An Exposé of Lascoux's theory of divided differences and Symmetric polynomials

Preamble

The spectacular results recently obtained by Lascoux-Lapointe-Morse on the theory of Macdonald polynomials has brought to evidence in the most inequivocal manner that the theory developped by Lascoux in his notes on "Interpolation" is a most powerfull tool for the study of symmetric functions. This given no serious researchers in this field can any longer afford not to be familiar with the basic results of this theory. These notes are an expansion of the often terse (better yet sybilline) messages contained in Lascoux notes.

1. Divided differences

The n^{th} order "*divided difference*" of a function f with respect to alphabet $A = \{a_1, a_2, a_3, \ldots\}$, denoted by $\Delta_A^{(n)}(f)$, is defined by setting

$$\Delta_A^{(n)}(f) \;=\; \Delta_{a_1,a_2,\ldots,a_n}(f) \;=\; \sum_{s=1}^{n} \frac{f(a_s)}{\prod_{r=1, r\neq s}^{n}(a_s - a_r)} \qquad 1.1$$

This linear functional has the following fundamental properties:

Proposition 1.1

$$\begin{aligned} a)\quad & \Delta_A^{(n)}(x^k) = 0 \qquad \text{for} \quad 0 \le k \le n-2 \\ b)\quad & \Delta_A^{(n)}(x^k) = 1 \qquad \text{for} \quad k = n-1 \end{aligned} \qquad 1.2$$

Proof

Note that for any $0 \le k \le n-1$ we have a partial fraction expansion of the form

$$\frac{x^{n-k-1}}{\prod_{r=1}^{n}(1-a_r x)} \;=\; \sum_{s=1}^{n} A_s \,\frac{1}{1-a_s x} \quad . \qquad 1.3$$

Multiplying both sides by $1 - a_s x$ and letting $x \to 1/a_s$ yields that

$$A_s \;=\; \frac{a_s^k}{\prod_{r=1}^{n}(a_s - a_r)} \quad ,$$

and 1.3 becomes

$$\frac{x^{n-k-1}}{\prod_{r=1}^{n}(1-a_r x)} \;=\; \sum_{s=1}^{n} \frac{a_s^k}{\prod_{r=1}^{n}(a_s - a_r)}\,\frac{1}{1-a_s x} \quad . \qquad 1.4$$

This given, both 1.2 a) and b) are obtained by setting $x = 0$ on both sides of this equation.

An immediate consequence of this proposition is what is referred to in Lascoux notes as the "*Newton interpolation* " formula.

Theorem 1.1

For any polynomial $P(x)$ of degree less or equal to n we have

$$P(x) \;=\; P(a_1) \;+\; \sum_{k=1}^{n} \Delta_A^{(k+1)}(P)\,(x-a_1)(x-a_2)\cdots(x-a_k) \tag{1.5}$$

Proof

Since polynomials

$$1\,,\;(x-a_1)\,,\;(x-a_1)(x-a_2)\,,\;\cdots\,,\;(x-a_1)(x-a_2)\cdots(x-a_n)$$

form a basis for the polynomials of degree $\leq n$, and the right hand side of 2.4 is linear in P, the assertion will imediately follow if 1.5 is true for each and everyone of these polynomials. To this end note that the definition in 1.1 yields that

$$\Delta_A^{(k+1)}\big((x-a_1)(x-a_2)\cdots(x-a_s)\big) \;=\; 0 \qquad \text{for all } k<s\;. \tag{1.6}$$

On the other hand, 1.2 a) yields that we must have

$$\Delta_A^{(k+1)}\big((x-a_1)(x-a_2)\cdots(x-a_s)\big) \;=\; 0 \qquad \text{for all } k>s\;. \tag{1.7}$$

Finally, the linearity of the operator $\Delta_A^{(k+1)}$ combined with 1.2 a) yields that

$$\Delta_A^{(s+1)}\big((x-a_1)(x-a_2)\cdots(x-a_s)\big) \;=\; \Delta_A^{(s+1)}\big(x^s\big) \;=\; 1 \tag{1.8}$$

where the last equality is an instance of 1.2 b).

In summary we see that when $P(x)=(x-a_1)(x-a_2)\cdots(x-a_s)$ and $s\leq n$ the identities in 1.6, 1.7 and 1.8 reduce 1.5 to the assertion that

$$(x-a_1)(x-a_2)\cdots(x-a_s) \;=\; (x-a_1)(x-a_2)\cdots(x-a_s)$$

Finally, the fact that 1.5 is also true when $P=1$ is an immediate consequence of 1.2 a) for $k=0$. This completes our proof of 1.5

In Lascoux's notes, the determinant

$$W_A(f_1,f_2,\ldots,f_n) \;=\; det\,\|\,f_i(a_1)\,,\;\Delta_{a_1,a_2}(f_i)\,,\;\cdots\,,\;\Delta_{a_1,a_2,\ldots,a_{n-1}}(f_i)\,\|_{i=1}^{n} \tag{1.9}$$

is referred to as the "*Wronskian* " of the functions $f_1,f_2,\ldots,f_n$.

A further immediate consequence of the definition in 1.1 is the following basic identity.

Theorem 1.2

For any n^{tuple} of functions $f_1,f_2,\ldots,f_n$ we have

$$W_A(f_1,f_2,\ldots,f_n) \;=\; \frac{(-1)^{\binom{n}{2}}}{\Delta(A_n)}\,det\,\|f_i(a_j)\|_{i,j=1}^{n} \tag{1.10}$$

where $\Delta(A_n)$ denotes the Vandermonde determinant in $a_1,a_2,\ldots,a_n$ that is

$$\Delta(A_n) \;=\; \prod_{1\leq i<j\leq n}(a_i-a_j) \tag{1.11}$$

Proof

We see from the definition in 1.1 that the determinant in 1.9 will not change if the j^{th} element of the i^{th} row of the matrix is replaced by

$$\frac{f_i(a_j)}{\prod_{r=1}^{j-1}(a_j - a_r)} \ .$$

But if we do this all the elemnts of the j^{yh} column have the common factor

$$\frac{1}{\prod_{r=1}^{j-1}(a_j - a_r)}$$

which may be extracted yielding that

$$W_A(f_1, f_2, \ldots, f_n) \ = \ \frac{1}{\prod_{j=2}^{n}\prod_{r=1}^{j-1}(a_j - a_r)} \, det \, \| f_i(a_j) \|_{i,j=1}^{n}$$

and this is simply another way of writing 1.10.

