Notes on Schubert Polynomials

I. G. Macdonald

(Document compiled on 19 February 2025.)

To cite this work, please use the original publication details, as provided in the BibTeX entry below.

```
@book{Macdonald1991,
     author    = {I. G. Macdonald},
     title     = {Notes on Schubert Polynomials},
     publisher = {Publications du LaCIM},
     address   = {Laboratoire de combinatoire et d'informatique math\'ematique
                  (LaCIM), Universit\'e du Qu\'ebec \`a Montr\'eal},
     year      = 1991,
     volume    = 6,
     ISBN      = {2-89276-086-0 LACIM Montr\'eal},
  }
```

Notice of Recompilation (February 2025).

This is a modern recompilation of the TeX source used to produce the original version published in 1991. Changes to the source were made to ensure compatibility with pdfTeX (2024) and to add hyperlinks to the document. Except for a footnote added in the appendix to Chapter IV, no changes were made to the original content of the document itself, and the original numbering for equations, sections and chapters were preserved. However, the page numbers do not match the original version.

We extend our sincere thanks to the *Estate of Ian Grant Macdonald* for providing permission to recompile and distribute this work online, and to the Trustees of The Combinatorics Consortium (TCC), *Sara Billey*, *Vic Reiner*, *Volkmar Welker*, and *Mike Zabrocki*, for their invaluable support in navigating the legal and technical complexities of the copyright transfer.

Copyright information.

Dépôt légal, premier semestre 1991, Bibliothèque nationale du Québec.
ISBN 2-89276-086-0 LACIM Montral

Foreword

These notes are the fruit of the author's attempts to understand and develop from scratch the elegant theory of Schubert polynomials created by A. Lascoux and M.P.-Schützenberger in recent years. Most of the results expounded here occur somewhere in the publications of these authors, though not always accompanied by proof, and I have not attempted to give chapter and verse at each point. Brief indications to the literature will be found in the notes and references at the end.

Topics *not* covered in these notes include (i) the interpretation of Schubert polynomials as traces of functors (from filtered vector spaces to vector spaces) for which we refer to [KP]; and (ii) the non-commutative theory, for which we refer to [LS8].

Most of this material was presented in a course of lectures at the University of California, San Diego in the winter quarter of 1990, and I would like to take this opportunity to thank the audience, especially Adriano Garsia and Jeff Remmel, for their support.

San Diego, 1991 I. G. Macdonald

Contents

Chapter I

Permutations

For each integer $n \geq 1$, let S_n denote the symmetric group of degree n, that is to say the group of all permutations of the set $[1, n] = \{1, 2, \ldots, n\}$. Each $w \in S_n$ is a mapping of $[1, n]$ onto itself. As is customary, we write all mappings on the left of their arguments, so that the image of $i \in [1, n]$ under w is $w(i)$. We shall sometimes denote w by the sequence $(w(1), w(2), \ldots, w(n))$. Thus for example (53214) is the element of S_5 that sends 1 to 5, 2 to 3, 3 to 2, 4 to 1 and 5 to 4.

For $i = 1, 2, \ldots, n-1$ let s_i denote the transposition that interchanges i and $i+1$, and fixes all other elements of $[1, n]$. We have

$$\textbf{(1.1)} \qquad \begin{cases} s_i^2 = 1, & \\ s_i s_j = s_j s_i & \text{if } |i-j| > 1, \\ s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1} & (1 \leq i \leq n-2). \end{cases}$$

Also, for each $w \in S_n$, let

$$I(w) = \{\, (i, j) : 1 \leq i < j \leq n \text{ and } w(i) > w(j) \,\}.$$

We regard $I(w)$ as a subset of the square $\Sigma_n = [1, n] \times [1, n]$, and we shall adopt the convention of matrices, that in Σ_n the first coordinate increases from north to south, and the second coordinate from west to east. The group $S_n \times S_n$ acts on $\Sigma_n : (u \times v)(i, j) = (u(i), v(j))$. In particular, S_n acts diagonally: $w(i, j) = (w \times w)(i, j) = (w(i), w(j))$.

Let $w \in S_n$, $1 \leq r \leq n-1$. Then ws_r is the permutation

$$(w(1), \ldots, w(r+1), w(r), \ldots, w(n))$$

and it is clear that

$$\textbf{(1.2)} \qquad I(ws_r) = \begin{cases} s_r I(w) \cup \{(r, r+1)\} & \text{if } w(r) < w(r+1), \\ s_r I(w) - \{(r+1, r)\} & \text{if } w(r) > w(r+1). \end{cases}$$

Let $\ell(w)$= Card $I(w)$. Then from (1.2) we have

(1.3)
$$\ell(ws_r) = \begin{cases} \ell(w)+1 & \text{if } w(r) < w(r+1), \\ \ell(w)-1 & \text{if } w(r) > w(r+1). \end{cases}$$

(1.4) *$s_1, \ldots, s_{n-1}$ generate the group S_n.*

Proof: We shall show by induction on $\ell(w)$that each $w \in S_n$ is a product of s's. If $\ell(w) = 0$, then $w = 1$ and there is nothing to prove. If $\ell(w) > 0$ then $w(r) > w(r+1)$ for some r, and hence $\ell(ws_r) = \ell(w) - 1$ by (1.3). Hence $ws_r = s_{a_1} \ldots s_{a_p}$ say, and therefore (as $s_r^2 = 1$) $w = s_{a_1} \ldots s_{a_p} s_r$. ‖

For each $w \in S_n$, the *length* of w is the minimal length of a sequence $(a_1, \ldots, a_p)$ such that $w = s_{a_1} \ldots s_{a_p}$.

(1.5) *The length of $w \in S_n$ is equal to $\ell(w) =$ Card $I(w)$.*

Proof: Let $\ell'(w)$ temporarily denote the length of w. The proof of (1.4) shows that w can be written as a word of length $\ell(w)$ in the s_i, so that $\ell'(w) \le \ell(w)$. Conversely, let $w = s_{a_1} \ldots s_{a_p}$ be *any* expression of w as a product of s_i. To show that $\ell(w) \le \ell'(w)$ it is enough to show that $\ell(w) \le p$. Let $w' = s_{a_1} \ldots s_{a_{p-1}}$; from (1.3) we have $\ell(w) \le \ell(w') + 1$ and hence

$$\ell(w') \le p-1 \Rightarrow \ell(w) \le p.$$

Hence the proof is completed by induction on p. ‖

(1.6) *Let $w \in S_n$. Then*

(i) $\ell(w) = 0$ *if and only if* $w = 1$.

(ii) $\ell(w) = 1$ *if and only if* $w = s_r$ $(1 \le r \le n-1)$.

(iii) $\ell(w^{-1}) = \ell(w)$.

(iv) *Let* $w_0 = (n, n-1, \ldots, 2, 1) \in S_n$. *Then*

$\ell(w_0 w) = \ell(w w_0) = \frac{1}{2}n(n-1) - \ell(w)$.

Proof: (i), (ii) require no comment. Also (iii) is clear, since $w = s_{a_1} \ldots s_{a_p}$ if and only if $w^{-1} = s_{a_p} \ldots s_{a_1}$.

(iv) The set $I(w_0)$ consists of all $(i,j) \in \Sigma_n$ such that $i < j$, so that $\ell(w_0) = \frac{1}{2}n(n-1)$. Next, we have

$$ww_0 = (w(n), w(n-1), \ldots, w(1))$$

so that $I(ww_0)$ is the complement of $I(w)$ in $I(w_0)$, and therefore

$$\ell(ww_0) = \frac{1}{2}n(n-1) - \ell(w).$$

Finally, since $w_0^2 = 1$ we have, by virtue of (iii) above,

$$\begin{aligned}\ell(w_0 w) &= \ell(w^{-1}w_0)\\ &= \frac{1}{2}n(n-1) - \ell(w^{-1})\\ &= \frac{1}{2}n(n-1) - \ell(w). \ \|\end{aligned}$$

The element w_0 is called the *longest element* of S_n.

For each $w \in S_n$ let $R(w)$ denote the set of all sequences $(a_1, \ldots, a_p)$ of length $p = \ell(w)$ such that $w = s_{a_1} \ldots s_{a_p}$. Such sequences are called *reduced words* for w. Clearly,

$$(a_1, \ldots, a_p) \in R(w) \iff (a_p, \ldots, a_1) \in R(w^{-1}).$$

(1.7) *Let* $(a_1, \ldots, a_p) \in R(w)$. *Then*

$$I(w) = \{s_{a_p} \ldots s_{a_{r+1}}(a_r, a_r + 1) : \ 1 \le r \le p\}.$$

Proof: Let $w' = ws_{a_p} = s_{a_1} \ldots s_{a_{p-1}}$. Then $\ell(w') = p - 1$ and hence by (1.2) and (1.3) we have

$$I(w) = s_{a_p} I(w') \cup \{(a_p, a_p + 1)\}$$

from which (1.7) follows by induction on p. ‖

(1.8) (Exchange Lemma). *Let* $(a_1, \ldots, a_p), (b_1, \ldots, b_p) \in R(w)$. *Then*

$$(b_1, a_1, \ldots, \hat{a}_i, \ldots, a_p) \in R(w) \textit{ for some } i = 1, 2, \ldots, p.$$

Proof: By (1.7), applied to w^{-1}, we have $(b_1, b_1 + 1) \in I(w^{-1})$ and hence

$$(b_1, b_1 + 1) = s_{a_1} \ldots s_{a_{i-1}}(a_i, a_i + 1)$$

for some $i = 1, \ldots, p$. It follows that

$$s_{b_1} = s_{a_1} \ldots s_{a_{i-1}} s_{a_i} (s_{a_1} \ldots s_{a_{i-1}})^{-1},$$

so that $s_{b_1} s_{a_1} \ldots s_{a_{i-1}} = s_{a_1} \ldots s_{a_i}$ and therefore

$$s_{b_1} s_{a_1} \ldots \hat{s}_{a_i} \ldots s_{a_p} = s_{a_1} \ldots s_{a_p} = w. \ \|$$

(1.9) *Let* $w = s_{a_1} \cdots s_{a_r}$ *where* $r > \ell(w)$. *Then*

$$w = s_{a_1} \cdots \hat{s}_{a_p} \cdots \hat{s}_{a_q} \cdots s_{a_r}$$

for some pair (p,q) *such that* $1 \le p < q \le r$.

Proof: Since $\ell(s_{a_1}) = 1$ and $\ell(s_{a_1} \cdots s_{a_r}) < r$ there exists $q \ge 2$ such that

$$\ell(s_{a_1} \cdots s_{a_{q-1}}) = q-1, \ \ell(s_{a_1} \cdots s_{a_q}) < q.$$

Let $v = s_{a_1} \cdots s_{a_{q-1}}$, so that $\ell(v) = q-1$ and $\ell(vs_{a_q}) \le q-1$, whence by (1.3) we have $\ell(vs_{a_q}) = q-2$. Let $(b_1, \ldots, b_{q-2})$ be a reduced word for vs_{a_q}, then $(b_1, \ldots, b_{q-2}, a_q)$ and $(a_1, \ldots, a_{q-1})$ are reduced words for v. By (1.8) (applied to v^{-1}) it follows that $v = s_{a_1} \cdots \hat{s}_{a_p} \cdots s_{a_{q-1}}$ for some $p = 1, 2, \ldots, q-1$, and hence

$$w = vs_{a_q} \cdots s_{a_r} = s_{a_1} \cdots \hat{s}_{a_p} \cdots \hat{s}_{a_q} \cdots s_{a_r}. \ \|$$

If $i < j$, let t_{ij} denote the transposition that interchanges i and j and fixes each $k \ne i, j$. For each permutation w, let $e_{ij}(w)$ denote the number of k such that $i < k < j$ and $w(k)$ lies between $w(i)$ and $w(j)$. Consideration of $I(w)$ and $I(wt_{ij})$ shows that

(1.10)
$$\ell(wt_{ij}) = \begin{cases} \ell(w) - 2e_{ij}(w) - 1 & \text{if } w(i) > w(j), \\ \ell(w) + 2e_{ij}(w) + 1 & \text{if } w(i) < w(j). \end{cases}$$

In particular, $\ell(wt_{ij}) = \ell(w) \pm 1$ if and only if $e_{ij} = 0$.

(1.11) *Let* v, w *be permutations and let* $(a_1, \ldots, a_p)$ *be a reduced word for* w. *Then the following conditions are equivalent :*

(i) $\ell(v) < \ell(w)$ *and* $v^{-1}w$ *is a transposition,*

(ii) $v = s_{a_1} \cdots \hat{s}_{a_r} \cdots s_{a_p}$ *for some* $r = 1, 2, \ldots, p$.

Proof: (i) $\Rightarrow$ (ii). Suppose that $v^{-1}w = t_{ij}$, so that $v = wt_{ij}$. Then (1.10) shows that $w(i) > w(j)$, so that $(i,j) \in I(w)$. Hence by (1.7) we have $(i,j) = s_{a_p} \cdots s_{a_{r+1}}(a_r, a_{r+1})$ for some $r = 1, 2, \ldots, p$, and therefore

(1)
$$\begin{aligned} t_{ij} &= (s_{a_p} \cdots s_{a_{r+1}}) s_{a_r} (s_{a_p} \cdots s_{a_{r+1}})^{-1} \\ &= s_{a_p} \cdots s_{a_{r+1}} s_{a_r} s_{a_{r+1}} \cdots s_{a_p}. \end{aligned}$$

Consequently

$$\begin{aligned} v = wt_{ij} &= (s_{a_1} \cdots s_{a_p})(s_{a_p} \cdots s_{a_r} \cdots s_{a_p}) \\ &= s_{a_1} \cdots \hat{s}_{a_r} \cdots s_{a_p}. \end{aligned}$$

(ii) $\Rightarrow$ (i). Clearly $\ell(v) < \ell(w)$, and the calculation above shows that $v^{-1}w$ is the transposition (1). $\|$

The Bruhat order

Let v, w be permutations such that

(a) $\ell(w) = \ell(v) + 1$,

(b) $w = tv$ *where* t *is a transposition.*

Since $tv = vt'$ with $t' = v^{-1}tv$ also a transposition, we can replace (b) by

(b') $w = vt'$ *where* t' *is also a transposition.*

If (a) and (b) (or (b')) are satisfied we shall say that w *covers* v and write $v \to w$.

(1.12) *Let* $v, w \in S_n$ *and let* w_0 *be the longest element of* S_n*. Then the following conditions are equivalent:*

$(a) v \to w$; $(b) v^{-1} \to w^{-1}$; $(c) ww_0 \to vw_0$; $(d) w_0 w \to w_0 v$.

This follows from the definition and (1.6)(iii),(iv). ‖

(1.13) *Let* $(a_1, \dots, a_p)$ *be a reduced word for* w*. Then* $v \to w$ *if and only if* $v = s_{a_1} \cdots \hat{s}_{a_i} \cdots s_{a_p}$ *for some* $i = 1, 2, \dots, p$ *such that* $(a_1, \dots, \hat{a}_i, \dots, a_p)$ *is reduced.*

This follows from (1.11).

