} $$

Clearly, (1.5.2) is equivalent to the matrix equation ${}^{r\dagger}U = (\det U)^{r-1}\, {}^{r}\widetilde{U}$, and we verify the latter.

First, the matrix equation follows immediately from (1.5.6), if $\det U$ is invertible.

Next, if $u = \det U$ is regular, then the matrix equation holds, since we have just seen that it holds over the ring $R[u^{-1}]$.

Finally, to verify the equation in general, embed R in the polynomial ring $R[X]$. Then the matrix equation holds for the matrix $U + X1_n$. So the substitution $X := 0$ in the latter equation yields the equation for U. □

References

1. C. G. Jacobi, *De functionibus alternantibus earumque divisione per productum e differentiis elementorum conflatum*, J. Reine Angw. Math. **22** (1841), 360–371, Also in: “Werke **3**”, pp. 439–452..
2. I. G. Macdonald, *Symmetric functions and Hall polynomials*, Oxford Univ. Press, Oxford, 1979.
3. I.G. Macdonald, *Notes on Schubert polynomials*, Université de Québec à Montréal, 1991.

I. Index.