2. Brief review of plethystic notation

The connection between divided differences and symmetric polynomials is best understood by working out a few examples. The latter are best expressed by means of pleththystic notation. To this end let us recall that the plethystic substitution of a rational expression

$$E \ = \ E(t_1, t_2, t_3 \ldots)$$

in any number of variables $t_1, t_2, t_3 \ldots$ into the power symmetric function p_k, denoted $p_k[E]$, is simply defined by setting

$$p_k[E] \ = \ E[t_1^k, t_2^k, t_3^k, \ldots] \ . \qquad 2.1$$

This operation extends to any power basis element

$$p_\lambda \ = \ p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_k}$$

by setting

$$p_\lambda[E] \ = \ p_{\lambda_1}[E] p_{\lambda_2}[E] \cdots p_{\lambda_k}[E] \qquad 2.2$$

This given, we may extend this operation to any symmetric function by first expressing it in terms of the power basis and then using 2.2. With this device all the familiar identities of symmetric function theory will immediately extend to a variety of new identities. It should be noted that the operations in 2.1 and 2.2 are easily implemented on the computer so that experimentation becomes possible in a large scale.

Given an alphabet $A = \{a_1, a_2, \ldots\}$ we set

$$\Omega[A] \ = \ \prod_{a \in A} \frac{1}{1-a} \ . \qquad 2.3$$

Note that we also have

$$\Omega[A] \;=\; exp\Big(\sum_{k\geq 1} \frac{p_k[A]}{k}\Big) \qquad 2.4$$

with

$$p_k[A] \;=\; \sum_{a\in A} a^k \;. \qquad 2.5$$

Now we see that all of these relations may be viewed as the plthystic substitution of the sum $A = a_1 + a_2 + \cdots$ into the symmetric function

$$\Omega(x_1, x_2, \ldots) \;=\; \prod_i \frac{1}{1-x_i} \;=\; exp\Big(\sum_{k\geq 1} \frac{p_k(x_1,x_2,\ldots)}{k}\Big) \;.$$

But we can go further. For instance, given two alphabets $A = a_1 + a_2 + \cdots$ and $B = b_1 + b_2 + \cdots$, substitute their difference in Ω and obtain that

$$\begin{aligned}
\Omega[A-B] \;&=\; exp\Big(\sum_{k\geq 1} \frac{p_k[A-B]}{k}\Big) \\
&=\; exp\Big(\sum_{k\geq 1} \frac{p_k[A]-p_k[B]}{k}\Big) \\
&=\; exp\Big(\sum_{k\geq 1} \frac{p_k[A]}{k}\Big) \Big/ exp\Big(\sum_{k\geq 1} \frac{p_k[B]}{k}\Big) \\
&=\; \Big(\prod_{a\in A} \frac{1}{1-a}\Big) \prod_{b\in B} (1-b)
\end{aligned} \qquad 2.6$$

The familiar expansions

$$\prod_i \frac{1}{1-x_i} \;=\; \sum_{m\geq 0} h_m(x) \quad , \qquad \prod_i (1-x_i) \;=\; \sum_{m\geq 0} e_m(x)$$

then yield that

$$h_m[A-B] \;=\; \sum_{i=0}^{m} h_i[A]\, (-1)^{m-i} e_{m-i}[B] \qquad 2.7$$

In particular, letting A be the empty alphabet we derive that

$$h_m[-B] \;=\; (-1)^m e_m[B]$$

This relation can be written in a more suggestive manner. Note that the ω involution defined by setting $\omega h_m = e_m$ may now simbly be written as

$$\omega h_m[X] \;=\; (-1)^m h_m[-X]$$

Better yet if we agree that evaluating a symmetric function $f(x_1, x_2, \ldots)$ at $-x_1, -x_2, \ldots$ in the ordinary sense should be expressed by writing

$$f(x_1, x_2, \ldots)\,\big|_{x_i \to -x_i} \;=\; f[\,^-X]$$

Then the ω transformation applied to any symmetric function, (whether homogeneous or not) may be written as

$$\omega\, f[X] \;=\; f[-^{-}X]\ .$$

Thus we may write the ω relation between skew Schur functions in the form

$$S_{\lambda'/\mu'}[A] \;=\; \omega S_{\lambda/\mu}[A] \;=\; (-1)^{|\lambda/\mu|} S_{\lambda/\mu}[-A] \;=\; S_{\lambda/\mu}[-^{-}A]\ . \tag{2.8}$$

In this vein note that the familiar addition formula for Schur functions now yields that

$$\begin{aligned} S_\lambda[A-B] \;&=\; \sum_{\mu\subseteq\lambda} S_\mu[A]\, S_{\lambda/\mu}[-B] \\ &=\; \sum_{\mu\subseteq\lambda} S_\mu[A]\, (-1)^{|\lambda/\mu|} S_{\lambda'/\mu'}[B] \end{aligned} \tag{2.9}$$

Here and after it will be good to use Schur function notation and write $h_m = S_m$, $e_m = S_{1^m}$. This given, 2.7 may be rewritten as an instance of 2.9, that is

$$S_m[A-B] \;=\; \sum_{i=0}^{m} S_i[A]\, S_{1^{m-i}}[-B]\ . \tag{2.10}$$

We are now ready to use divided difference methods in symmetric function theory.