(1.14) *Let* w *be a permutation and let* $i \geq 1$*. Then either* $w \to s_i w$ *or* $s_i w \to w$*. Moreover we have* $s_i w \to w$ *if and only if there is a reduced word for* w *starting with* i*.*

Proof: The first statement follows from (1.3) and (1.6)(iii). If $s_i w \to w$, let $(a_1, \dots, a_p)$ be a reduced word for $s_i w$; then $w = s_i s_{a_1} \cdots s_{a_p}$ is a reduced expression for w. Conversely if $w = s_i s_{a_1} \cdots s_{a_p}$ is reduced, it is clear that $\ell(s_i w) = \ell(w) - 1$, and hence $s_i w \to w$. ‖

(1.15) *Let* v, w *be permutations and let* $i \geq 1$ *be such that*

$$v \to s_i v \neq w.$$

Then $v \to w$ *if and only if both* $w \to s_i w$ *and* $s_i v \to s_i w$*.*

Proof: Assume that $v \to w$, and let $(a_1, \dots, a_p)$ be a reduced word for w. Suppose that $a_1 = i$. By (1.13) we have $v = s_{a_1} \cdots \hat{s}_{a_r} \cdots s_{a_p}$ for some r. If $r = 1$ then $s_i v = s_{a_1} v = w$, and if $r > 1$ then $s_i v = s_{a_2} \cdots \hat{s}_{a_r} \cdots s_{a_p}$ has length $< p - 1 = \ell(v)$, so that $s_i v \to v$ by (1.14). Since both these possibilities are excluded by our hypothesis, we can conclude that $a_1 \neq i$. Hence (1.14) shows that $w \to s_i w$. It follows that $s_i s_{a_1} \cdots s_{a_p}$ is a reduced expression for $s_i w$, and $s_i s_{a_1} \cdots \hat{s}_{a_r} \cdots s_{a_p}$ is one for $s_i v$. Hence (1.13) shows that $s_i v \to s_i w$.

Conversely, assume that $w \to s_i w$ and $s_i v \to s_i w$. As before, let $w = s_{a_1} \cdots s_{a_p}$ be a reduced expression. Then $s_i w = s_i s_{a_1} \cdots s_{a_p}$ is reduced, and since $s_i v \neq w$ it follows from (1.13) that

$s_i v = s_i s_{a_1} \cdots \hat{s}_{a_r} \cdots s_{a_p}$ for some $r = 1, 2, \ldots, p$. Hence $v = s_{a_1} \cdots \hat{s}_{a_r} \cdots s_{a_p}$ and so $v \to w$ by (1.13) again. ‖

The *Bruhat order* , denoted by $\leq$, is the partial order on S_n that is the transitive closure of the relation $\to$. In other words, if v and w are permutations, $v \leq w$ means that there exists $r \geq 0$ and $v_0, v_1, \ldots, v_r$ in S_n such that

$$v = v_0 \to v_1 \to \cdots \to v_r = w$$

(which implies that $\ell(w) = \ell(v) + r$).

(1.16) *Let $v, w \in S_n$ and $i \geq 1$ be such that $s_i v \to v$ and $s_i w \to w$. Then the following conditions are equivalent :*

(i) $v \leq w$, (ii) $s_i v < w$, (iii) $s_i v \leq s_i w$.

Proof: (i) $\Rightarrow$ (ii). We have $s_i v < v \leq w$, hence $s_i v < w$.

(ii) $\Rightarrow$ (i). By definition there exist $v_0, v_1, \ldots, v_m$, where $m \geq 1$, such that

$$s_i v = v_0 \to v_1 \to \cdots \to v_m = w.$$

We have $v_0 \to s_i v_0$ and $s_i v_m \to v_m$. Hence there exists $k = 1, 2, \ldots, m$ such that $v_j \to s_i v_j$ for $0 \leq j \leq k-1$, and $s_i v_k \to v_k$.

Suppose $1 \leq j \leq k-1$. Then $v_{j-1} \to s_i v_{j-1}$ and $v_{j-1} \to v_j$; also $v_j \neq s_i v_{j-1}$, otherwise we should have $s_i v_j = v_j - 1$ and hence $s_i v_j \to v_j$. Hence by (1.15) we have

$$\textbf{(1)} \qquad s_i v_{j-1} \neq s_i v_j \qquad (1 \leq j \leq k-1).$$

Next, we have $v_{k-1} \to s_i v_{k-1}$ and $v_{k-1} \to v_k$. If $v_k \neq s_i v_{k-1}$ we should by (1.15) have $v_k \to s_i v_k$, contradicting the definition of k. Hence

$$\textbf{(2)} \qquad v_k = s_i v_{k-1}.$$

From (1) and (2) it follows that

$$v = s_i v_0 \to s_i v_1 \to \cdots \to s_i v_{k-1} = v_k \to \cdots \to v_m = w$$

and hence $v \leq w$.

This shows that (i) and (iii) are equivalent. To show that (ii) and (iii) are equivalent, assume that $v, w \in S_n$ for some $n \geq 1$, let w_0 be the longest element of S_n, and replace v, w respectively by $s_i w w_0$ and $s_i v w_0$. Then we have

$$\begin{aligned} s_i v \leq s_i w &\iff s_i w w_0 \leq s_i v w_0 \quad &&\text{(by (1.12))} \\ &\iff w w_0 < s_i v w_0 &&\text{(by } (a) \Leftrightarrow (b)) \\ &\iff s_i v < w &&\text{(by (1.12) again)} \end{aligned}$$

and the proof is complete. ‖

(1.17) *Let v, w be permutations and let $\boldsymbol{a} = (a_1, \ldots, a_p)$ be a reduced word for w. Then the following conditions are equivalent:*

(i) $v \leq w$;

(ii) *there exists a subsequence $\boldsymbol{b} = (b_1, \ldots, b_q)$ of $\boldsymbol{a}$ such that $v = s_{b_1} \cdots s_{b_q}$;*

(iii) *there exits a reduced subsequence $\boldsymbol{b} = (b_1, \ldots, b_q)$ of $\boldsymbol{a}$ such that $v = s_{b_1} \cdots s_{b_q}$.*

Proof: If follows from (1.13) that (i) $\Rightarrow$ (iii), and from (1.9) that (ii) and (iii) are equivalent. Thus it remains to prove that (iii) $\Rightarrow$ (i).

We proceed by induction on $r = p + q = \ell(v) + \ell(w)$. If $r = 0$, we have $v = w = 1$, so assume that $r \geq 1$. We distinguish two cases:

(a) $v \to s_{a_1} v$. In this case we have $b_1 \neq a_1$, hence $(b_1, \ldots, b_q)$ is a subsequence of $(a_2, \ldots, a_p)$, which is a reduced word for $s_{a_1} w$. By the inductive hypothesis we have $v \leq s_{a_1} w < w$, hence $v < w$.

$(b) s_{a_1} v \to v$. In this case $\ell(s_{a_1} v) + \ell(w) = p - 1 + q = r - 1$, and $s_{a_1} v = s_{a_1} s_{b_1} \cdots s_{b_q}$. If $a_1 = b_1$ we have $s_{a_1} v = s_{b_2} \cdots s_{b_q}$, and if $a_1 \neq b_1$ then $(a_1, b_1, \ldots, b_q)$ is a non-reduced subsequence of $(a_1, \ldots, a_p)$. Hence the inductive hypothesis implies that $s_{a_1} v < w$. But also $s_{a_1} w \to w$, hence $v \leq w$ by (1.16). ‖

(1.18) *Let $w \in S_n$ and let t be a transposition. Then*

$$\ell(wt) < \ell(w) \Rightarrow wt < w.$$

This follows from (1.11) and (1.17). ‖

To recognize when two permutations are comparable for the Bruhat order, the following rule may be used. For each $w \in S_n$ let $K(w)$ denote the column-strict tableau (of shape $\delta = (n-1, n-2, ..., 1)$) whose jth column, for $1 \leq j \leq n-1$, consists of the numbers $w(1), ..., w(n-j)$ arranged in increasing order from north to south.

(1.19) *Let $v, w \in S_n$. then $v \leq w$ if and only if $K(v) \leq K(w)$ (i.e., each entry in $K(v)$ is less than or equal to the corresponding entry in $K(w)$).*

Proof: If $v \to w$ it is easily seen that $K(v) \leq K(w)$, and hence $v \leq w$ implies $K(v) \leq K(w)$.

Conversely, suppose that $K(v) \le K(w)$ and let $j = j(v,w)$ be the smallest integer ≥ 1 such that $v(j) \ne w(j)$. (If $v = w$ we define $j(v,w) = n$.) We proceed by descending induction on $j(v,w)$. If $j(v,w) = n$ we have $v = w$, so assume $j(v,w) = j < n$. Then $w(j)$ is not equal to any $v(1), ..., v(j)$ and hence is equal to $v(k)$ for some $k > j$. For each $i < j$ the $(n-i)$th columns of $K(v)$ and $K(w)$ are identical, and since $K(v) \le K(w)$ it follows that $v(j) < w(j)$, i.e. $v(j) < v(k)$. Let $v' = vt_{jk}$, then by (1.10) we have $\ell(v) < \ell(v')$ and hence $v < v'$ by (1.18). Also $v'(i) = v(i) = w(i)$ for $i < j$, and $v'(j) = v(k) = w(j)$ so that $j(v', w) > j$. Hence $v' \le w$ by the inductive hypothesis, and therefore $v < w$. ‖

The diagram of a permutation

We may regard $I(w)$ as a "diagram" of $w \in S_n$. However, for many purposes it is more convenient to define the diagram of w to be

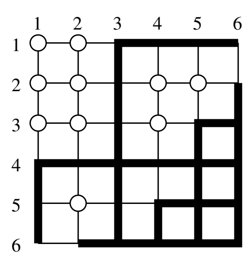
$$D(w) = (1 \times w)I(w).$$

Thus we have $(i,j) \in D(w)$ if and only if $(i, w^{-1}j) \in I(w)$; that is

(1.20) $$(i,j) \in D(w) \iff i < w^{-1}j \text{ and } j < wi.$$

Hence the points (i,j) in the square $\Sigma_n = [1,n]^2$ *not* in $D(w)$ are those for which either $i \ge w^{-1}$ or $j \ge wi$.

The *graph* $G(w)$ of w is the set of points $(i, w(i))$ $(1 \le i \le n)$, or equivalently $(w^{-1}j, j)$ $(1 \le j \le n)$. The complement of $D(w)$ in Σ_n therefore consists of all the lattice points due south or due east of some point of $G(w)$, hence is the union of the hooks with corners at the points of $G(w)$. For example, if $w = (365142)$ and $n = 6$, the diagram $D(w)$ consists of the points circled in the picture below:



If $m > n$, we shall identify S_n with the subgroup of permutations $w \in S_m$ that fix $n+1$, $n+2$, $\ldots, m$. We may then form the group

$$S_\infty = \bigcup_{n \ge 1} S_n$$

consisting of all permutations of the set of positive integers that fix all but a finite number of them.

The diagram $D(w)$ of $w \in S_n$ is unchanged by this identification of S_n with the subgroup of S_∞ fixing all $m > n$, and hence is well-defined for all $w \in S_\infty$. Also, it is clear from the definitions and (1.7) that

(1.21) (i) *$D(w^{-1})$ is the transpose of $D(w)$ (i.e., we have $(i,j) \in D(w^{-1})$ if and only if $(j,i) \in D(w)$).*

(ii) *Card $D(w) = \ell(w)$.*

(iii) *If $(a_1, \ldots, a_p) \in R(w)$, then $D(w)$ consists of the lattice points*

$$(s_{a_p} \ldots s_{a_{r+1}}(a_r), s_{a_1} \ldots s_{a_{r-1}}(a_r))$$

for $r = 1, 2, \ldots, p$. ‖

In particular, it follows from (iii) above that

(1.22) (i) *If $\ell(ws_r) > \ell(w))$, then $D(ws_r) = (s_r \times 1)D(w) \cup \{(r, wr)\}$.*

(ii) *If $\ell(s_r w) > \ell(w))$, then $D(ws_r) = (1 \times s_r)D(w) \cup \{(w^{-1}r, r)\}$.* ‖

The code of a permutation

Let $w \in S_n$, and for each $i \geq 1$ let

$$c_i(w) = \text{Card}\{j : j > i \text{ and } w(j) < w(i)\}.$$

Thus $c_i(w)$ is the number of points in the i^{th} row of $I(w)$, or equivalently the number of points in the i^{th} row of $D(w)$. The vector

$$c(w) = (c_1(w), \ldots, c_n(w)) \in \mathbf{N}^n$$

is called the *code* of w. As with partitions, we may disregard any string of zeros at the right-hand end of $c(w)$, and with this convention the code $c(w)$ (like the diagram $D(w)$) is unchanged by the embedding of S_n in S_m where $m > n$ and is well-defined for all $w \in S_\infty$.

The permutation w may be reconstructed from its code $c(w) = (c_1, c_2, \ldots)$ as follows :– for each $i \geq 1$, $w(i)$ is the $(c_i + 1)^{\text{th}}$ element, in increasing order, of the sequence of positive integers from which $w(1), w(2), \ldots, w(i-1)$ have been deleted. The sum $|c| = c_1 + c_2 + \cdots$ is equal to $\ell(w)$. Each sequence $c = (c_1, c_2, \ldots)$ of non-negative integers such that $|c| < \infty$ occurs as the code of a unique permutation $w \in S_\infty$.

The *length* of $c(w)$ is the largest r such that $c_r(w) \neq 0$. From the definition, r is the *last descent* of the permutation w, that is to say $w(r) > w(r+1)$ and $w(r+1) < w(r+2) < \ldots$.

(1.23) (i) *If* $\ell(ws_r) > \ell(w)$ (i.e., *if* $w(r) < w(r+1)$) *then*

$$c(ws_r) = s_r c(w) + \epsilon_r,$$

where ϵ_r *is the sequence with 1 in the* r^{th} *place and 0 elsewhere.*

(ii) *If* $(a_1, \ldots a_p) \in R(w)$ *then*

$$c(w) = \sum_{i=1}^{p} s_{a_p} \ldots s_{a_{i+1}}(\epsilon_{a_i}).$$

Proof: (i) follows from (1.21)(i), and (ii) follows from (i) by induction on p. ‖

(1.24) *Let* $i \geq 1$. *Then*

$$c_i(w) > c_{i+1}(w) \iff w(i) > w(i+1).$$

Proof: Suppose that $w(i) > w(i+1)$. Then the $(i+1)^{\text{th}}$ row of $I(w)$ is strictly contained in the i^{th} row, whence $c_i(w) > c_{i+1}(w)$. Conversely, if $w(i) < w(i+1)$, then the i^{th} row of $I(w)$ is contained in the $(i+1)^{\text{th}}$ row, so that $c_i(w) \leq c_{i+1}(w)$. ‖

To compute the code of w^{-1} in terms of the code $(c_1, c_2, \ldots)$ of w, we introduce the following notation. If $u = (u_1, u_2, \ldots)$ is any sequence and r is an integer ≥ 0, let

$$\zeta_r u = (u_1, u_2, \ldots, u_r, 0, u_{r+1}, u_{r+2}, \ldots)$$

so that the operation ζ_r introduces a zero *after* the r^{th} place. Then we have

(1.25)
$$c(w^{-1}) = \sum_{i \geq 1} \zeta_{c_1} \cdots \zeta_{c_{i-1}} (1^{c_i})$$

where (1^{c_i}) *is the sequence consisting of* c_i $1's$.

Proof: By induction on the length of $c(w)$ it is enough to show that if w_1 is the permutation whose code is $(c_2, c_3, \ldots)$ then

(1)
$$c(w^{-1}) = (1^{c_1}) + \zeta_{c_1} c(w_1^{-1}).$$

Now the diagram of w_1 is obtained from that of w by deleting the first row (of length c_1) and then moving each column after the c_1^{th} one space to the left. On reading the diagrams of w and w_1 by columns, we obtain (1). ‖

The *shape* $\lambda(w)$ of a permutation w is the partition whose parts are the non-zero $c_i(w)$, arranged in weakly decreasing order. We have

$$|\lambda(w)| = \text{Card } D(w) = \ell(w).$$

Next, recall that for two partitions $\lambda = (\lambda_1, \lambda_2, \ldots)$ and $\mu = (\mu_1, \mu_2, \ldots)$ the relation $\lambda \geq \mu$ means that $|\lambda| = |\mu|$ and $\lambda_1 + \cdots + \lambda_i \geq \mu_1 + \cdots + \mu_i$ for all $i \geq 1$ [M, Ch.I]. With this understood, the shapes of w and w^{-1} are related by

(1.26) $\lambda(w)' \geq \lambda(w^{-1})$.