3. Divided differences and symmetric polynomials

The point of departure here is the following basic identity which allows the calculation of divided differences by a recursive procedure. To state it, it will be convenient to set

$$\Delta_A^{(1)}(f) \;=\; \Delta_{a_1}(f) \;=\; f(a_1)\ . \tag{3.1}$$

Proposition 3.1

For any alphabet $A = a_1 + a_2 + a_3 + \cdots$ and any function $f(x)$ we have for $n \geq 0$

$$\Delta_{a_1,a_2,\ldots,a_{n+1}}(f) \;=\; \frac{\Delta_{a_1,a_2,\ldots,a_{n-1},a_n}(f) - \Delta_{a_1,a_2,\ldots,a_{n-1},a_{n+1}}(f)}{a_n - a_{n+1}}\ . \tag{3.2}$$

Proof

Note that 1.1 for $n = 2$ reduces to

$$\Delta_{a_1,a_2}(f) \;=\; \frac{f(a_1)}{a_1-a_2} \;+\; \frac{f(a_2)}{a_2-a_1} \;=\; \frac{f(a_1) \;-\; f(a_2)}{a_1-a_2}$$

which, in view of 3.1, is simply 3.2 for $n = 2$. More generally, from the definition 1.1 we get that

$$\Delta_{a_1,a_2,\ldots,a_{n-1},a_n}(f) \;=\; \sum_{s=1}^{n-1} \frac{f(a_s)}{\prod_{r=1;r\neq s}^{n-1}(a_s-a_r)} \times \frac{1}{a_s-a_n} \;+\; \frac{f(a_n)}{\prod_{r=1}^{n-1}(a_n-a_r)}\ . \tag{3.3}$$

Similarly we have

$$\Delta_{a_1,a_2,\ldots,a_{n-1},a_{n+1}}(f) \;=\; \sum_{s=1}^{n-1} \frac{f(a_s)}{\prod_{r=1;r\neq s}^{n-1}(a_s-a_r)} \times \frac{1}{a_s-a_{n+1}} \;+\; \frac{f(a_{n+1})}{\prod_{r=1}^{n-1}(a_{n+1}-a_r)} \;. \qquad 3.4$$

Since

$$\frac{\frac{1}{a_s-a_n} - \frac{1}{a_s-a_{n+1}}}{(a_n-a_{n+1})} \;=\; \frac{1}{(a_s-a_n)(a_s-a_{n+1})}$$

we see that, taking the difference of 3.3 and 3.4, we may write the right hand side of 3.2 in the form

$$\begin{aligned} &\sum_{s=1}^{n-1} \frac{f(a_s)}{\prod_{r=1;r\neq s}^{n-1}(a_s-a_r)} \times \frac{1}{(a_s-a_n)(a_s-a_{n+1})} \;+ \\ &\qquad + \frac{f(a_n)}{\prod_{r=1}^{n-1}(a_{n+1}-a_r)} \times \frac{1}{(a_n-a_{n+1})} \;. \qquad 3.4 \\ &\qquad\qquad - \frac{f(a_{n+1})}{\prod_{r=1}^{n-1}(a_{n+1}-a_r)} \times \frac{1}{(a_n-a_{n+1})} \;. \end{aligned}$$

But this is just another way of writing 1.1 for $n \to n+1$. Thus 3.2 must hold true in full generality.

An immediate consequence of 3.2 is the following remarkable recurrence

Theorem 3.1

For given alphabets $A = a_1 + a_2 + \cdots$, $B = b_1 + b_2 + \cdots$, $C = c_1 + c_2 + \cdots$, *setting*

$$A_k \;=\; a_1 + a_2 + \cdots + a_k \qquad\qquad (\textit{with } A_o = 0) \qquad 3.5$$

we have for any $k, j \geq 1$

$$\frac{S_j[a_k + A_{k-1} + B - C] \;-\; S_j[a_{k+1} + A_{k-1} + B - C]}{a_k - a_{k+1}} \;=\; S_{j-1}[A_{k+1} + B - C] \qquad 3.6$$

In particular we derive that

$$\Delta_{a_1,a_2,\ldots,a_{k+1}}(S_m[x+B-C]) \;=\; S_{m-k}[A_{k+1} + B - C] \qquad 3.7$$

Proof

Note that the addition formula for the complete homogeneous function gives

$$S_m[x+B-C] \;=\; \sum_{i=0}^{m} S_i[x]\, S_{m-i}\,[B-C] \;=\; \sum_{i=0}^{m} x^i\, S_{m-i}\,[B-C]$$

Thus we may write

$$\begin{aligned} \frac{S_j[a_k + A_{k-1} + B - C] - S_j[a_{k+1} + A_{k-1} + B - C]}{a_k - a_{k+1}} &= \sum_{i=0}^{j} S_{j-i}\,[A_{k-1} + B - C]\, \frac{a_k^i - a_{k+1}^i}{a_k - a_{k+1}} \\ &= \sum_{i=0}^{j} S_{j-i}\,[A_{k-1} + B - C]\, S_{i-1}[a_k + a_{k+1}] \\ &= S_{j-1}\,[A_{k-1} + a_k + a_{k+1} + B - C] \end{aligned}$$

and this is 3.6. Now by the convention in 3.1 we have

$$\Delta_{a_1}\big(S_m[x+B-C]\big) \;=\; S_m[a_1+B-C] \;=\; S_m[A_1+B-C]$$

Thus an application of 3.6 for $k=1$ and $j=m$ gives

$$\Delta_{a_1,a_2}\big(S_m[x+B-C]\big) \;=\; S_{m-1}[A_2+B-C]$$

and 3.7 follows by successive uses of 3.6.

We are now ready to establish the first non-trivial symmetric function consequence of the Newton interpolation formula. To appreciate the result we should start by noting that for any alphabets A, and B we have

$$S_m[x-B] \;=\; \sum_{k=0}^{m} S_{m-k}[A-B]\; S_k[x-A]$$

This is easily established by writing

$$S_m[x-B] \;=\; S_m[A-B \;+\; x-A]$$

and using the addition formula.

Now it develops that we also have

Proposition 3.2

$$S_m[x-B] \;=\; \sum_{k=0}^{m} S_{m-k}[A_{k+1}-B]\; S_k[x-A_k] \qquad 3.8$$

Proof

The expansion in 1.5 for $f(x)=S[x-B]$ gives, taking account of 3.7 for $B=0$ and $C\to B$:

$$S_m[x-B] \;=\; \sum_{k=0}^{m} S_{m-k}[A_{k+1}-B]\; (x-a_1)(x-a_2)\cdots(x-a_k)\ . \qquad 3.9$$

However it is deduced from 2.10 (or 2.7) that for any alphabet $A=a_1+a_2+\cdots+a_k$ we have

$$S_k[x-A_k] \;=\; (x-a_1)(x-a_2)\cdots(x-a_k)\ .$$

Thus 3.9 is simply another way of writing 3.8. This completes our proof.

4. Left coset representatives and divided differences

If G is a finite group and H is one of its subgroups then we have an equivalence relation "$\cong_H$" on G defined by setting for $\alpha, \beta \in G$

$$\alpha \cong_H \beta \qquad \Longleftrightarrow \qquad \alpha^{-1}\beta \in H \ \ .$$

This divides up G into a disjoint union of equivalence classes which are usuallly referred to as *"left cosets"* of H in G. We shall here and after denote by G/H any suitably chosen complete set of distinct representatives of the left cosets of H in G. Note that if

$$G/H \ = \ \{\tau_1, \tau_2, \tau_3, \ldots, \tau_k\} \qquad\qquad (\tau_1 = \text{ the identity of } G)$$

then we have the decomposition

$$G \ = \ H \ + \ \tau_2 H \ + \ \tau_3 H \ + \ \cdots \ + \ \tau_k H \ , \tag{4.1}$$

where by this we mean that every element $\sigma \in G$ can be uniquely written in the form $\sigma = \tau_i \, h$ for some $h \in H$ and $1 \leq i \leq k$.

The typical example we shall extensively use here is the case

$$G = S_n \qquad\qquad \text{and} \qquad\qquad H \ = \ S_{[1,2,\ldots,k\,]} \times S_{[k+1,k+2,\ldots,n\,]}$$

That is H is the subgroup consisting of the elements of the symmetric group S_n which leave invariant the two subsets

$$U_k = \{1, 2, \ldots, k\} \qquad\qquad \text{and} \qquad\qquad V_k = \{k+1, k+2, \ldots, n\}$$

It will be convenient to use a number of abreviations here. To begin with, for any integer m we set

$$[m] \ = \ \{1, 2, \ldots, m\}$$

For a given subset $E \subseteq [n]$ we let ${}^cE$ denote the complement of E in $[n]$. Since n will be kept fixed throughout this should cause no confusion. Note that if

$$E = \{1 \leq i_1 < i_2 < \cdots < i_k \leq n\} \qquad \text{and} \qquad {}^cE = \{1 \leq j_1 < j_2 < \cdots < j_{n-k} \leq n\}$$

then the symbol

$$\begin{bmatrix} [\,k] & {}^c[k] \\ E & {}^cE \end{bmatrix}$$

will represent the permutation

$$\tau(E) \ = \ \begin{bmatrix} 1 & 2 & \cdots & k & k+1 & k+2 & \cdots & n \\ i_1 & i_2 & \cdots & i_k & j_1 & j_2 & \cdots & j_{n-k} \end{bmatrix} \ . \tag{4.2}$$

Note further that every element $\sigma \in S_n$ may be uniquely written in the form

$$\sigma \;=\; \begin{bmatrix} 1 & 2 & \cdots & k & k+1 & k+2 & \cdots & n \\ i_{\alpha_1} & i_{\alpha_2} & \cdots & i_{\alpha_k} & j_{\beta_1} & j_{\beta_2} & \cdots & j_{\beta_{n-k}} \end{bmatrix}$$

with $E = \{1 \le i_1 < i_2 < \cdots < i_k \le n\}$ and ${}^cE = \{1 \le j_1 < j_2 < \cdots < j_{n-k} \le n\}$ complementary subsets of $\{1, 2, \ldots, n\}$ and with $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ a permutation of $(1, 2, \ldots, k)$ and $(\beta_1, \beta_2, \ldots, \beta_{n-k})$ a permutation of $(k+1, k+2, \ldots, n)$. In other words σ has a unique factorization of the form

$$\sigma \;=\; \begin{bmatrix} 1 & 2 & \cdots & k & k+1 & k+2 & \cdots & n \\ i_1 & i_2 & \cdots & i_k & j_1 & j_2 & \cdots & j_{n-k} \end{bmatrix} \times \begin{bmatrix} 1 & 2 & \cdots & k & k+1 & k+2 & \cdots & n \\ \alpha_1 & \alpha_2 & \cdots & \alpha_k & \beta_1 & \beta_2 & \cdots & \beta_{n-k} \end{bmatrix}$$

this shows that the collection

$$\left\{ \tau(E) = \begin{bmatrix} [\,k] & {}^c[k] \\ E & {}^cE \end{bmatrix} \;:\; E = \{1 \le i_1 < i_2 < \cdots < i_k \le n\} \;\right\}$$

gives a complete set of representatives of the left cosets of $S_{[k]} \times S_{\,{}^c[k]}$ in S_n. Thus using Young's notation we may succintly express the identity in 4.1 by writing

$$[1, 2, \ldots, n] \;=\; \Big(\sum_{\substack{E \subseteq [n] \\ |E| = k}} \tau(E) \Big) \times \;\; [1, 2, \ldots, k] \times [k+1, k+2, \ldots, n] \;\; . \tag{4.3}$$

Likewise we have

$$[1, 2, \ldots, n]' \;=\; \Big(\sum_{\substack{E \subseteq [n] \\ |E| = k}} sign\big(\tau(E)\big)\, \tau(E) \Big) \times \;\; [1, 2, \ldots, k]' \times [k+1, k+2, \ldots, n]' \;\; . \tag{4.4}$$

This given we are finally in a position to deal with the basic goals of this section.