Proof: Let $\lambda = \lambda(w), \mu = \lambda(w^{-1})$. Define a matrix $M = (m_{ij})$ as follows: $m_{ij} = 1$ if $(i,j) \in D(w)$, and $m_{ij} = 0$ otherwise. Then M is a $(0,1)$ matrix with row-sums $\lambda_1, \lambda_2, \ldots$ in some order, and column-sums $\mu_1, \mu_2, \ldots$ in some order. Hence (see e.g. [M, Ch.I, §6]) we have $\lambda' \geq \mu$. ‖

Vexillary permutations

Special interest attaches to those permutations $w \in S_\infty$ for which $\lambda(w)' = \lambda(w^{-1})$. They may be characterized in various ways:

(1.27) *The following conditions on a permutation $w \in S_\infty$ are equivalent:*

- (i) *the set of rows of $D(w)$ is totally ordered by inclusion*;
- (i)$'$ *the set of rows of $I(w)$ is totally ordered by inclusion*;
- (ii) *the set of columns of $D(w)$ is totally ordered by inclusion*;
- (ii)$'$ *the set of columns of $I(w)$ is totally ordered by inclusion*;
- (iii) *there do not exist a, b, c, d such that $1 \leq a < b < c < d$ and $w(b) < w(a) < w(d) < w(c)$*;
- (iv) *there exist $u, v \in S_\infty$ such that $(u \times v)D(w)$ is the diagram $D(\lambda)$ of a partition λ*;
- (v) $\lambda(w)' = \lambda(w^{-1})$.

Proof: Since $D(w) = (1 \times w)I(w)$ it is clear that (i)$\Leftrightarrow$(i)$'$ and (ii)$\Leftrightarrow$(ii)$'$. Morever (i) $\Leftrightarrow$ (ii), for either is false if and only if there exist $a, \beta, c, \delta \in [1, n]$ such that $a < c$, $\beta < \delta$ and $(a, \beta), (c, \delta)$ belong to $D(w)$, whilst (a, δ) and (c, β) do not. Let $b = w^{-1}(\beta)$ and $d = w^{-1}(\delta)$; then we have $a < b < c < d$ and $w(b) < w(a) < w(d) < w(c)$. Thus (i), (ii) and (iii) are all equivalent.

Next, it is clear that the conjunction of (i) and (ii) is equivalent to (iv). Thus it remains to show that (iv) and (v) are equivalent. If (iv) is satisfied, then $\lambda(w) = \lambda$ and $\lambda(w^{-1}) = \lambda'$, whence (v) is satisfied. Conversely, if $\lambda(w) = \lambda$ and $\lambda(w^{-1}) = \lambda'$, then $D(w)$ can be brought into coincidence with $D(\lambda)$ by suitable permutations of the rows and of the columns, whence (iv) is satisfied. ‖

An element $w \in S_\infty$ is said to be *vexillary* it it satisfies the equivalent conditions of (1.27). By (1.27) (iii), the first non-vexillary permutation is (2143) in S_4.

For each $w \in S_n$ let

$$\overline{w} = w_0 w w_0$$

where as before $w_0 = (n, n-1, \dots, 2, 1)$ is the longest element of S_n. Then

(1.28) (i) $\ell(\overline{w}) = \ell(w)$.

(ii) *$I(\overline{w})$ is the reflection of $I(w)$ in the "antidiagonal" $i + j = n + 1$.*

(iii) $\lambda(\overline{w}) = \lambda(w)'$.

Proof: (i) follows from (1.6) (or from (ii) below).

(ii) If $i < j$ then

$$\begin{aligned}(i,j) \in I(\overline{w}) &\iff w_0 w w_0(i) > w_0 w w_0(j) \\ &\iff w(n+1-i) < w(n+1-j) \\ &\iff (n+1-j, n+1-i) \in I(w).\end{aligned}$$

(iii) now follows from (ii). ‖

From (1.27) and (1.28) it follows that

(1.29) *w is vexillary $\iff$ w^{-1} is vexillary $\iff$ $\overline{w}$ is vexillary.*

Dominant permutations

We consider next two particular types of vexillary permutations.

(1.30) *Let $w \in S_\infty$. Then the following conditions are equivalent:*

(i) *the code of w is a partition;*

(ii) *the code of w^{-1} is a partition;*

(iii) *$D(w)$ is the diagram of a partition.*

Proof: Clearly (iii) implies (i) and (ii).

Conversely, suppose that $c(w)$ is a partition $\lambda = (\lambda_1, \dots, \lambda_m)$, where $\lambda_1 \geq \dots \geq \lambda_m \geq 0$. We shall show by induction on i that

$$(i,j) \in D(w) \iff 1 \leq j \leq \lambda_i.$$

This is true for $i = 1$, so assume that $1 < i \leq m$ and that the statement is true for $i - 1$. Then we have $w(k) \leq \lambda_{i-1}$ for $1 \leq k \leq i - 1$, and $w(k) = \lambda_{i-1}$ for some $k \leq i - 1$. Since $\lambda_i \leq \lambda_{i-1}$ it follows that the i^{th} row of $D(w)$ consists of the points (i, j), $1 \leq j \leq \lambda_i$, as required. Hence (i) implies (iii), and the same argument applied to w^{-1} shows that if the code of w^{-1} is a partition, then $D(w^{-1})$ is the diagram of a partition. Hence so is $D(w)$, by (1.21) (i), and the proof is complete. ‖

A permutation is said to be *dominant* if it satisfies the equivalent conditions of (1.30). Dominant permutations are clearly vexillary, and w is dominant if and only if w^{-1} is dominant.

Grassmannian permutations

(1.31) Let $w \in S_\infty$. *Then the following conditions are equivalent:*

(i) $c_1(w) \leq \ldots \leq c_r(w)$ and $c_i(w) = 0$ for $i > r$;

(ii) $w(i) < w(i+1)$ *unless* $i = r$.

Proof: (i) $\Rightarrow$ (ii). By (1.15) we have $w(1) < \ldots < w(r)$ and $w(r+1) < \ldots < w(n)$.

(ii) $\Rightarrow$ (i). We have

$$c(w) = (w(1) - 1, \ldots, w(r) - r). \;\|$$

If w satisfies the equivalent conditions of (1.31), w is called a *Grassmannian permutation*. By (1.27)(iii), Grassmannian permutations are vexillary, and $w \in S_n$ is Grassmannian if and only if $\overline{w} = w_0 w w_0$ is Grassmannian.

Enumeration of vexillary permutations

Let w be a permutation, $c = c(w) = (c_1, c_2, \ldots)$ its code. Consider the following two conditions on the sequence c:

(V1) *If* $i < j$ *and* $c_i > c_j$, *then*

$$\text{Card}\,\{k : i < k < j \text{ and } c_k < c_j\} \leq c_i - c_j;$$

(V2) *If* $i < j$ *and* $c_i \leq c_j$, *then* $c_k \geq c_i$ *whenever* $i < k < j$.

(1.32) *A permutation* w *is vexillary if and only if its code* $c(w)$ *satisfies* (V1) *and* (V2).

Proof: For each $i \geq 1$, let

$$\rho_i = \{j : (i,j) \in D(w)\}$$

be the i^{th} row of $D(w)$.

Suppose first that w is vexillary with code $c = (c_1, c_2, \ldots)$. Let $i < k < j$ be such that $c_i \geq c_j > c_k$. Then $\rho_i \supseteq \rho_j \supset \rho_k$ (where $\supset$ denotes strict containment), hence there exists $t \in \rho_j, t \notin \rho_k$. Let $s = w(k)$, then $s < t$ and (since $t \in \rho_i$) we have $s \in \rho_i$ and $s \notin \rho_j$. Hence for fixed (i,j) such that $i < j$ and $c_i \geq c_j$, the number of k between i and j such that $c_j > c_k$ is at most $\text{Card}(\rho_i - \rho_j) = c_i - c_j$, so that (V1) is satisfied.

Next let w be vexillary, $i < k < j$ and $c_i < c_j$, so that $\rho_i \subseteq \rho_j$. Let $s \in \rho_i$. If $s \notin \rho_k$ then $w(k) \leq s < w(i)$, so that $w(k)$ lies in ρ_i but not in ρ_j, which is impossible. Hence $s \in \rho_k$ and therefore $\rho_i \subseteq \rho_k$. So we have $c_i \leq c_k$, and (V2) is satisfied.

Conversely, suppose that the code c of w satisfies (V1) and (V2). Then so does the sequence $(c_2, c_3, \ldots)$ and we may therefore assume that the set $\{\rho_2, \rho_3, \ldots\}$ is totally ordered by inclusion.

Let $j > 1$ and suppose first that $c_1 \geq c_j$. If $\rho_1 \not\supseteq \rho_j$, there exists $s \in \rho_j$ such that $s \notin \rho_1$, so that $w(1) < s < w(j)$. There are at least $c_1 - c_j + 1$ elements $t \in \rho_1$ such that $t \notin \rho_j$, and since each such t satisfies $t < w(1) < w(j)$, it is of the form $t = w(k)$ for some k between 1 and j. Since $w(k) = t < w(1) < s$, it follows that $s \notin \rho_k$. Since either $\rho_k \subseteq \rho_j$ or $\rho_j \subseteq \rho_k$, we conclude that $\rho_k \subset \rho_j$ (strict inclusion) and hence that $c_k < c_j$. Hence there are at least $c_1 - c_j + 1$ values of k between 1 and j for which $c_k < c_j$, contradicting (V1). Hence $\rho_1 \supseteq \rho_j$.

Finally, let $j > 1$ and $c_1 < c_j$, so that $w(1) < w(j)$. If $\rho_1 \not\subseteq \rho_j$ there exists $s \in \rho_1$ such that $s \notin \rho_j$; we have $s = w(k)$ for some k between 1 and j, and since $w(k) < w(1)$ we have $c_k < c_1$, contradicting (V2). Hence $\rho_1 \subseteq \rho_j$ in this case, and the proof is complete. ‖

Remark. It is stated in [LS4, prop. 2.4] that w is vexillary if and only if $c(w)$ satisfies (V1) and

(V3) *If* $c_i > c_{i+1}$ *for some* $i \geq 1$, *then* $c_i > c_j$ *for all* $j > i$.

Since (V3) is implied by (V2), it follows from (1.32) that every vexillary code satisfies (V1) and (V3). However, the conjuction of (V1) and (V3) is not sufficient for vexillarity: for example, the permutation $w = (2571634)$ is not vexillary (since e.g. it contains the subword 2163) but its code is $c = (13402)$, which satisfies (V1) and (V3) (but not (V2)).

Let w be a permutation with code $c(w) = (c_1, c_2, \ldots)$. For each $i \geq 1$ such that $c_i \neq 0$, let

$$e_i = max\{j : j \geq i \text{ and } c_j \geq c_i\}.$$

Arrange the numbers e_i in increasing order of magnitude, say $\phi_1 \leq \ldots \leq \phi_m$. The sequence

$$\phi(w) = (\phi_1, \ldots, \phi_m)$$

is called the *flag* of w. It is a sequence of length equal to $\ell(\lambda)$, where λ is the shape of w.

Remark. There is another definition of the flag of a permutation w, due to M. Wachs [W]. For each $i \geq 1$ such that $c_i \neq 0$, let

$$d_i = min\{j : j > i \text{ and } w(j) < w(i)\}.$$

Arrange the numbers $d_i - 1$ in increasing order of magnitude, say $\phi_1^* \leq \ldots \leq \phi_m^*$, and let

$$\phi^*(w) = (\phi_1^*, \ldots, \phi_m^*).$$

These two notions are not equivalent. In fact

(1.33) (J. Alfano) *We have* $\phi(w) = \phi^*(w)$ *if and only if the permutation* w *satisfies* (V2).

Proof: If $c_i \neq 0$ we have $w(j) > w(i)$ for $i < j < d_i$, and hence $c_j \geq c_i$ for these values of j. Hence $d_i - 1 \leq e_i$ in all cases, and we shall have $\phi(w) = \phi^*(w)$ if and only if $d_i - 1 = e_i$ for each i. But this condition means that, for each $i \geq 1$, the set of $j \geq i$ such that $c_j \geq c_i$ is an *interval*; and this is just a restatement of the condition (V2). ∥

We shall show that a vexillary permutation is uniquely determined by its *shape* $\lambda(w)$ and its *flag* $\phi(w)$.

Let us write $\lambda = \lambda(w)$ in the form

(1.34) $$\lambda = (p_1^{m_1}, p_2^{m_2}, \ldots, p_k^{m_k})$$

where $p_1 > p_2 > \ldots > p_k > 0$ and each $m_i \geq 1$. For $1 \leq r \leq k$ let

$$f_r = max\{j : c_j \geq p_r\}$$

so that $f_1 \leq \ldots \leq f_k$. If $c = (c_1, c_2, \ldots)$ is the code of w, each nonzero c_i is equal to p_r for some r, and

$$e_i = max\{j : j \geq i \text{ and } c_j \geq p_r\} = f_r.$$

It follows that (whether w is vexillary or not)

(1.35) $$\phi(w) = (f_1^{m_1}, f_2^{m_2}, \ldots, f_k^{m_k}).$$

Moreover we must have

(1.36) $$f_r \geq m_1 + \cdots + m_r \qquad (1 \leq r \leq k)$$

since in the sequence $(c_1, c_2, \ldots)$ there are $m_1 + \cdots + m_r$ terms $\geq p_r$, and they must all occur in the first f_r places of the sequence.

(1.37) *Suppose w is a vexillary permutation with shape λ and flag ϕ given by (1.34) and (1.35). Then the f_r must satisfy the inequalities*

$$0 \leq f_r - f_{r-1} \leq m_r + p_{r-1} - p_r.$$

Proof: If $f_{r-1} = f_r$ there is nothing to prove, so assume that $f_{r-1} < f_r$ and therefore $c_{f_r} = p_r$. Let

$$s = max\{i : c_i = p_{r-1}\} \leq f_{r-1}.$$

Since $c_s = p_{r-1} > p_r = c_{f_r}$ and w is vexillary, we have by (V1)

(1) $$\text{Card } \{k : s < k \leq f_r \text{ and } c_k < p_r\} \leq p_{r-1} - p_r.$$

Also

$$(\mathbf{2}) \qquad \text{Card}\,\{k : s < k \le f_r \text{ and } c_k = p_r\} \le m_r,$$

since exactly m_r terms of the sequence c are equal to p_r.

Finally we have

$$(\mathbf{3}) \qquad \text{Card}\,\{k : s < k \le f_r \text{ and } c_k > p_r\} = f_{r-1} - s$$

because $c_k \le p_r$ for all $k > f_{r-1}$, and $c_k \ge p_{r-1}$ for all k such that $s < k \le f_{r-1}$, by virtue of (V2).

From (1), (2), and (3) we deduce that

$$f_r - s \le p_{r-1} - p_r + m_r + f_{r-1} - s$$

which proves (1.37). ∥

(1.38) *For each sequence* $(f_1, \ldots, f_k)$ *satisfying* (1.36) *and* (1.37) *there is a unique vexillary permutation* w *with shape* λ *and flag* $\phi = (f_1^{m_1}, \ldots, f_k^{m_k})$. *The code* c *of* w *is constructed as follows: first the* m_1 *entries equal to* p_1 *are inserted at the right-hand end of the interval* $[1, f_1]$;*then the* m_2 *entries in* c *equal to* p_2 *are inserted in the rightmost available spaces in the interval* $[1, f_2]$, *and so on: for each* $r \ge 1$, *when all the terms* $> p_r$ *in the sequence* c *have been inserted, the* m_r *entries equal to* p_r *are inserted in the rightmost available spaces of the interval* $[1, f_r]$.

Proof: Suppose first that w is vexillary. If $1 \le i \le f_r$ and $c_i = p_r$, then by (V2) we have $c_j \ge p_r$ for all j such that $i < j < f_r$. Hence the entries equal to p_r in the sequence c must be inserted as described above.

Conversely, if the sequence c is constructed as above, we claim that c satisfies (V1) and (V2), and hence w is vexillary by (1.32). Suppose first that $i < j$ and $c_i \ge c_j$: say $c_i = p_r, c_j = p_s, r \le s$. Then the number of k such that $i < k < j$ and $c_k < p_s$ is equal to the number of blank spaces in the interval $[f_r, f_s]$ after all the entries p_i, $r+1 \le i \le s$ have been inserted, hence is at most

$$f_s - f_r - (m_{r+1} + \cdots + m_s)$$

which by (1.37) is $\le p_r - p_s$. Hence the sequence c satisfies (V1). Suppose next that $i < j$ and $c_i < c_j$: say $c_i = p_s, c_j = p_r$ with $r < s$. Then we have $j \le f_r \le f_s$. From the definition of the sequence c, it follows that for each k such that $i \le k \le f_s$ we have $c_k \ge p_s$, and hence $c_k \ge c_i$ whenever $i < k < j$. Consequently the condition (V2) is satisified, and the proof is complete. ∥

If w is a permutation and $r \geq 0$, we denote by $1_r \times w$ the permutation

$$1_r \times w = (1, 2, \ldots, r, r + w(1), r + w(2), \ldots).$$

Let us say that two permutations w, w' are *diagonally equivalent* if either $w' = 1_r \times w$ or $w = 1_r \times w'$ for some $r \geq 0$. Graphically, this means that the diagram of w' can be brought into coincidence with that of w by a translation along the diagonal $i = j$, and w' is vexillary if and only if w is vexillary. The equivalence classes of vexillary permutations of a given shape λ are then determined by the differences $f_r - f_{r-1}$ $(2 \leq r \leq k)$, and hence it follows from (1.37) and (1.38) that

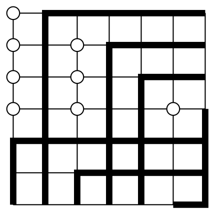
(1.39) *The number of diagonal equivalence classes of vexillary permutations of shape* $\lambda = (p_1^{m_1}, \ldots, p_k^{m_k})$ *is*

$$\prod_{r=2}^{k} (p_{r-1} - p_r + m_r + 1).$$

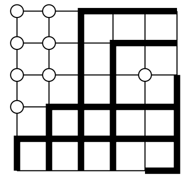
We may remark that this number is the product of the hook lengths at the re-entrant nodes of the border of the diagram of λ (i.e., the nodes with coordinates $(m_1 + \cdots + m_{r-1}, p_r)$, $2 \leq r \leq k$).

Example. If $\lambda = (32^21)$ the flag $\phi = (f_1, f_2^2, f_3)$ must satisfy $0 \leq f_2 - f_1 \leq 3$, $0 \leq f_3 - f_2 \leq 2$. Hence there are $(3+1)(2+1) = 12$ vexillary classes, and the representatives of these classes for which $w(1) \neq 1$ (or equivalently $c_1(w) \neq 0$) are as follows:

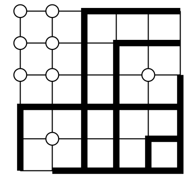
$\phi(w)$	$c(w)$	w
4444	1223	2457136
3444	1232	246513
2444	1322	254613
1444	3122	425613
3334	2231	346215
2334	2321	35421
1334	3221	43521
1445	30221	415632
3335	22301	346152
2335	23201	354162
1335	32201	435162
1446	302201	4156273



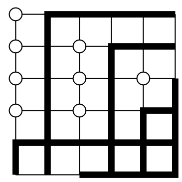
1223



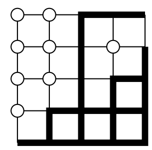
2231



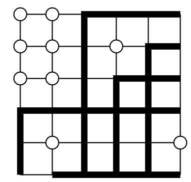
22301



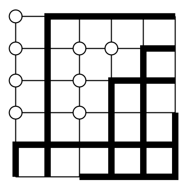
1232



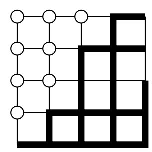
2321



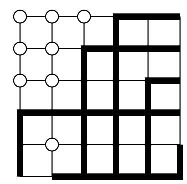
23201



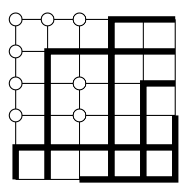
1322



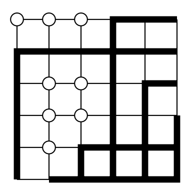
3221



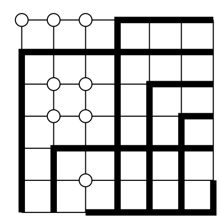
32201



3122



302201



302201

Let $\lambda = (p_1^{m_1}, \ldots, p_k^{m_k})$ as before and let

$$\lambda' = (q_1^{n_1}, q_2^{n_2}, \ldots, q_k^{n_k})$$

be the conjugate partition, where $q_1 > q_2 > \ldots > q_k > 0$ and each $n_i \geq 1$. We have

$$\textbf{(1.40)} \qquad \begin{cases} p_r = n_1 + \cdots + n_s, \\ q_r = m_1 + \cdots + m_s, \end{cases}$$

where $s = k + 1 - r$ $(1 \leq r \leq k)$. The border of the diagram of λ is a staircase with risers of heights $m_1, m_2, \ldots, m_k$ (starting from the top) and treads of lengths $n_1, n_2, \ldots, n_k$ (starting at the bottom).

Recall (1.27) that if w is vexillary of shape λ, then w^{-1} is vexillary of shape λ'.

(1.41) *Let w be a vexillary permutation of shape λ and flag $\phi(w) = (f_1^{m_1}, \ldots, f_k^{m_k})$. Then the flag of w^{-1} is*

$$\phi(w^{-1}) = (g_1^{n_1}, \ldots, g_k^{n_k})$$

where

$$(*) \qquad g_i + q_i = f_{k+1-i} + p_{k+1-i} \qquad (1 \le i \le k).$$

Proof: We proceed by induction on $\ell(w) = |\lambda|$. Let $c = (c_1, c_2, \ldots)$ be the code of w, and let w' be the permutation with code $c' = (c_2, c_3, \ldots)$. We may assume that $c_1 \neq 0$. Then $c_1 = p_r$ for some r, and we have

$$\lambda(w') = (p_1^{m_1}, \ldots, p_r^{m_r - 1}, \ldots, p_k^{m_k}),$$

$$\phi(w') = ((f_1 - 1)^{m_1}, \ldots, (f_r - 1)^{m_r - 1}, \ldots, (f_k - 1)^{m_k}).$$

Since w is vexillary, its code c satisfies the conditions (V1) and (V2). Hence c' also satisfies these conditions, and therefore w' is vexillary. It follows that $\lambda(w'^{-1}) = \lambda(w')'$, so that

$$\lambda(w'^{-1}) = ((q_1 - 1)^{n_1}, \ldots, (q_s - 1)^{n_s}, q_{s+1}^{n_{s+1}}, \ldots, q_k^{n_k})$$

where $s = k + 1 - r$. We have $\ell(w') = \ell(w) - c_1$, so that the inductive hypothesis applies to w'. Hence if $g_1, \ldots, g_k$ are defined by the formula $(*)$, we have

$$(\mathbf{1}) \qquad \phi(w'^{-1}) = (g_1^{n_1}, \ldots, g_s^{n_s}, (g_{s+1} - 1)^{n_{s+1}}, \ldots, (g_k - 1)^{n_k}).$$

But if w'^{-1} has code $c(w'^{-1}) = (d_1, d_2, \ldots)$ then by (1.25) we have

$$(\mathbf{2}) \qquad c(w^{-1}) = (d_1 + 1, \ldots, d_{p_r} + 1, 0, d_{p_r+1}, d_{p_r+2}, \ldots).$$

From (1) and (2) and (1.40) it follows that

$$\phi(w^{-1}) = (g_1^{n_1}, \ldots, g_s^{n_s}, g_{s+1}^{n_{s+1}}, \ldots, g_k^{n_k})$$

as required. ‖

If $w \in S_n$, let $\overline{w}_n = w_0 w w_0$, where w_0 is the longest element in S_n. If w is vexillary, of shape λ, then $\overline{w}_n$ is vexillary of shape λ',by (1.27) and (1.28). Let

$$\phi(\overline{w}_n) = (\overline{f_1}^{n_1}, \ldots, \overline{f_k}^{n_k})$$

be the flag of $\overline{w}_n$. Then we have

$$(\mathbf{1.42}) \qquad \overline{f_i} = n - f_{k+1-i} \qquad (1 \le i \le k).$$

For once we shall leave the proof to the reader.

Let N_n denote the number of non-vexillary $w \in S_n$, and let

$$P_n = N_n/n!$$

be the probability that an element of S_n is non-vexillary. The first few values of N_n and P_n are

n	N_n	P_n
1	0	0
2	0	0
3	0	0
4	1	.042
5	17	.142
6	207	.288
7	2279*	.452

If we divide up the sequence $(w(1), \ldots, w(n))$ into consecutive blocks of length 4, and observe that the probability that such a block satisfies the vexillarity condition (1.27)(iii) is 23/24 (because S_4 contains only one non-vexillary permutation), we see that the probability that $w \in S_n$ is vexillary is at most $(23/24)^{[n/4]}$, hence decreases exponentially to zero. (A. Lascoux.) Thus the vexillary permutations in S_n become sparser and sparser as n increases.

Instead of counting non-vexillary permutations, we may attempt to count vexillary permutations. Let us say that a permutation $w \in S_n$ is *primitive* if $w(1) \neq 1$ and $w(n) \neq n$. For each $n \geq 1$, let V_n (resp. U_n) denote the number of vexillary (resp. primitive vexillary) permutations $w \in S_n$. Since each primitive vexillary $w \in S_n$ gives rise to $r+1$ imprimitive vexillary permutations in S_{n+r}, namely $1_p \times w \times 1_q$ where $p, q \geq 0$ and $p + q = r$, it follows that

$$V_n = 1 + U_n + 2U_{n-1} + 3U_{n-2} + \cdots$$

Hence the generating functions

$$V(t) = \sum_{n \geq 1} V_n t^n$$

$$U(t) = \sum_{n \geq 1} U_n t^n$$

are related by

$$V(t) = \frac{t}{1-t} + \frac{U(t)}{(1-t)^2}. \tag{1.43}$$

* N_7 was computed by A. Garsia. I would guess that N_8 is of the order of 24000.

For each partition $\lambda \neq 0$, let $U_{n,\lambda}$ denote the number of primitive vexillary permutations of shape λ in S_n, and let

$$U_\lambda(t) = \sum_{n \geq 1} U_{n,\lambda} t^n,$$

so that

$$U(t) = \sum_{\lambda \neq 0} U_\lambda(t). \tag{1.44}$$

Each $U_\lambda(t)$ is a polynomial, and we shall now show how to compute it. Write λ in the form

$$\lambda = (p_1^{m_1}, p_2^{m_2}, \ldots, p_k^{m_k})$$

as before, where $p_1 > p_2 > \ldots > p_k > 0$. By (1.37) a vexillary permutation w of shape λ is uniquely determined by its flag $\phi(w) = (f_1^{m_1}, \ldots, f_k^{m_k})$, where $(f_1, \ldots, f_k)$ is any vector of positive integers satisfying the inequalities (1.36),(1.37):

$$f_r \geq m_1 + \cdots + m_r \qquad (1 \leq r \leq k),$$

$$0 < f_r - f_{r-1} \leq m_r + p_{r-1} - p_r \qquad (2 \leq r \leq k).$$

Moreover we shall have $w(1) \neq 1$ if and only if the first element of the code of w is not zero, and this will be the case if and only if

$$f_r = m_1 + \ldots + m_r \qquad \text{for some } r = 1, \ldots, k. \tag{1}$$

In general, if $c = (c_1, c_2, \ldots)$ is the code of a permutation w, then $w \in S_n$ if and only if $n \geq c_i + i$ for $1 \leq i \leq r$, where r is the length of c. In other words, the least n for which $w \in S_n$ is $n = max\{c_i + i : 1 \leq i \leq r\}$. In the case of a vexillary permutation w as above, with flag $(f_1^{m_1}, \ldots, f_k^{m_k})$, the numbers $c_i + i$ will increase strictly as i runs through each non-empty interval $[f_{r-1}+1, f_r]$ $(r = 1, \ldots, k)$, and hence w will be primitive in S_n if and only if w satisfies (1) above and

$$n = max\{p_r + f_r : 1 \leq r \leq k\}. \tag{2}$$

Let $\pi_r = m_1 + \cdots + m_r$ for $1 \leq r \leq k$ and put

$$u_r = f_r - \pi_r$$

so that $u_r \geq 0$ for each r. From (1.36) we have

$$\pi_1 + u_1 \leq \pi_2 + u_2 \leq \ldots \leq \pi_k + u_k \tag{3}$$

and

$$\begin{aligned} m_r + p_{r-1} - p_r \geq f_r - f_{r-1} \\ &= (u_r + \pi_r) - (u_{r-1} + \pi_{r-1}) \\ &= m_r + u_r - u_{r-1} \end{aligned}$$

so that

$$p_1 + u_1 \geq p_2 + u_2 \geq \ldots \geq p_k + u_k. \tag{4}$$

It now follows that

$$U_\lambda(t) = \sum_u t^{max\{p_r + \pi_r + u_r : 1 \leq r \leq k\}} \tag{1.45}$$

summed over the integer vectors $u = (u_1, \ldots, u_k) \in \mathbf{N}^k$ having at least one zero component, and satisfying the inequalities (3), (4) above. We have

$$U_\lambda(1) = \prod_{r=2}^{k} (m_r + p_{r-1} - p_r + 1)$$

and

$$U_\lambda(t) = U_{\lambda'}(t)$$

(since $w \in S_n$ is primitive vexillary of shape λ if and only if w^{-1} is primitive vexillary of shape λ').

Added in proof

Julian West, a student of R. Stanley, has recently shown that

$$V_n = \sum_{\substack{|\lambda|=n \\ \ell(\lambda) \leq 3}} (f^\lambda)^2 \tag{1}$$

where f^λ is the degree of the irreducible representation of the symmetric group S_n indexed by the partition λ. From this and results of A. Regev (Advances in Math. **41** (1981) 115–136) it follows that

$$V_n \sim c 9^n n^{-4} \tag{2}$$

as $n \to \infty$, where c is a constant that Regev determines explicitly.

The formula (1) gives that $N_8 = 24553$.

Chapter II

Divided differences

If f is a function of x and y (and possibly other variables), let

$$\partial_{xy} f = \frac{f(x,y) - f(y,x)}{x - y}$$

("divided difference"). Equivalently

$$\partial_{xy} f = (x-y)^{-1}(1 - s_{xy})$$

where s_{xy} interchanges x and y. The operator ∂_{xy} takes polynomials to polynomials, and has degree -1 (i.e., if f is homogeneous of degree d, then $\partial_{xy} f$ is homogeneous of degree $d-1$). Explicitly, if $f = x^r y^s$ we have

(2.1)
$$\begin{aligned}\partial_{xy}(x^r y^s) &= \frac{x^r y^s - x^s y^r}{x - y} \\ &= \sigma(r-s) \sum x^p y^q\end{aligned}$$

where the sum is over (p,q) such that $p + q = r + s - 1$ and $max(p,q) < max(r,s)$, and $\sigma(r-s)$ is $+1, 0$ or -1 according as $r - s$ is positive, zero or negative.

On a product fg, ∂_{xy} acts according to the rule

(2.2)
$$\partial_{xy}(fg) = (\partial_{xy} f) g + (s_{xy} f)(\partial_{xy} g).$$

In particular we have

(2.2′)
$$\partial_{xy}(fg) = f \partial_{xy} g$$

if $f(x,y) = f(y,x)$.