Lascoux defines the "Sylvester" operator $\delta_{k|n-k}$ by setting for every polynomial P in $x_1, x_2, \ldots, x_n$

$$\delta_{k|n-k}\, P \;=\; \sum_{\substack{E \subseteq [n] \\ |E| = k|}} \Big(\prod_{\substack{i \in E \\ j \in {}^cE}} \frac{1}{x_i - x_j} \Big)\, \tau(E)\, P \tag{4.5}$$

Note that this may also be rewritten as

$$\delta_{k|n-k}\, P \;=\; \sum_{\substack{E \subseteq [n] \\ |E| = k|}} \tau(E) \Big(\prod_{\substack{1 \le i \le k \\ k+1 \le j \le n}} \frac{1}{x_i - x_j} \Big)\, P \;\; . \tag{4.6}$$

Now it is easy to see that we have

$$\Big(\prod_{\substack{1 \le i \le k \\ k+1 \le j \le n}} \frac{1}{x_i - x_j} \Big) \;=\; \frac{\Delta[X_k]\Delta[{}^cX_k]}{\Delta[X_n]}$$

Where, $\Delta[X_k]$, $\Delta[{}^cX_k]$ and $\Delta[X_n]$ respectively denote the Vandermonde determinants in the alphabets $X_k = x_1 + x_2 + \cdots + x_k$, ${}^cX_k = x_{k+1} + x_{k+2} + \cdots + x_n$ and $X_n = x_1 + x_2 + \cdots + x_n$. Substituting this in 4.6 and pulling out the Vandermonde determinant $\Delta[X_n]$, shows that in ultimate analysis the Sylvester operator is none other than

$$\delta_{k|n-k} \;=\; \frac{1}{\Delta[X_n]} \sum_{\substack{E\subseteq[n]\\ |E|=k}} sign\big(\tau(E)\big)\; \tau(E)\; \Delta[X_k]\Delta[{}^cX_k] \;\; . \tag{4.7}$$

Let us recall now that for any composition $p = (p_1, p_2, \ldots, p_n)$ we set

$$S_p[X_n] \;=\; \frac{det\, \big\|x_i^{p_j+n-j}\big\|_{i,j=1}^{n}}{\Delta[X_n]} \tag{4.8}$$

Thus for $p = \mu$ a partition $(\mu_1 \geq \mu_2 \geq \ldots \geq \mu_n \geq 0)$ this ratio simply gives the customary Schur function $S_\mu[X_n]$, and if the vector

$$(p_1 + n - 1, p_2 + n - 2, \ldots, p_n)$$

has distict non-negative components, we get again a Schur function but multiplied by the sign of the permutation that does this rearrangement, otherwise 4.8 simply reduces to 0.

Keeping this in mind the identity in 4.8 immediately gives the following basic result

Proposition 4.1

For any partitions $\mu = (\mu_1, \mu_2, \ldots, \mu_k)$, $\nu = (\nu_1, \nu_2, \ldots, \nu_{n-k})$ *we have*

$$\delta_{k|n-k}\;\; S_\mu[X_k]\, S_\nu[{}^cX_k] \;=\; S_{\mu_1^*,\mu_2^*,\ldots,\mu_k^*,\nu_1,\nu_2,\ldots,\nu_{n-k}}[X_n] \tag{4.9}$$

Where for convenience we have set

$$\mu_i^* \;=\; \mu_i - (n-k) \qquad\qquad (\text{for } i = 1, 2, \ldots, k) \tag{4.10}$$

Proof

Since

$$\Delta[X_k] S_\mu[X_k] \,=\, det\, \big\|x_i^{\mu_j+k-j}\big\|_{i,j=1}^{k} \quad \text{and} \quad \Delta[{}^cX_k] S_\nu[{}^cX_k] \,=\, det\, \big\|x_{k+i}^{\nu_j+n-k-j}\big\|_{i,j=1}^{n-k}$$

from 4.7 we get that

$$\delta_{k|n-k}\; S_\mu[X_k]\, S_\nu[{}^cX_k] \;=\; \frac{1}{\Delta[X_n]} \sum_{\substack{E\subseteq[n]\\ |E|=k}} sign\big(\tau(E)\big)\; \tau(E)\; det\, \big\|x_i^{\mu_j+k-j}\big\|_{i,j=1}^{k} det\, \big\|x_{k+i}^{\nu_j+n-k-j}\big\|_{i,j=1}^{n-k} \;\; . \tag{4.11}$$

But since we can write

$$det\, \big\|x_i^{\mu_j+k-j}\big\|_{i,j=1}^{k} \,=\, [1, 2, .., k]'\; x_1^{\mu_1+k-1} x_2^{\mu_2+k-2} \cdots x_k^{\mu_k+k-k}$$

and

$$det\, \big\|x_{k+i}^{\nu_j+n-k-j}\big\|_{i,j=1}^{n-k} \,=\, [k+1, k+2, .., n]'\; x_{k+1}^{\nu_1+n-k-1} x_{k+2}^{\nu_2+n-k-2} \cdots x_n^{\nu_{n-k}+n-n}$$

we can use 4.4 and rewrite 4.11 in the form

$$\delta_{k|n-k}\ S_\mu[X_k]\,S_\nu[{}^cX_k] \;=\; \frac{1}{\Delta[X_n]}\ \ [1,2,..,n]'\ \ x_1^{\mu_1+k-1}\cdots x_k^{\mu_k+k-k}\ x_{k+1}^{\nu_1+n-k-1}\cdots x_n^{\nu_{n-k}+n-n}\ \ .$$

Better yet we can also write

$$\delta_{k|n-k}\ S_\mu[X_k]\,S_\nu[{}^cX_k] \;=\; \frac{1}{\Delta[X_n]}\ \ [1,2,..,n]'\ \ x_1^{\mu_1^*+n-1}\cdots x_k^{\mu_k^*+n-k}\ x_{k+1}^{\nu_1+n-k-1}\cdots x_n^{\nu_{n-k}+n-n}\ \ .$$

In other words we have

$$\delta_{k|n-k}\ S_\mu[X_k]\,S_\nu[{}^cX_k] \;=\; \frac{det\ \left\| x_i^{\mu_1^*+n-1}\,,\ \cdots\,,\ x_i^{\mu_k^*+n-k}\,,\ \ x_i^{\nu_1+n-k-1}\,,\ \cdots\,,\ x_i^{\nu_{n-k}+n-n}\ \right\|_{i=1}^n}{\Delta[X_n]}$$

Which is what we wanted to prove.