(2.3) (i) $\partial_{xy} s_{xy} = -\partial_{xy}, \qquad s_{xy}\partial_{xy} = \partial_{xy},$

(ii) $\partial_{xy}^2 = 0,$

(iii) $\partial_{xy}\partial_{yz}\partial_{xy} = \partial_{yz}\partial_{xy}\partial_{yz}.$

Proof: (i) and (ii) are immediate from the definitions, and (iii) is verified by direct calculation: each side is equal to

$$(x-y)^{-1}(x-z)^{-1}(y-z)^{-1}\sum_{w\in S_3}\epsilon(w)w,$$

where the symmetric group S_3 permutes x, y and z, and $\epsilon(w)$ is the sign of the permutation w. ‖

Let $x_1, x_2, \ldots, x_n, \ldots$ be independent variables, and let

$$P_n = \mathbf{Z}[x_1, x_2, \ldots, x_n]$$

for each $n \geq 1$, and

$$\begin{aligned} P_\infty &= \mathbf{Z}[x_1, x_2, \ldots] \\ &= \bigcup_{n=1}^{\infty} P_n. \end{aligned}$$

For each $i \geq 1$ let

$$\partial_i = \partial_{x_i, x_{i+1}}.$$

Each ∂_i is a linear operator on P_∞ (and on P_n for $n > i$) of degree -1. From (2.3) we have (compare with (1.1))

$$\textbf{(2.4)} \qquad \begin{cases} \partial_i^2 = 0, & \\ \partial_i\partial_j = \partial_j\partial_i & \text{if } |i-j| > 1, \\ \partial_i\partial_{i+1}\partial_i = \partial_{i+1}\partial_i\partial_{i+1} & \end{cases}$$

For any sequence $\boldsymbol{a} = (a_1, \ldots, a_p)$ of positive integers, we define

$$\partial_{\boldsymbol{a}} = \partial_{a_1} \ldots \partial_{a_p}.$$

Recall that if w is any permutation, $R(w)$ denotes the set of *reduced words* for w, i.e. sequences $(a_1, \ldots, a_p)$ such that $w = s_{a_1} \ldots s_{a_p}$ and $p = \ell(w)$.

(2.5) *If* $\boldsymbol{a}, \boldsymbol{b} \in R(w)$ *then* $\partial_{\boldsymbol{a}} = \partial_{\boldsymbol{b}}$.

Proof: We proceed by induction on $p = \ell(w)$. Let us write $\boldsymbol{a} \equiv \boldsymbol{b}$ to mean that $\partial_{\boldsymbol{a}} = \partial_{\boldsymbol{b}}$. The inductive hypothesis then implies that

$$(*) \qquad \boldsymbol{a} \equiv \boldsymbol{b} \text{ if either } a_1 = b_1 \text{ or } a_p = b_p.$$

By the exchange lemma (1.8) we have

$$\boldsymbol{c}_i = (b_1, a_1, \ldots, \hat{a}_i, \ldots, a_p) \in R(w)$$

for some $i = 1, \dots, p$. If $i \neq p$ then $\boldsymbol{b} \equiv \boldsymbol{c}_i \equiv \boldsymbol{a}$ by virtue of $(*)$, so that $\boldsymbol{a} \equiv \boldsymbol{b}$. If $i = p$ and $|b_1 - a_1| > 1$ then by (2.4) and (1.1)

$$\boldsymbol{c}'_p = (a_1, b_1, a_2, \dots, a_{p-1}) \in R(w)$$

and $\boldsymbol{a} \equiv \boldsymbol{c}'_p \equiv \boldsymbol{c}_p \equiv \boldsymbol{b}$, so that again $\boldsymbol{a} \equiv \boldsymbol{b}$.

Finally, if $i = p$ and $|b_1 - a_1| = 1$, we apply the exchange lemma again, this time to $\boldsymbol{c}_p$ and $\boldsymbol{a}$; this shows that

$$\boldsymbol{d}_i = (a_1, b_1, a_1, \dots, \hat{a}_i, \dots, a_{p-1}) \in R(w)$$

for some $i = 2, \dots, p-1$. But then by (2.4) and (1.1) we have

$$\boldsymbol{d}'_i = (b_1, a_1, b_1, a_2, \dots, \hat{a}_i, \dots, a_{p-1}) \in R(w)$$

and $\boldsymbol{a} \equiv \boldsymbol{d}_i \equiv \boldsymbol{d}'_i \equiv \boldsymbol{b}$. Hence $\boldsymbol{a} \equiv \boldsymbol{b}$ in all cases. ‖

Remark. For any permutation w, let $GR(w)$ denote the graph whose vertices are the reduced words for w, and in which a reduced word $\boldsymbol{a}$ is joined by an edge to each of the words obtained from $\boldsymbol{a}$ by either interchanging two consecutive terms i, j such that $|i-j| > 1$, or by replacing three consecutive terms i, j, i such that $|i-j| = 1$ by j, i, j. Then the proof of (2.5) shows that

(2.5') *The graph* $GR(w)$ *is connected.* ‖

From (2.5) it follows that we may define

$$\partial_w = \partial_{\boldsymbol{a}}$$

unambiguously, where $\boldsymbol{a}$ is any reduced word for w. By $(2.2')$, the operators ∂_w for $w \in S_n$ are Λ_n linear, where

$$\Lambda_n = \mathbf{Z}[x_1, \dots, x_n]^{S_n} \subset P_n$$

is the ring of symmetric polynomials in $x_1, \dots, x_n$.

A sequence $\boldsymbol{a} = (a_1, \dots, a_p)$ will be said to be *reduced* if $\boldsymbol{a} \in R(w)$ for some permutation w.

(2.6) *If* $\boldsymbol{a} = (a_1, \dots, a_p)$ *is not reduced, then* $\partial_{\boldsymbol{a}} = 0$.

Proof: By induction on p. If $\boldsymbol{a}' = (a_1, \dots, a_{p-1})$ is not reduced, then $\partial_{\boldsymbol{a}'} = 0$ and hence $\partial_{\boldsymbol{a}} = \partial_{\boldsymbol{a}'}\partial_{a_p} = 0$. So we may assume that $\boldsymbol{a}'$ is reduced. Let $v = s_{a_1} \dots s_{a_{p-1}}$, $w = s_{a_1} \dots s_{a_p}$. We have $\ell(v) = p-1$ and $\ell(w) \leq p-1$, hence by (1.3) $\ell(w) = p-2$, so that $\ell(v) = \ell(ws_{a_p}) = \ell(w) + 1$. Consequently $\partial_v = \partial_w \partial_{a_p}$ and therefore $\partial_{\boldsymbol{a}} = \partial_v \partial_{a_p} = \partial_w \partial^2_{a_p} = 0$. ‖

(2.7) *Let u, v be permutations. Then*

$$\partial_u \partial_v = \begin{cases} \partial_{uv} & \text{if } \ell(uv) = \ell(u) + \ell(v), \\ 0 & \text{otherwise .} \end{cases}$$

Proof: (2.5), (2.6). ‖

(2.8) *Let w be a permutation, $i \geq 1$. Then*

$$s_i \partial_w = \partial_w \iff \ell(s_i w) = \ell(w) - 1.$$

Proof: We have $s_i \partial_w = \partial_w \iff \partial_i \partial_w = 0$, hence the result follows from (2.7). ‖

As before let $w_0 = (n, n-1, \ldots, 2, 1)$ be the longest element of S_n. One element of $R(w_0)$ is the sequence

(**2.9**) $$(1, 2, \ldots, n-1, 1, 2, \ldots, n-2, \ldots, 1, 2, 3, 1, 2, 1).$$

(2.10) *We have*

$$\partial_{w_0} = a_\delta^{-1} \sum_{w \in S_n} \epsilon(w) w$$

where $a_\delta = \prod_{1 \leq i < j \leq n} (x_i - x_j)$, and $\epsilon(w) = \pm 1$ is the sign of w.

Proof: From the definition it follows that ∂_{w_0} is of the form

(**1**) $$\partial_{w_0} = \sum_{w \in S_n} c_w w$$

with coefficients c_w rational functions of $x_1, \ldots, x_n$. By (2.8) we have $s_i \partial_{w_0} = \partial_{w_0}$ for $1 \leq i \leq n-1$, so that $v \partial_{w_0} = \partial_{w_0}$ for all $v \in S_n$, and therefore

(**2**) $$\partial_{w_0} = \sum_{w \in S_n} v(c_w) v w.$$

Comparison of (1) and (2) shows that

(**3**) $$c_{vw} = v(c_w) \qquad (v, w \in S_n).$$

Hence all the coefficients c_w are determined by one of them, say c_{w_0}. From the sequence (2.9) for w_0 it is easily checked that the coefficient of w_0 in ∂_0 is

$$c_{w_0} = \epsilon(w_0) a_\delta^{-1}.$$

Hence from (3) we have

$$c_w = w w_0 (c_{w_0}) = \epsilon(w) a_\delta^{-1}$$

which proves (2.10). ‖

From (2.10) it follows that, for any $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbf{N}^n$,

(2.11) $$\partial_{w_0} x^\alpha = s_{\alpha-\delta}(x_1, \dots, x_n)$$

where x^α means $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, $\delta = (n-1, n-2, \dots, 1, 0)$ and $s_{\alpha-\delta}$ is the Schur function indexed by $\alpha - \delta$. Thus ∂_{w_0} is a Λ_n-linear mapping of P_n onto Λ_n.

For $w \in S_n$, let $\overline{w} = w_0 w w_0$. Then

(2.12) $$\partial_{\overline{w}} = \epsilon(w) w_0 \partial_w w_0.$$

Proof: From the definition of ∂_i we have

$$w_0 \partial_i w_0 = -\partial_{n-i}$$

from which (2.12) follows easily, since $w_0^2 = 1$. ‖

If f and g are polynomials in $x_1, x_2, \dots$, the expression of $\partial_w(fg)$ as a sum of polynomials $\partial_u f \cdot \partial_v g$ (i.e. the "Leibnitz formula" for ∂_w) is in general rather complicated. However, there is one case in which it is reasonably simple, namely when one of the factors f, g is linear:

(2.13) *If* $f = \sum \alpha_i x_i$ *then*

$$\partial_w(fg) = w(f)\partial_w g + \sum (\alpha_i - \alpha_j)\partial_{wt_{ij}} g$$

summed over all pairs $i < j$ *such that* $\ell(wt_{ij}) = \ell(w) - 1$, *where* t_{ij} *is the transposition that interchanges* i *and* j.

Proof: Let $(a_1, \dots, a_p)$ be a reduced word for w. Since f is linear it follows from (2.2) that

$$\begin{aligned}\partial_w(fg) &= \partial_{a_1} \cdots \partial_{a_p}(fg) \\ &= s_{a_1} \cdots s_{a_p}(f)\partial_{a_1} \cdots \partial_{a_p} g + \sum_{r=1}^{p} s_{a_1} \cdots \partial_{a_r} \cdots \ s_{a_p}(f)\partial_{a_1} \cdots \hat{\partial}_{a_r} \cdots \partial_{a_p} g.\end{aligned}$$

Now $\partial_{a_1} \cdots \hat{\partial}_{a_r} \cdots \partial_{a_p} = 0$ unless $(a_1, \dots, \hat{a}_r, \dots, a_p)$ is reduced, and then by (1.11) it is equal to ∂_{wt}, where $wt = s_{a_p} \cdots \hat{s}_{a_r} \cdots s_{a_p}$ has length $p - 1 = \ell(w) - 1$, and $t = s_{a_p} \cdots s_{a_r} \cdots s_{a_p} = t_{ij}$ where $(i, j) = s_{a_p} \cdots s_{a_{r+1}}(a_r, a_{r+1})$, so that

$$s_{a_1} \cdots s_{a_{r-1}} \partial_{a_r} s_{a_{r+1}} \cdots \ s_{a_p}(f) = \alpha_i - \alpha_j. \ \|$$

We also introduce the operators $\pi_i (i \geq 1)$ defined by

$$\pi_i f = \partial_i(x_i f).$$

In place of (2.4) we have

$$(\mathbf{2.14}) \qquad \begin{cases} \pi_i^2 = \pi_i, \\ \pi_i \pi_j = \pi_j \pi_i & \text{if } |i-j| > 1, \\ \pi_i \pi_{i+1} \pi_i = \pi_{i+1} \pi_i \pi_{i+1}. \end{cases}$$

If we define $\pi_{\boldsymbol{a}}$ to be $\pi_{a_1} \cdots \pi_{a_p}$ for any sequence $\boldsymbol{a} = (a_1, \ldots, a_p)$ of positive integers, then corresponding to (2.5) we have

(2.15) If $\boldsymbol{a}, \boldsymbol{b} \in R(w)$ then $\pi_{\boldsymbol{a}} = \pi_{\boldsymbol{b}}$.

The proof is the same as that of (2.5), and rests only on the second and third of the relations (2.14). From (2.15) it follows that we may define

$$\pi_w = \pi_{\boldsymbol{a}}$$

unambiguously, where $\boldsymbol{a}$ is any reduced word for w.

In place of (2.10) we have

(2.16) *For any* $f \in P_n$,

$$\pi_{w_0} f = a_\delta^{-1} \sum_{w \in S_n} \epsilon(w) w(x^\delta f) = \partial_{w_0}(x^\delta f).$$

In particular, if $\alpha \in \mathbf{N}^n$,

$$(\mathbf{2.16'}) \qquad \pi_{w_0} x^\alpha = s_\alpha(x_1, \ldots, x_n).$$

Proof: We have

$$\pi_1 f = \partial_1(x_1 f),$$
$$\pi_1 \pi_2 f = \partial_1(x_1 \partial_2(x_2 f)) = \partial_1 \partial_2(x_1 x_2 f)$$

and generally

$$\pi_1 \cdots \pi_r f = \partial_1 \cdots \partial_r (x_1 \cdots x_r f)$$

for each $r \geq 1$. From this and (2.10) it follows easily that $\pi_{w_0} f = \partial_{w_0}(x^\delta f)$. ‖

Let $(a_1, \ldots, a_p)$ be a reduced word for w. Then

$$\begin{aligned} \partial_w &= \partial_{a_1} \cdots \partial_{a_p} \\ &= (x_{a_1} - x_{a_1+1})^{-1}(1 - s_{a_1})(x_{a_2} - x_{a_2+1})^{-1}(1 - s_{a_2}) \cdots \end{aligned}$$

which shows on expansion that ∂_w is of the form

$$\partial_w = \sum_{v \leq w} f_{vw} v$$

where f_{vw} are rational functions of $x_1, x_2, \ldots$, and in particular (by (1.7))

$$f_{ww} = (-1)^p \prod_{(i,j)\in I(w^{-1})} (x_i - x_j)^{-1}$$

and thus is $\neq 0$. It follows that the ∂_w are linearly independent over the field of rational functions $\mathbf{Q}_\infty = \mathbf{Q}(x_1, x_2, \ldots)$.