This result has a number of important corollaries. To begin with it yields the following basic fact:

Theorem 4.1

The operator $\delta_{k|n-k}$ sends a polynomial P that is separately symmetric in the alphabets X_k and ${}^cX_k$ in to a polynomial Q that is symmetric in the alphabet X_n. Moreover, if P is homogeneous of degree d then Q will be homogeneous of degree $d-k(n-k)$.

Proof

Every polynomial P that is symmetric in X_k and ${}^cX_k$ may be expressed as a linear combination of products of the form $S_\mu[X_k]\,S_\nu[{}^cX_k]$, moreover if P is homogeneous of degree d the terms in its expansion will involve pairs of partition μ,ν with $|\mu|+|\nu|=d$. Thus all these assertions follow immediately from 4.9. In particular from 4.9 we get that the degree of the result will be

$$d-\binom{n}{2}\ +\ \binom{k}{2}\ +\ \binom{n-k}{2}\ =\ d\ -\ k(n-k)$$

precisely as asserted.

Following Lascoux let us denote by $\mathcal{SYM}(k|n-k)$ the ring of polynomials in $x_1,x_2,\ldots,x_n$ which are separately symmetric in X_k and ${}^cX_k$. Likewise $\mathcal{SYM}(n)$ will denote the subring of $\mathcal{SYM}(k|n-k)$ consisting of polynomials which are symmetric in $x_1,x_2,\ldots,x_n$. We have the following important consequence of 4.7.

Proposition 4.2

Symmetric polynomials behave like constants with respect to the Sylvester operator. More precisely, for any polynomial $P[X_n]$ and any $S[X_n]\in\mathcal{SYM}(n)$ we have

$$\delta_{k|n-k}\big(S[X_n]\,P[X_n]\big)\ =\ S[X_n]\ \delta_{k|n-k}\big(P[X_n]\big) \qquad 4.12$$

In particular

$$\delta_{k|n-k}\big(S[X_n]\big)\ =\ 0\ \ . \qquad 4.13$$

Proof

We need only verify 4.13 or equivalently that

$$\delta_{k|n-k}(1) \;=\; 0$$

But this is follows from Theorem 4.1 which in particular asserts that $\delta_{k|n-k}$ decreases the degree. Alternatively we can also derive this identity directly from formula 4.9 by choosing μ and ν to be the zero partitions.

It develops that Proposition 4.1 yields a number of interesting results concerning the quotient ring $\mathcal{SYM}(k|n-k)/\mathcal{SYM}(n)$. The crucial identity underlying these developments may be stated as follows.

Proposition 4.3

If $\mu = (\mu_1 \geq \mu_2 \geq \cdots \geq \mu_k \geq 0)$ and $\nu = (\nu_1 \geq \nu_2 \geq \cdots \geq \nu_{n-k} \geq 0)$ are any two partitions satisfying

$$\mu_1 \leq n-k \qquad \text{and} \qquad \nu_1 \leq k \tag{4.14}$$

Then

$$\delta(k|n-k)\; S_\mu[X_k] S_\nu[\,{}^cX_k] \;=\; 0 \tag{4.15}$$

except when diagrams of μ and ν are complementary in the $k \times n$ rectangle. More precisely we have 4.15 except when ν' (the conjugate of ν) is the partition

$$\nu' \;=\; (n-k-\mu_k\,,\, n-k-\mu_{k-1}\,,\,\ldots\,,\, n-k-\mu_1) \tag{4.16}$$

and in that case we have

$$\delta(k|n-k)\; S_\mu[X_k]\, S_\nu[\,{}^cX_k] \;=\; (-1)^{|\nu|} \tag{4.17}$$

Proof

Note first that 4.14 implies that for any $1 \leq i \leq k$ and any $1 \leq j \leq n-k$ we have

$$\mu_i + k - i \leq n-k+k-i \leq n-1$$

$$\nu_j + n-k-j \leq k+n-k-i \leq n-1$$

In other words the numbers

$$\begin{aligned} &e_1 = \mu_1 + k - 1\,, e_2 = \mu_2 + k - 1\,,\,\ldots\,,\, e_k = \mu_k + k - k\,, \\ &\qquad e_{k+1} = \nu_1 + n-k-1\,,\, e_{k+2} = \nu_2 + n-k-1\,,\,\ldots\,,\, e_n = \nu_{n-k} + n-k-(n-k)\,, \end{aligned} \tag{4.18}$$

are all in the interval $[0, n-1]$. Since 4.9, may equivalently be written as

$$\delta(k|n-k)\; S_\mu[X_k]\, S_\nu[\,{}^cX_k] \;=\; \frac{det\|x_i^{e_j}\|_{i,j=1}^n}{det\|x_i^{n-j}\|_{i,j=1}^n} \tag{4.19}$$

we see that will necessarily have 4.16 unless the exponents $e_1, e_2, \ldots, e_n$ are distinct and therefore simply a rearrangement of $0, 1, 2, \ldots, n-1$ and in this case the ratio in 4.19 will evaluate to the sign of the permutation that rearranges $e_1, e_2, \ldots, e_n$ in decreasing order. To complete the proof we must show that this sign is in fact given by $(-1)^{|\nu|}$ and that condition that $e_1, e_2, \ldots, e_n$ can be rearranged to $0, 1, 2, \ldots, n-1$ is equivalent to requiring that μ and ν be related by the equations in 4.16.