Now from (2.2) we have

$$\partial_a(fg) = (\partial_a f)g + (s_a f)(\partial_a g)$$

or equivalently, if $\mu : P_\infty \otimes P_\infty \to P_\infty$ is the multiplication map,

$$\partial_a \circ \mu = \mu \circ (\partial_a \otimes 1 + s_a \otimes \partial_a).$$

From this it follows that

$$\partial_w \circ \mu = \mu \circ (\partial_{a_1} \otimes 1 + s_{a_1} \otimes \partial_{a_1}) \circ \cdots \circ (\partial_{a_p} \otimes 1 + s_{a_p} \otimes \partial_{a_p})$$

On expansion this is a sum over subsequences $\boldsymbol{b}$ of $\boldsymbol{a} = (a_1, \ldots, a_p)$, say

(**1**) $$\partial_w \circ \mu = \mu \circ \sum_{\boldsymbol{b}\subset\boldsymbol{a}} \phi(\boldsymbol{a}, \boldsymbol{b}) \otimes \partial_{\boldsymbol{b}}$$

where

$$\phi(\boldsymbol{a}, \boldsymbol{b}) = \phi_1(\boldsymbol{a}, \boldsymbol{b}) \circ \cdots \circ \phi_p(\boldsymbol{a}, \boldsymbol{b})$$

and

$$\phi_i(\boldsymbol{a}, \boldsymbol{b}) = \begin{cases} s_{a_i} & \text{if } a_i \in \boldsymbol{b}, \\ \partial_{a_i} & \text{if } a_i \notin \boldsymbol{b}. \end{cases}$$

Since $\partial_{\boldsymbol{b}} = 0$ if $\boldsymbol{b}$ is not reduced (2.6), the sum is over *reduced* subsequences $\boldsymbol{b}$ of $\boldsymbol{a}$, and by (1.17) we can write

(**2**) $$\partial_w \circ \mu = \mu \circ \sum_{v \leq w} v\partial_{w/v} \otimes \partial_v$$

where for $v \leq w$

(**3**) $$\partial_{w/v} = v^{-1} \sum \phi(\boldsymbol{a}, \boldsymbol{b})$$

summed over subsequences $\boldsymbol{b} \subset \boldsymbol{a}$ such that $\boldsymbol{b}$ is a reduced word for v.

So for each pair of permutations w, v such that $w \geq v$ we have a well-defined operator $\partial_{w/v}$ on P_∞, defined by (3). Since the ∂_v are linearly independent, the definition (3) is *independent* of the reduced word $\boldsymbol{a} \in R(w)$.

(2.17) *For each pair* $w, v \in S_\infty$ *such that* $w \geq v$ *there is a linear operator* $\partial_{w/v}$ *on* P_∞ *such that*

$$\partial_w(fg) = \sum_{v \leq w} v(\partial_{w/v} f) \cdot \partial_v g.$$

$\partial_{w/v}$ *has degree* $-\ell(w) + \ell(v)$.

Examples.

1. Let $v = w$, then

$$\partial_{w/w} = w^{-1}\phi(\boldsymbol{a}, \boldsymbol{a}) = w^{-1} s_{a_1} \cdots s_{a_p} = 1.$$

2. Let $v = 1$, then

$$\partial_{w/1} = \phi(\boldsymbol{a}, \emptyset) = \partial_{a_1} \cdots \partial_{a_p} = \partial_w.$$

3. Suppose that $v \to w$, so that $v = s_{a_1} \cdots \hat{s}_{a_r} \cdots s_{a_p}$ for an unique $r \in [1, p]$. Then $\boldsymbol{b} = (a_1, \ldots, \hat{a}_r, \ldots, a_p)$ and

$$\begin{aligned}\partial_{w/v} &= v^{-1}\phi(\boldsymbol{a}, \boldsymbol{b}) \\ &= v^{-1} s_{a_1} \cdots s_{a_{r-1}} \partial_{a_r} s_{a_{r+1}} \cdots s_{a_p} \\ &= s_{a_p} \cdots s_{a_{r+1}} \partial_{a_r} s_{a_{r+1}} \cdots s_{a_p}\end{aligned}$$

.

Now $w = vt$ where t is the transposition

$$t = t_{ij} = s_{a_p} \cdots s_{a_r} \cdots s_{a_p} \qquad (i < j)$$

so that $(i, j) = s_{a_p} \cdots s_{a_{r+1}}(a_r, a_r + 1)$ and therefore

$$\begin{aligned}\partial_{w/v} &= s_{a_p} \cdots s_{a_{r+1}} (x_{a_r} - x_{a_r+1})^{-1}(1 - s_{a_r}) s_{a_{r+1}} \cdots s_{a_p} \\ &= (x_i - x_j)^{-1}(1 - t_{ij})\end{aligned}$$

is the divided difference operator ∂_{x_i, x_j}.

The product formula for $\partial_{w/u}$ is

(2.18)
$$\partial_{w/u}(fg) = \sum_{u \leq v \leq w} u^{-1} v(\partial_{w/v} f) \partial_{w/u} g.$$

Proof: We have

$$\partial_w(fgh) = \sum_{u \leq w} u \partial_{w/u}(fg) \partial_u h \tag{1}$$

and on the other hand

$$\partial_w(fgh) = \sum_{v \le w} v\partial_{w/u}(f)\partial_v(gh)$$
$$= \sum_{u \le v \le w} v\partial_{w/v}(f) \cdot u\partial_{v/u}(g) \cdot \partial_u h. \tag{2}$$

Comparison of (1) and (2) gives

$$u\partial_{w/u}(fg) = \sum_{u \le v \le w} v\partial_{w/u}(f) \cdot u\partial_{v/u}(g)$$

which gives the result. ‖

When $u = 1$, this reduces to (2.17).

Chapter III

Multi-Schur functions

For the time being we shall work in an arbitrary $\lambda-$ring R, but we shall use the notation of symmetric functions [M] rather than that of $\lambda-$rings. Thus for $X \in R$ we shall write $e_r(X)$ in place of $\lambda^r(X)$ for the r^{th} exterior power, and $h_r(X)$ in place of $\sigma^r(X)(= (-1)^r\lambda^r(-X))$ for the r^{th} symmetric power of X. We have $e_0(X) = h_0(X) = 1; e_1(X) = h_1(X) = X$; and $e_r(X) = h_r(X) = 0$ if $r < 0$.

Recall that if λ, μ are partitions and $X \in R$, the skew Schur function $s_{\lambda/\mu}(X)$ is defined by the formula

$$s_{\lambda/\mu}(X) = det(h_{\lambda_i-\mu_j-i+j}(X))_{1\leq i,j\leq n}$$

where $n \geq max(\ell(\lambda), \ell(\mu))$. It is zero unless $\lambda \supset \mu$.

We generalize this definition as follows : let $X_1, \ldots, X_n \in R$ and let λ, μ be partitions of length $\leq n$; then the *multi-Schur function* $s_{\lambda/\mu}(X_1, \ldots, X_n)$ is defined by

(3.1) $$s_{\lambda/\mu}(X_1, \ldots, X_n) = det(h_{\lambda_i-\mu_j-i+j}(X_i))_{1\leq i,j\leq n} \ .$$

We also define

(3.1′) $$s_\alpha(X_1, \ldots, X_n) = det(h_{\alpha_i-i+j}(X_i))_{1\leq i,j\leq n}$$

for any sequence $\alpha = (\alpha_1, \ldots, \alpha_n)$ of integers of length n.

Remark. In the definition (3.1) the argument X_i is constant in each *row* of the determinant. We might therefore also define

$$\hat{s}_{\lambda/\mu}(X_1, \ldots, X_n) = det(h_{\lambda_i-\mu_j-i+j}(X_j))_{1\leq i,j\leq n}$$

with arguments constant in each *column*. However, we get nothing essentially new: if we define partitions $\hat{\lambda}$, $\hat{\mu}$ by

$$\hat{\lambda}_i = N - \lambda_{n+1-i}, \quad \hat{\mu}_i = N - \mu_{n+1-i} \qquad (1 \leq i \leq n)$$

where $N \geq max(\lambda_1, \mu_1)$ (so that $\hat{\lambda}$ and $\hat{\mu}$ are the respective complements of λ and μ in the rectangle (N^n)), then we have

$$\hat{s}_{\lambda/\mu}(X_1, \ldots, X_n) = s_{\hat{\mu}/\hat{\lambda}}(X_n, ..., X_1)$$

as one sees by replacing (i, j) by $(n+1-j, n+1-i)$ in the determinant (3.1).

(3.2) *We have*

$$s_{\lambda/\mu}(X_1, \ldots, X_n) = 0$$

unless $\lambda \supset \mu$.

Proof: If $\lambda \not\supset \mu$ then $\lambda_r < \mu_r$ for some $r \leq n$, and hence

$$\lambda_i - \mu_j - i + j \leq \lambda_r - \mu_r < 0$$

whenever $i \geq r$ and $j \leq r$. It follows that the matrix $(h_{\lambda_i - \mu_j - i + j}(X_i))$ has an $(n - r + 1) \times r$ block of zeros in the south-west corner, and hence its determinant vanishes. ‖

(3.3) *If* $\lambda \supset \mu$ *and* $\ell(\lambda) = r < n$ *then*

$$s_{\lambda/\mu}(X_1, \ldots X_n) = s_{\lambda/\mu}(X_1, \ldots, X_r)$$

Proof: We have $\lambda_s = \mu_s = 0$ for $r + 1 \leq s \leq n$. Hence for each $s > r$ the s^{th} row of the matrix $(h_{\lambda_i - \mu_j - i + j}(X_i))$ has zeros in the first $s - 1$ places, and 1 in the s^{th} place. ‖

An element $X \in R$ is said to have *finite rank* if $e_n(X) = 0$ for all sufficiently large n. We then define the *rank* $rk(X)$ of X to be the largest r such that $e_r(X) \neq 0$. If X, Y both have finite rank, the formula

$$e_r(X + Y) = \sum_{p+q=r} e_p(X) e_q(Y)$$

shows that $X + Y$ has finite rank, and that

$$rk(X + Y) \leq rk(X) + rk(Y).$$

(3.4) *Let* $X_1, \ldots, X_n, Y_1, \ldots, Y_n \in R$ *with* $rk(Y_j) \leq j - 1$ $(1 \leq j \leq n)$ *(so that* $Y_1 = 0$*). Then for* all $\alpha \in \mathbf{Z}^n$,

$$s_\alpha(X_1, \ldots, X_n) = det(h_{\alpha_i - i + j}(X_i - Y_j)).$$

Proof: We have

$$h_{\alpha_i - i + j}(X_i - Y_j) = \sum_{k=1}^{j} h_{\alpha_i - i + k}(X_i) h_{j-k}(-Y_j),$$

since $h_r(-Y_j) = (-1)^r e_r(Y_j) = 0$ if $r \geq j$. Hence the matrix

$$(h_{\alpha_i - i + j}(X_i - Y_j))_{1 \leq i,j \leq n}$$

is the product of the matrix

$$(h_{\alpha_i - i + j}(X_i))_{1 \leq i,j \leq n}$$

and the matrix

$$(h_{i-j}(-Y_j))_{1 \leq i,j \leq n}$$

which is unitriangular. Now take determinants. $\|$

So far the X_i have been arbitrary elements of the λ-ring R. But it seems that $s_\alpha(X_1, \ldots, X_n)$ is mainly of interest when $X_1, \ldots, X_n$ is an *increasing* sequence in R, in the sense that $rk(X_{i+1} - X_i) < \infty$ for $1 \leq i \leq n-1$.

(3.5) *Let* x_i, y_i $(i \geq 1)$ *be elements of* R, *each of rank* ≤ 1, *and let*

$$X_i = x_1 + \cdots + x_i, \; Y_i = y_1 + \cdots + y_i$$

for each $i \geq 0$. *Then for all* $\alpha \in \mathbf{N}^n$ *we have*

$$s_\alpha(X_1 - Y_{\alpha_1}, \ldots, X_n - Y_{\alpha_n}) = \prod_{i=1}^{n} \prod_{j=1}^{\alpha_i} (x_i - y_j).$$

In particular, if λ *is a partition of length* $\leq n$,

$$s_\lambda(X_1 - Y_{\lambda_1}, \ldots, X_n - Y_{\lambda_n}) = \prod_{(i,j) \in \lambda} (x_i - y_j).$$

Proof: From (3.4) we have

$$(*) \qquad s_\alpha(X_1 - Y_{\alpha_1}, \ldots, X_n - Y_{\alpha_n}) = det(h_{\alpha_i - i + j}(X_i - Y_{\alpha_i} - X_{j-1}))_{1 \leq i,j \leq n} \; .$$

If $j > i$, then

$$h_{\alpha_i - i + j}(X_i - Y_{\alpha_i} - X_{j-1}) = \pm e_{\alpha_i - i + j}(Y_{\alpha_i} + X_{j-1} - X_i)$$

which is 0 because

$$rk(Y_{\alpha_i} + X_{j-1} - X_i) \leq \alpha_i + (j-1) - i < \alpha_i - i + j \; .$$

Hence the determinant at $(*)$ is triangular, with diagonal elements

$$\begin{aligned} h_{\alpha_i}(X_i - Y_{\alpha_i} - X_{i-1}) &= h_{\alpha_i}(x_i - Y_{\alpha_i}) \\ &= \sum_{r\geq 0} h_r(-Y_{\alpha_i}) h_{\alpha_i - r}(x_i) \\ &= \sum_{r\geq 0} (-1)^r e_r(Y_{\alpha_i}) x_i^{\alpha_i - r} \\ &= \prod_{j=1}^{\alpha_i} (x_i - y_j). \end{aligned}$$

The formula (3.5) now follows. ‖

In particular, when all the y_i are zero we have

(3.5′)
$$s_\alpha(X_1, \ldots, X_n) = \prod_{i=1}^{n} x_i^{\alpha_i} = x^\alpha$$

for all $\alpha \in \mathbf{N}^n$. Also, when all the x_i are zero (and α is a partition λ) we have

$$\begin{aligned} s_\lambda(-Y_{\lambda_1}, \ldots, -Y_{\lambda_n}) &= (-1)^{|\lambda|} \prod_{(i,j)\in\lambda} y_j \\ &= (-1)^{|\lambda|} y^{\lambda'}. \end{aligned}$$

If we replace the y's by x's, and λ by λ', this becomes

(3.5″)
$$x^\lambda = (-1)^{|\lambda|} s_{\lambda'}(-X_{\lambda'_1}, \ldots, -X_{\lambda'_n}).$$

(3.6) *Let $\lambda = (\lambda_1, \lambda_2, \ldots)$ be a partition of length $\leq n$, and $X_1, \ldots, X_n$ elements of a λ-ring R. Suppose that $i < j$ are such that $\lambda_i = \lambda_{i+1} = \cdots = \lambda_j$ and*

$$rk(X_j - X_k) \leq j - k \qquad \textit{for } i \leq k \leq j.$$

Then

$$s_{\lambda/\mu}(X_1, \ldots, X_n) = s_{\lambda/\mu}(X_1, \ldots, X_{i-1}, X_j, \ldots, X_j, X_{j+1}, \ldots, X_n),$$

that is to say we can replace each X_k $(i \leq k \leq j)$ by X_j without changing the value of the multi-Schur function.

Proof: Let $Y = X_j - X_i$, so that $rk(Y) \leq j - 1$. For all $m \geq 0$ we have

$$\begin{aligned} h_m(X_i) &= h_m(X_j - Y) \\ &= \sum_{k=0}^{j-i} (-1)^k e_k(Y) h_{m-k}(X_j). \end{aligned}$$

It follows that if we replace the i^{th} row of the determinant $s_{\lambda/\mu}(X_1, \ldots, X_{i-1}, X_j, \ldots, X_j, X_{j+1}, \ldots, X_n)$ by

$$\sum_{k=0}^{j-i} (-1)^k e_k(Y) row_{i+k}$$

we shall obtain

$$s_{\lambda/\mu}(X_1, \ldots, X_{i-1}, X_i, X_j, \ldots, X_j, X_{j+1}, \ldots, X_n)$$

with $j-i$ arguments equal to X_j. The proof is now completed by induction on $j-i$. ‖

Duality

Let $(X_n)_{n \in \mathbf{Z}}$ be a sequence in the λ-ring R such that

$$rk(X_n - X_{n-1}) \leq 1$$

for all $n \in \mathbf{Z}$.

(3.7) *Let I be any interval in $\mathbf{Z}$. Then the inverse of the matrix*

$$H = (h_{i-j}(-X_i))_{i,j \in I}$$

is $H^{-1} = (h_{i-j}(X_{j+1}))_{i,j \in I}$.