Now this is best understood by working out a significant example. To this end let us take $n = 9$ and $k = 5$. Consider the staircase diagram depicted below

$$\begin{array}{llllllll} X \\ X & X \\ X & X & X \\ X & X & X & X \\ X & X & X & X & X \\ X & X & X & X & X & X \\ X & X & X & X & X & X & X \\ X & X & X & X & X & X & X & X \end{array} \tag{4.19}$$

It is easily seen that the removal of a column or a row from such a diagram always results in a staircase diagram. This given, a way of obtaining a collection of numbers $e_1 > e_2 > e_3 > e_4 > e_5$ and $e_6 > e_7 > e_8 > e_9$ (as given in 4.18) which rearrange to $0, 1, 2, 4, 5, 6, 7, 8$ is to peel off the diagram in 4.19 successively either a row or a column, with 5 total row removals and 4 total column removals. We have indicated a particular choice of doing this by using the symbol "$\rightarrow$" to denote row removals and symbol "$\uparrow$" for column removals:

$$\begin{array}{llllllll} \uparrow \\ \uparrow & \uparrow \\ \uparrow & \uparrow & \uparrow \\ \uparrow & \uparrow & \uparrow & \rightarrow \\ \uparrow & \uparrow & \uparrow & \rightarrow & \rightarrow \\ \uparrow & \uparrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\ \uparrow & \uparrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\ \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow \end{array} \tag{4.19}$$

This particular construction is to represent the selection

$$e_1 = 8\ ,\ e_2 = 5\ ,\ e_3 = 4\ ,\ e_4 = 2\ ,\ e_5 = 1\ ,$$
$$e_6 = 7\ ,\ e_7 = 6\ ,\ e_8 = 3\ ,\ e_9 = 0\ ,$$

Now to obtain μ we need only subtract $4, 3, 2, 1, 0$ from e_1, e_2, e_3, e_4, e_5 and to obtain ν we subtract $3, 2, 1, 0$ from e_6, e_7, e_8, e_9. This process can be visualized as drawing the lines which cut out the 5×4 rectangle out of our staircase, as depicted below.

$$\begin{array}{llll|llll} \uparrow \\ \uparrow & \uparrow \\ \uparrow & \uparrow & \uparrow \\ \hline \uparrow & \uparrow & \uparrow & \rightarrow \\ \uparrow & \uparrow & \uparrow & \rightarrow & \rightarrow \\ \uparrow & \uparrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\ \uparrow & \uparrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\ \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow & \rightarrow \end{array} \tag{4.20}$$

removing the extra arrows and replacing each "$\rightarrow$" by "μ" and each "$\uparrow$" by "ν'" yields the diagram below

$$\begin{array}{cccc|} - & - & - & - \\ \nu' & \nu' & \nu' & \mu \\ \nu' & \nu' & \nu' & \mu \\ \nu' & \nu' & \mu & \mu \\ \nu' & \nu' & \mu & \mu \\ \mu & \mu & \mu & \mu \end{array} \qquad 4.21$$

which should inequivocally explain why the condition that $e_1, e_2, \ldots, e_n$ rearrange to $0, 1, 2, \ldots, n-1$ is equivalent to the condition that the diagrams of μ and ν be complementary in the $k \times n-k$ rectangle, yielding the equations in 4.16.

Finally, a look at the diagram in 4.20 should reveal that the number of "$\uparrow$" in the top row of the square is equal to the number of transposition to move e_6 right after e_1, similarly the number of "$\uparrow$" in the second (from the top) row of the rectangle is equal to the number of transposition to move e_7 right after e_6. Proceeding in this manner we obtain that the total number of "$\uparrow$" within the rectangle is gives the number of transpositions needed to rearrange $e_1, e_2, \ldots, e_n$ in decreasing order. In summary the sign of the rearranging permutation will necessarily be -1 to the power the number of ν' in the diagram 4.21 which is of course given by $|\nu|$. This yields 4.17 and completes the proof of the Proposition.

For convenience let $B_{k|n-k}$ denote the $k \times (n-k)$ box, and for a given $\mu \in B_{k|n-k}$ let μ^* denote the partition ν whose conjugate ν' is given by 4.16. With this notation Propositition 4.3 may be simply restated as

$$(-1)^{|\nu|}\delta(k|n-k)\, S_\mu[X_k]\, S_\nu[\,{}^cX_k] \;=\; \begin{cases} 0 & \text{if } \nu \neq \mu^* \\ 1 & \text{if } \nu = \mu^* \end{cases} \qquad 4.22$$

This given we have the following beautiful consequence of this result.

Theorem 4.2

(1) The polynomials $\{S_\mu[X_k]\}_{\mu \in B_{k|n-k}}$ *are a basis for the quotient ring*

$$\mathcal{SYM}(k, n-k)/\mathcal{SYM}(n)$$

(2) Every polynomial $P \in \mathcal{SYM}(k, n-k)$ *has a unique expansion of the form*

$$P \;=\; \sum_{\mu \in B_{k|n-k}} c_\mu[X_n]\, S_\mu[X_k] \qquad \textit{with } c_\mu[X_n] \in \mathcal{SYM}(n)$$

(3) The coefficients in this expansion may be computed by the formula

$$c_\mu[X_n] \;=\; (-1)^{k(n-k)-|\mu|}\, \delta(k|n-k) P S_{\mu^*}[X_k] \qquad 4.23$$

Proof

Note that any polynomial $P=P[x_1,x_2,\ldots,x_n]$ may be written in the form

$$P \;=\; \sum_{p,q} c_{p,q}\; X_k^p \;\; {}^cX_k^q \;\; ,$$

if $P\in\mathcal{SYM}(k,n-k)$ then we necessarily have

$$P \;=\; \frac{1}{k!(n-k)!}\sum_{p,q} c_{p,q}\Big([1,2,\ldots k]X_k^p\Big)\;\Big([k+1,k+2,\ldots n]\;{}^cX_k^q\Big) \;\; .$$

This shows that every $P\in\mathcal{SYM}(k,n-k)$ may be expanded in the form

$$P \;=\; \sum_{\mu,\nu} c_{\mu,\nu}\, S_\mu[X_k]\, S_\nu[{}^cX_k] \;\; . \qquad 4.24$$

Note further that the identity

$$\begin{aligned} S_\mu[X_k] \;&=\; S_\mu[X_n-{}^c X_k] \;=\; \sum_{\rho\subseteq\mu} S_{\mu/\rho}[X_n]\, S_\rho[-{}^cX_k] \\ &=\; \sum_{\rho\subseteq\mu} S_{\mu/\rho}[X_n]\,(-1)^{|\rho|} S_{\rho'}[{}^cX_k] \end{aligned}$$

implies the congruence

$$S_\mu[X_k]\cong(-1)^{|\mu|}S_{\mu'}[{}^cX_k] \qquad\qquad \mathrm{mod}\;\big(\mathcal{SYM}(n)\big) \;\; . \qquad 4.25$$

This shows that, modulo the ideal generated by $\mathcal{SYM}(n)$ the expansion in 4.24 reduces to

$$P \;\cong\; \sum_{\mu,\nu} c_{\mu,\nu}\, S_\mu[X_k]\,(-1)^{|\nu|} S_{\nu'}[X_k] \;\; .$$

Expanding the product $S_\mu[X_k]S_{\nu'}[X_k]$ in terms of Schur functions, we derive that in the quotient ring $\mathcal{SYM}(k,n-k)/\mathcal{SYM}(n)$ every element f has an expansion of the form

$$f \;\cong\; \sum_{\lambda\in B_{k|n-k}} c_\lambda\; S_\lambda[X_k] \;\; . \qquad 4.26$$

The reason there is no loss in restricting λ to $B_{k|n-k}$ is that $S_\lambda[X_k]\cong 0$ if λ has more than k non-zero parts and 4.26 gives that $S_\lambda[X_k]\;\cong\;(-1)^{\lambda|}[{}^cX_k]\;=\;0$ if μ' has more than $n-k$ non-zero parts or equivalently if $\mu_1>n-k$. This proves that the polynomials $\{S_\mu[X_k]\}_{\mu\in B_{k|n-k}}$ span the quotient $\mathcal{SYM}(k,n-k)/\mathcal{SYM}(n)$.

Next, suppose that we have the equality

$$\sum_{\substack{\lambda\in B_{k|n-k}\\ |\lambda|=m}} c_\lambda\; S_\lambda[X_k] \;=\; \sum_{\substack{\mu\in B_{k|n-k}\\ |\mu|<m}} Q_\mu[X_n]\; S_\mu[X_k] \qquad\qquad \big(\text{with } Q_\mu[X_n]\in\mathcal{SYM}(n)\big) \;\; . \qquad 4.27$$

Then acting on both sides by $\delta(k|n-k)S_{\lambda^*}[X_k]$ and applying 4.22 we obtain (using 4.12)

$$(-1)^{|\lambda^*|}c_\lambda = \sum_{\substack{\mu\in B_{k|n-k}\\ |\mu|<m}} Q_\mu[X_n]\ \delta(k|n-k)S_{\lambda^*}[X_k]S_\mu[X_k]$$

However, since none of the μ in this sum can be equal to λ, the relation in 4.22 forces $c_\lambda = 0$, proving the independence of the polynomials $\{S_\mu[X_k]\}_{\mu\in B_{k|n-k}}$ $\big(\mathrm{mod}\ (\mathcal{SYM}(n)\big)$. (†) This completes the proof of part *(1)*.

Using this fact we derive that every $f \in \mathcal{SYM}(k, n-k)$ that is homogeneous of degree m may be written in the form

$$f \;=\; \sum_{\substack{\lambda\in B_{k|n-k}\\ |\lambda|=m}} c_\lambda\, S_\lambda[X_k] \;+\; \sum_{i=1}^{n} e_i[X_n]\, f_i[X_k] \qquad \big(\text{with } f_i[X_n]\in\mathcal{SYM}(k,n-k)\big)\ . \qquad 4.27$$

where the c_λ are scalars and $e_1[X_n], e_2[X_n], \ldots, e_n[X_n]$ denote the elementary symmetric functions in $x_1, x_2, \ldots, x_n$. Iterating this result on the $f_i[X_n]$ themselves and repeating this process as many times as necessary we ultimately obtain that f may be given an expansion of the form

$$f \;=\; \sum_{\lambda\in B_{k|n-k}} c_\lambda[X_n]\, S_\lambda[X_k] \qquad \big(\text{with } c_\lambda[X_n]\in\mathcal{SYM}(n)\big)\ . \qquad 4.28$$

To complete the proof of *(2)* we need to show that the coefficients $c_\lambda[X_n]$ can be uniquely determined by f. However this is a direct consequence of *(3)* which is immediately obtained by applying to both sides of 4.28 the operator $\delta(k|n-k)S_{\lambda^*}[X_n]$ and using 4.22 again.

One last remark we need to add in this section is that the case $k = 1$ of Proposition 4.1 may given an interesting reformulation. Namely we have

Theorem 4.3

For any partition $\mu = (\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n)$ *we have*

$$S_{\mu_1,\mu_2,\ldots,\mu_n}[X_n] \;=\; \delta(1|n)\ x_1^{\mu_1+n-1}\ S_{\mu_2,\ldots,\mu_n}[X_n - x_1] \qquad 4.29$$

which may be viewed as a "Rodriguez formula" for Schur functions.

(†) The relation in 4.27 must hold true with some non-zero scalars c_λ if a non-trivial linear combination of these polynomials is in the ideal generated by the symmetric functions in $x_1, x_2, \ldots, x_n$.