Proof: Let K denote the matrix $(h_{i-j}(X_{j+1}))$. The (i,k) element of HK is then

$$\sum_j h_{i-j}(-X_i) h_{j-k}(X_{k+1}) = h_{i-k}(-X_i + X_{k+1}).$$

If $i < k$ this is zero; if $i = k$ it is equal to 1; and if $i > k$ it is equal to

$$(-1)^{i-k} e_{i-k}(X_i - X_{k+1})$$

which is zero because $rk(X_i - X_{k+1}) \leq i - (k+1) < i - k$. ‖

(3.8) (Duality Theorem, 1st version) *Let $\lambda \supset \mu$ be partitions of length $\leq n$, such that $\ell(\lambda') \leq m$.* Then

$$s_{\lambda/\mu}(-X_{\lambda_1 - 1}, \ldots, -X_{\lambda_n - n}) = (-1)^{|\lambda - \mu|} s_{\lambda'/\mu'}(X_{1-\lambda'_1}, \ldots, X_{m-\lambda'_m}).$$

Proof: Let

$$\xi_i = \lambda_i - i, \quad \eta_i = \mu_i - i \qquad (1 \leq i \leq n),$$

$$\xi'_j = \lambda'_j - j, \quad \eta'_j = \mu'_j - j \qquad (1 \leq j \leq n),$$

Then the integers ξ_i $(1 \leq i \leq n)$ and $-\xi'_j - 1$ $(1 \leq j \leq m)$ fill up the interval $[-m, n-1]$, and so do the η_i and the $-\eta'_j - 1$.

The $(\xi_1, \ldots, \xi_n; \eta_1, \ldots, \eta_n)$ minor of the matrix H is

$$det(h_{\xi_i-\eta_j}(-X_{\xi_i})) = s_{\lambda/\mu}(-X_{\xi_i}, \ldots, -X_{\xi_n}).$$

The complementary cofactor of $(H^{-1})' = (h_{j-i}(X_{i+1}))_{-m \le i,j \le n-1}$ has row indices $-\xi_i' - 1$ $(1 \le i \le m)$ and column indices $-\eta_j' - 1$ $(1 \le j \le m)$. Hence it is

$$(-1)^{|\lambda-\mu|} s_{\lambda'/\mu'}(-X_{\xi_1'}, \ldots, -X_{\xi_m'}).$$

Since each minor of H is equal to the complementary cofactor of $(H^{-1})'$ (because det $H = 1$) the result follows. ‖

Remark. Observe that

$$rk(X_{\lambda_i - i} - X_{\lambda_{i+1}-i-1}) \le (\lambda_i - i) - (\lambda_{i+1} - i - 1) = \lambda_i - \lambda_{i+1} + 1.$$

Hence (3.8) gives a duality theorem for the multi-Schur function $s_{\lambda/\mu}(Y_1, \ldots, Y_n)$ provided that $rk(Y_{i+1} - Y_i) \le \lambda_i - \lambda_{i+1} + 1$ for $1 \le i \le n-1$.

At first sight the formula (3.8) is disconcerting, becuase the arguments $-X_{\lambda_i - i}$ on the left are not in general the negatives of the arguments $X_{i-\lambda_i'}$ on the right. However, we can use (3.6) to rewrite (3.8), as follows. As in Chapter I, let us write the partition λ in the form

$$\lambda = (p_1^{m_1}, p_2^{m_2}, \ldots, p_k^{m_k})$$

where $p_1 > p_2 > \ldots > p_k > 0$ and each $m_i \ge 1$. Then in

$$s_{\lambda/\mu}(-X_{\lambda_1-1}, \ldots, -X_{\lambda_n-n})$$

the first m_1 arguments are

$$-X_{p_1-1}, \ldots, -X_{p_1-m_1}$$

which by (3.6) may all be replaced by $-X_{c_1}$, where $c_1 = p_1 - m_1$. The next m_2 arguments are

$$-X_{p_2-m_1-1}, \ldots, -X_{p_2-m_1-m_2}$$

which by (3.6) may all be replaced by $-X_{c_2}$, where $c_2 = p_2 - m_1 - m_2$. In general, for each $i = 1, 2, \ldots, k$ the i^{th} group of m_i arguments may all be replaced by $-X_{c_i}$, where $c_i = p_i - (m_1 + \cdots + m_i)$. Now if

$$\lambda' = (q_1^{n_1}, q_2^{n_2}, \ldots, q_k^{n_k})$$

is the conjugate partition, we have $m_1 + \cdots + m_i = q_{k+1-i}$, and $c_i = p_i - q_{k+1-i}$ is the *content* of the square $s_i = (q_{k+1-i}, p_i)$ in the diagram of λ. The squares $s_1, \ldots, s_k$ are the "salients" of the border of λ, read in sequence from north-east to south-west. Hence the duality theorem (3.8) now takes the form

(3.8′) (Duality Theorem, 2nd version). *With the above notation, we have*

$$s_{\lambda/\mu}((-X_{c_1})^{m_1}, \ldots, (-X_{c_k})^{m_k}) = (-1)^{|\lambda-\mu|} s_{\lambda'/\mu'}((X_{c_k})^{n_1}, \ldots, (X_{c_1})^{n_k}) \ .$$

Finally, if we set $Z_i = -X_{c_i}$ $(1 \le i \le k)$ we have

$$s_{\lambda/\mu}(Z_1^{m_1}, \ldots, Z_k^{m_k}) = (-1)^{|\lambda-\mu|} s_{\lambda'/\mu'}((-Z_k)^{n_1}, \ldots, (-Z_1)^{n_k}) \tag{3.8''}$$

provided

$$rk(Z_{i+1} - Z_i) \le m_{i+1} + n_{k+1-i} \qquad (1 \le i \le k-1). \ \|$$

Let now $x_1, x_2, \ldots$ be independent indeterminates over $\mathbf{Z}$. We may regard $\mathbf{Z}[x_1, x_2, \ldots]$ as a λ-ring by requiring that each x_i has rank 1. Let $X_i = x_1 + \cdots + x_i$ for each $i \ge 1$. Then we have

$$\begin{aligned} \partial_i h_r(X_i) &= h_{r-1}(X_{i+1}), \\ \partial_i e_r(X_i) &= e_{r-1}(X_{i-1}), \\ \pi_i h_r(X_i) &= h_r(X_{i+1}). \end{aligned} \tag{3.9}$$

Proof: Consider the generating functions: $\partial_i h_r(X_i)$ is the coefficient of t^r in

$$\begin{aligned} \partial_i(\sum_{r\ge 0} h_r(X_i)t^r) &= \partial_i \prod_{j=1}^{i} (1 - x_j t)^{-1} \\ &= \prod_{j=1}^{i-1} (1 - x_j t)^{-1} \cdot \partial_i\left(\frac{1}{1 - x_i t}\right), \end{aligned}$$

and

$$\begin{aligned} \partial_i\left(\frac{1}{1 - x_i t}\right) &= \frac{1}{x_i - x_{i+1}}\left(\frac{1}{1 - x_i t} - \frac{1}{1 - x_{i+1} t}\right) \\ &= \frac{t}{(1 - x_i t)(1 - x_{i+1} t)} \end{aligned}$$

so that

$$\partial_i(\sum_r h_r(X_i)t^r) = t \prod_{j=1}^{i+1} (1 - x_j t)^{-1} = \sum_s h_s(X_{i+1}) t^{s+1}$$

in which the coefficient of t^r is $h_{r-1}(X_{i+1})$.

The other two relations are proved similarly. $\|$

(3.10) *Let $\alpha \in \mathbf{Z}^n$ and let $r_1, \ldots, r_n \geq 0$. If i is such that $r_i \neq r_j$ for all $j \neq i$ then*

$$\partial_{r_i} s_\alpha(X_{r_1}, \cdots, X_{r_n}) = s_{\alpha-\epsilon_i}(X_{r_1}, \ldots, X_{r_i+1}, \ldots, X_{r_n}),$$

$$\partial_{r_i} s_\alpha(-X_{r_1}, \cdots, -X_{r_n}) = -s_{\alpha-\epsilon_i}(-X_{r_1}, \ldots, -X_{r_i-1}, \ldots, -X_{r_n}),$$

$$\pi_{r_i} s_\alpha(X_{r_1}, \ldots, X_{r_n}) = s_\alpha(X_{r_1}, \ldots, X_{r_i+1}, \ldots, X_{r_n}),$$

where ϵ_i has ith coordinate equal to 1, and all other coordinates zero.

Proof: By definition, we have $s_\alpha = det(h_{\alpha_i-i+j}(X_{r_i}))$ and ∂_{r_i} acts only on the i^{th} row of the determinant, the entries in the other rows being symmetrical in x_{r_i} and x_{r_i+1} (because of the condition $r_j \neq r_i$ if $j \neq i$). Hence the first of the relations (3.10) follows from the first of the relations (3.9), and the other two are proved similarly. ‖

Remark. We can use the relations (3.10) to give another proof of duality (3.8) in the form

(3.8‴) *Let λ be a partition such that $\lambda_1 \leq m$ and $\lambda'_1 \leq n$. Then*

$$(*) \qquad s_\lambda(X_{m+1-\lambda_1}, \ldots, X_{m+n-\lambda_n}) = (-1)^{|\lambda|} s_{\lambda'}(-X_{m+\lambda'_1-1}, \ldots, -X_{\lambda'_m}).$$

Let (i,j) be a corner square of the diagram of λ, so that $j = \lambda_i$ and $i = \lambda'_j$. Let μ be the partition obtained from λ by removing the square (i,j). By operating on either side of $(*)$ with ∂_{m+i-j} we obtain the same relation with μ replacing λ. Hence it is enough to show that $(*)$ is true when $\lambda = (m^n)$, but in that case both sides are equal to $(X_1 \cdots X_m)^n$, by (3.5′), (3.5″) and (3.6). ‖

(3.11) *Let w_0 be the longest element of S_n. Then for any $\alpha \in \mathbf{Z}^n$ we have*

$$\partial_{w_0} s_\alpha(X_1 + Z_1, \ldots, X_n + Z_n) = s_{\alpha-\delta}(X_n + Z_1, \ldots, X_n + Z_n),$$

$$\pi_{w_0} s_\alpha(X_1 + Z_1, \ldots, X_n + Z_n) = s_\alpha(X_n + Z_1, \ldots, X_n + Z_n),$$

where $X_i = x_1 + \cdots + x_i$ $(1 \leq i \leq n)$ and the Z_i are independent of $x_1, \ldots, x_n$.

Proof: The sequence

$$(n-1, n-2, n-1, \ldots, 2, 3, \ldots, n-1, 1, 2, 3, \ldots, n-1)$$

is a reduced word for w_o, so that

$$\pi_{w_0} = \pi_{n-1}(\pi_{n-2}\pi_{n-1}) \cdots (\pi_2\pi_3 \cdots \pi_{n-1})(\pi_1\pi_2 \cdots \pi_{n-1})$$

and likewise for ∂_{w_o}. By (3.10), $\pi_1\pi_2 \ldots \pi_{n-1}$ applied to $s_\alpha(X_1 + Z_1, \ldots, X_n + Z_n)$ will produce

$$s_\alpha(X_2 + Z_1, X_3 + Z_2, \ldots, X_n + Z_{n-1}, X_n + Z_n).$$

We have next to operate on this with $\pi_2\pi_3\cdots\pi_{n-1}$, which will produce

$$s_\alpha(X_3+Z_1, X_4+Z_2, \ldots, X_n+Z_{n-2}, X_n+Z_{n-1}, X_n+Z_n).$$

By repeating this process we shall obtain the formula for $\pi_{w_0}s_\alpha$. That for $\partial_{w_0}s_\alpha$ is proved similarly. ‖

Remark. Let $\alpha \in \mathbf{N}^n$ and $Z_1 = \cdots = Z_n = 0$ in (3.11). Then by (3.5′) we have

$$\partial_{w_0}(x^\alpha) = \partial_{w_0}s_\alpha(X_1,\ldots,X_n) = s_{\alpha-\delta}(X_n),$$

$$\pi_{w_0}(x^\alpha) = \pi_{w_0}s_\alpha(X_1,\ldots,X_n) = s_\alpha(X_n).$$

Thus we have independent proofs of (2.11) and (2.16′) and hence (by linearity) of (2.10) and (2.16).

Sergeev's formula

Let $x_1,\ldots,x_m,y_1,\ldots,y_n$ be independent variables and let

$$X_i = x_1+\cdots+x_i,\ Y_i = y_1+\cdots+y_i$$

for all $i \geq 1$, with the understanding that $x_j = 0$ if $j > m$ and $y_j = 0$ if $j > n$.

(3.12) (Sergeev) *For all partitions* λ *we have*

$$s_\lambda(X_m - Y_n) = \sum_{w\in S_m\times S_n} w(f_\lambda(x,y)/D(x)D(y))$$

where

$$f_\lambda(x,y) = \prod_{(i,j)\in\lambda}(x_i - y_j),$$

$$D(x) = \prod_{1\leq i<j\leq m}(1-x_i^{-1}x_j), \quad D(y) = \prod_{1\leq i<j\leq n}(1-y_i^{-1}y_j).$$

Proof: Let $w_0^{(m)}$ (resp. $w_0^{(n)}$) be the longest element of S_m (resp. S_n) and let π_x (resp. π_y) denote $\pi_{w_o^{(m)}}$ acting on the x's (resp. $\pi_{w_0^{(n)}}$ acting on the y's). From (3.5) we have, if $r = \ell(\lambda)$,

(1) $$f_\lambda(x,y) = s_\lambda(X_1 - Y_{\lambda_1}, \ldots, X_r - Y_{\lambda_r})$$

and in view of (2.16) Sergeev's formula may be restated in the form

(3.12′) $$s_\lambda(X_m - Y_n) = \pi_y\pi_x f_\lambda(x,y).$$

From (3.11) and (1) above we have

(2) $$\pi_x f_\lambda(x,y) = s_\lambda(X_m - Y_{\lambda_1}, \ldots, X_m - Y_{\lambda_r}).$$

If $\lambda = (p_1^{m_1}, \dots, p_k^{m_k})$, (2) can be rewritten in the form

$$\pi_x f_\lambda(x, y) = s_\lambda(Z_1^{m_1}, \dots, Z_k^{m_k}) \tag{3}$$

where $Z_i = X_m - Y_{p_i}$. Since

$$rk(Z_{i+1} - Z_i) = rk(Y_{p_i} - Y_{p_{i+1}}) = p_i - p_{i+1},$$

the duality theorem (3.8″) applies, and gives

$$\begin{aligned} s_\lambda(Z_1^{m_1}, \dots, Z_k^{m_k}) &= (-1)^{|\lambda|} s_{\lambda'}((-Z_k)^{n_1}, \dots, (-Z_1)^{n_k}) \\ &= (-1)^{|\lambda|} s_{\lambda'}((Y_{p_k} - X_m)^{n_1}, \dots, (Y_{p_1} - X_m)^{n_k}) \\ &= (-1)^{|\lambda|} s_{\lambda'}(Y_1 - X_m, Y_2 - X_m, \dots, Y_s - X_m) \qquad (4) \end{aligned}$$

where $s = n_1 + \cdots + n_k = \ell(\lambda')$. We can now apply (3.11) again and obtain from (3) and (4)

$$\begin{aligned} \pi_y \pi_x f_\lambda &= (-1)^{|\lambda|} \pi_y s_{\lambda'}(Y_1 - X_m, \dots, Y_s - X_m) \\ &= (-1)^{|\lambda|} s_{\lambda'}(Y_n - X_m) \\ &= s_\lambda(X_m - Y_n). \ \| \end{aligned}$$

Chapter IV

Schubert Polynomials (1)

Let $\delta = \delta_n = (n-1, n-2, \ldots, 1, 0)$, so that

$$x^\delta = x_1^{n-1} x_2^{n-2} \cdots x_{n-1}.$$

For each permutation $w \in S_n$ the *Schubert polynomial* $\mathfrak{S}_w$ is defined to be

$$\mathfrak{S}_w = \partial_{w^{-1}w_0}(x^\delta) \tag{4.1}$$

where as usual w_0 is the longest element of S_n.

(4.2) *Let* $v, w \in S_n$. *Then*

$$\partial_v \mathfrak{S}_w = \begin{cases} \mathfrak{S}_{wv^{-1}} & \text{if } \ell(wv^{-1}) = \ell(w) - \ell(v), \\ 0 & \textit{otherwise.} \end{cases}$$

In particular,

$$\partial_i \mathfrak{S}_w = \begin{cases} \mathfrak{S}_{ws_i} & \text{if } w(i) > w(i+1), \\ 0 & \text{if } w(i) < w(i+1). \end{cases}$$

Proof: From (2.7) we have

$$\partial_v \partial_{w^{-1}w_0} = \begin{cases} \partial_{vw^{-1}w_0} & \text{if } \ell(v) + \ell(w^{-1}w_0) = \ell(vw^{-1}w_0), \\ 0 & \text{otherwise.} \end{cases}$$

Now

$$\ell(v) + \ell(w^{-1}w_0) = \ell(v) + \ell(w_0) - \ell(w)$$

and

$$\ell(vw^{-1}w_0) = \ell(w_0) - \ell(wv^{-1})$$

by (1.6). Hence $\partial_v \mathfrak{S}_w = \partial_v \partial_{w^{-1}w_0} x^\delta$ is equal to $\partial_{vw^{-1}w_0} x^\delta = \mathfrak{S}_{wv^{-1}}$ if $\ell(w) - \ell(v)$, and is zero otherwise. ‖

(4.3) (i) $\mathfrak{S}_{w_0} = x^\delta, \quad \mathfrak{S}_1 = 1.$

(ii) *For each* $w \in S_n$, $\mathfrak{S}_w$ *is a non-zero homogeneous polynomial in* $x_1, \ldots, x_{n-1}$ *of degree* $\ell(w)$, *of the form*

$$\mathfrak{S}_w = \sum_\alpha c_\alpha x^\alpha$$

summed over $\alpha \in \mathbf{N}^{n-1}$ *such that* $\alpha \subset \delta$ (i.e., $\alpha_i \le n - i$ *for each* i) *and* $|\alpha| = \ell(w)$.

(iii) $\mathfrak{S}_w$ *is symmetrical in* x_i, x_{i+1} *if and only if* $w(i) < w(i+1)$.

(iv) *If* r *is the last descent of* $w \in S_n$ (i.e., *if* $w(r) > w(r+1)$ *and* $w(r+1) < w(r+2) < \cdots < w(n)$), *then* $\mathfrak{S}_w \in P_r = \mathbf{Z}[x_1, \ldots, x_r]$, *and* $\mathfrak{S}_w \notin P_{r-1}$.

Proof: (i) That $\mathfrak{S}_{w_0} = x^\delta$ is clear from the definition (4.1). Also by (2.11) we have

$$\mathfrak{S}_1 = \partial_{w_0} x^\delta = s_{\delta-\delta} = 1.$$

(ii) The operator $\partial_{w^{-1}w_0}$ lowers degrees by $\ell(w^{-1}w_0) = \ell(w_0) - \ell(w^{-1}) = \frac{1}{2}n(n-1) - \ell(w)$. Hence $\mathfrak{S}_w = \partial_{w^{-1}w_0} x^\delta$ is homogeneous of degree $\ell(w)$. If now $\alpha \in \mathbf{N}^{n-1}$ is such that $\alpha \subset \delta$, then by (2.1) $\partial_r x^\alpha$ is a linear combination of monomials x^β such that $\beta_i = \alpha_i$ if $i \ne r, r+1$, and

$$max(\beta_i, \beta_{i+1}) \le max(\alpha_i, \alpha_{i+1}) - 1 \le n - i - 1,$$

so that $\beta \subset \delta$. Hence the linear span H_n of the monomials $x^\alpha, \alpha \subset \delta$ is mapped into itself by each ∂_r $(1 \le r \le n-1)$ and hence by each ∂_w, $w \in S_n$. Hence $\mathfrak{S}_w \in H_n$ for each $w \in S_n$.

(iii) $\mathfrak{S}_w$ is symmetrical in x_i and x_{i+1} if and only if $s_i \mathfrak{S}_w = \mathfrak{S}_w$, that is to say if and only if $\partial_i \mathfrak{S}_w = 0$, which by (4.2) is equivalent to $w(i) < w(i+1)$.

(iv) $\mathfrak{S}_w$ is symmetrical in $x_{r+1}, \ldots, x_n$ by (iii) above, but does not contain x_n, hence does not contain any of $x_{r+1}, \ldots, x_n$. $\|$

Remark. We shall show later (4.17) that the coefficients in (4.3)(ii) are always non-negative integers.

(4.4) *For* $i = 1, 2, \ldots, n-1$ *we have*

$$\mathfrak{S}_{s_i} = x_1 + x_2 + \cdots + x_i.$$

Proof: By (4.3), $\mathfrak{S}_{s_i}$ is a homogeneous symmetric polynomial of degree $\ell(s_i) = 1$ in $x_1, \ldots, x_i$, hence is equal to $c(x_1 + \cdots + x_i)$ for some integer c. But $\partial_i \mathfrak{S}_{s_i} = \mathfrak{S}_1 = 1$ by (4.2) and (4.3)(i), hence $c = 1$. $\|$

(4.5) (Stability) *Let* $m > n$ *and let* $i : S_n \hookrightarrow S_m$ *be the embedding. Then*

$$\mathfrak{S}_w = \mathfrak{S}_{i(w)}$$

for all $w \in S_n$.

Proof: We may assume that $m = n+1$. Let w_0' be the longest element of S_{n+1}, then $w_0' = w_0 s_n s_{n-1} \cdots s_1$, where w_0 is the longest element of S_n, and hence

$$\begin{aligned}\mathfrak{S}_{i(w)} &= \partial_{w^{-1}w_0'}(x_1^n x_2^{n-1} \cdots x_n) \\ &= \partial_{w^{-1}w_0}\partial_n \partial_{n-1} \cdots \partial_1 (x_1^n x_2^{n-1} \cdots x_n) \\ &= \partial_{w^{-1}w_0}(x_1^{n-1} x_2^{n-2} \cdots x_{n-1})\end{aligned}$$

(because $\partial_1 (x_1^n x_2^{n-1} \cdots x_n) = x_1^{n-1} x_2^{n-1} x_3^{n-2} \cdots x_n$, hence $\partial_2 \partial_1 (x_1^n x_2^{n-2} \cdots x_n) = x_1^{n-1} x_2^{n-2} x_3^{n-2} x_4^{n-3} \cdots x_n$, and so on.) $\|$

From (4.5) it follows that $\mathfrak{S}_w$ is a well-defined polynomial for each permutation $w \in S_\infty = \bigcup_n S_n$.

If $u \in S_m$ and $v \in S_n$, we denote by $u \times v$ the permutation

$$u \times v = (u(1), \ldots, u(m), v(1)+m, \ldots, v(n)+m)$$

in S_{m+n}. We have then

(4.6) $$\mathfrak{S}_{u\times v} = \mathfrak{S}_u \cdot \mathfrak{S}_{1_m \times v}$$

where 1_m is the identity element of S_m.

Proof: We shall make use of the following fact: if f is a polynomial in $x_1, x_2, \ldots$, and $\partial_i f = 0$ for all $i \geq 1$, then f is a constant. For $f \in P_n = \mathbf{Z}[x_1, \ldots, x_n]$ for some n, and is symmetric in $x_1, \ldots, x_{n+1}$ because $\partial_1 f = \cdots = \partial_n f = 0$.

To prove (4.6) we proceed by induction on $\ell(u) + \ell(v)$. If $\ell(u) = \ell(v) = 0$ then $u = 1_m,\ v = 1_n$, and both sides of (4.6) are equal to 1. Let

$$F(u,v) = \mathfrak{S}_{u\times v} - \mathfrak{S}_u \mathfrak{S}_{1_m \times v}.$$

By the remark above, it is enough to show that $\partial_i F(u,v) = 0$ for each i.

Suppose first that $i < m$. Then

$$\partial_i F(u,v) = \partial_i(\mathfrak{S}_{u\times v}) - \partial_i(\mathfrak{S}_u) \cdot \mathfrak{S}_{1_m \times v}$$

because $\partial_i(\mathfrak{S}_{1_m\times v}) = 0$ by (4.2). Hence we have $\partial_i F(u,v) = 0$ if $\ell(us_i) > \ell(u)$; and if $\ell(us_i) < \ell(u)$ then

$$\partial_i F(u,v) = F(us_i, v)$$

which is zero by the inductive hypothesis.

Likewise, if $i > m$ we have

$$\partial_i F(u,v) = \begin{cases} F(u, vs_i) & \text{if } \ell(vs_i) < \ell(v), \\ 0 & \text{otherwise,} \end{cases}$$

and so again $\partial_i F(u,v) = 0$ by the inductive hypothesis.

Finally, if $i = m$ we have $\ell((u \times v)s_m) > \ell(u \times v)$, because

$$(u \times v)(m) = u(m) < m + v(1) = (u \times v)(m+1),$$

and therefore ∂_m kills $\mathfrak{S}_{u\times v}$ and $\mathfrak{S}_{1_m \times v}$; moreover, $\partial_m \mathfrak{S}_u = 0$, because $\mathfrak{S}_u \in \mathbf{Z}[x_1, \dots, x_{m-1}]$. Hence $\partial_m F(u,v) = 0$, and the proof is complete. ‖

For certain classes of premutations there are explicit formulas for $\mathfrak{S}_w$. We consider first the case where w is dominant, of shape λ (so that the diagram of w coincides with the diagram of λ).

(4.7) *If w is dominant of shape λ, then*

$$\mathfrak{S}_w = x^\lambda.$$

Proof: We use descending induction on $\ell(w)$, where $w \in S_n$. The result is true for $w = w_0$ by (4.3)(i), since w_0 is dominant of shape δ.

Suppose $w \in S_n$, $w \neq w_0$ and w is dominant of shape λ. Then $\lambda \subset \delta$ and $\lambda \neq \delta$. Let $r \geq 0$ be the largest integer such that $\lambda'_i = n - i$ for $1 \leq i \leq r$, and let $a = \lambda'_{r+1} + 1 \leq n - r - 1$. Then ws_a is dominant of length $\ell(w) + 1$, and $\lambda(ws_a) = \lambda + \epsilon_a$, where ϵ_a is the vector whose a^{th} component is 1 and all other components zero. Hence we have

$$\mathfrak{S}_w = \partial_a \mathfrak{S}_{ws_a} = \partial_a (x_a x^\lambda) = x^\lambda,$$

because $\lambda_a = \lambda_{a+1}$. ‖

Conversely, every monomial x^λ (where λ is a partition) occurs as a Schubert polynomial, namely as $\mathfrak{S}_w$ where w is the permutation with code $c(w) = \lambda$.

Suppose next that w is Grassmannian, with descent at r.

(4.8) *If w is Grassmannian of shape λ,* then *$\mathfrak{S}_w$ is the Schur function $s_\lambda(X_r)$, where r is the unique descent of w, and $X_r = x_1 + \cdots + x_r$.*

Proof: We may assume that $w \neq 1$ (by (4.3)(i), $\mathfrak{S}_1 = 1$). Then $r \geq 1$ and the code of w is

$$(w(1) - 1, w(2) - 2, \dots, w(r) - r)$$

so that $\lambda = (w(r) - r, \dots, w(2) - 2, w(1) - 1)$. Let $u = w_0^{(r)}$ be the longest element of S_r. Then

$$wu = (w(r), \dots, w(1), w(r+1), w(r+2)\dots)$$

is dominant of shape $\lambda + \delta_r$, where $\delta_r = (r-1, r-2, \dots, 1, 0)$, and $\ell(wu) = \ell(w) + \ell(u)$. Hence

$$\mathfrak{S}_w = \partial_u \mathfrak{S}_{wu} = \partial_u(x^{\lambda+\delta_r}) = s_\lambda(X_r)$$

by (4.2), (4.7) and (2.11). ‖

Conversely, every Schur function $s_\lambda(X_r)$ (where λ is a partition of length $\le r$) occurs as a Schubert polynomial, namely as $\mathfrak{S}_w$ where w is the permutation with code $c(w) = (\lambda_r, \lambda_{r-1}, \dots, \lambda_1)$.

More generally, let w be vexillary with shape $\lambda = (\lambda_1, \dots, \lambda_m)$ (where $m = \ell(\lambda)$) and flag $\phi = (\phi_1, \dots, \phi_m)$ (Chapter I). Then $\mathfrak{S}_w$ is a multi-Schur function (Chapter III), namely

(4.9) $$\mathfrak{S}_w = s_\lambda(X_{\phi_1}, \dots, X_{\phi_m})$$

where $X_i = x_1 + \cdots + x_i$ *for each* $i \ge 1$.

Proof: The idea is to convert w systematically into a dominant permutation. Recall ((1.23), (1.24)) that if $c(w) = (c_1, c_2, \dots)$ and $c_i \le c_{i+1}$ for some $i \ge 1$, then $\ell(ws_i) = \ell(w) + 1$ and

$(*)$ $$c(ws_i) = (c_1, \dots, c_{i-1}, c_{i+1} + 1, c_i, c_{i+2}, c_{i+3}, \dots).$$

As in Chapter I let

$$\lambda(w) = (p_1^{m_1}, \dots, p_k^{m_k})$$

where $p_1 > \cdots > p_k > 0$ (and each $m_i \ge 1$), and let

$$\phi(w) = (f_1^{m_1}, \dots, f_k^{m_k})$$

where $f_1 \le \dots \le f_k$.

Consider first the terms equal to p_1 in the sequence $c(w)$. They occupy the positions $f_1 - m_1 + 1$, $\dots, f_1$. We shall use $(*)$ to move them all to the left until they occupy the first m_1 positions, by multiplying w on the right by

$$u_1 = (s_{f_1-m_1} \cdots s_2 s_1)(s_{f_1-m_1+1} \cdots s_3 s_2) \cdots \; (s_{f_1-1} \cdots s_{m_1+1} s_{m_1}).$$

Let $w_1 = wu_1$. In the code of w_1, the first m_1 entries will be equal to $p_1 + f_1 - m_1$; the shape of w_1 is

$$\lambda^{(1)} = \lambda(w_1) = ((p_1 + f_1 - m_1)^{m_1}, p_2^{m_2}, \dots, p_k^{m_k}),$$

and it follows from the description (1.38) of vexillary codes that the terms equal to p_2 in the sequence $c(w_1)$ will occupy the positions $f_2 - m_2 + 1, \ldots, f_2$. The next step is to move those to the left until they occupy the positions $m_1 + 1, \ldots, m_1 + m_2$ by multiplying w_1 on the right by

$$u_2 = (s_{f_2-m_2} \cdots s_{m_1+2} s_{m_1+1})(s_{f_2-m_2+1} \cdots s_{m_1+2}) \cdots (s_{f_2-1} \cdots s_{m_1+m_2}).$$

Let $w_2 = w_1 u_2$; the code of w_2 starts off with m_1 entries to $p_1 + f_1 - m_1$, then m_2 entries equal to $p_2 + f_2 - m_1 - m_2$; the shape of w_2 is

$$\lambda^{(2)} = \lambda(w_2) = ((p_1 + f_1 - m_1)^{m_1}, (p_1 + f_2 - m_1 - m_2)^{m_2}, p_3^{m_3}, \ldots, p_k^{m_k}),$$

and the terms equal to p_3 in the sequence $c(w_2)$ will occupy the positions $f_3 - m_3 + 1, \ldots, f_3 - m_3$.

We continue in this way; at the r^{th} stage we define $w_r = w_{r-1} u_r$, where

$$u_r = (s_{f_r - m_r} \cdots s_{m_1 + \cdots + m_{r-1} + 1}) \cdots (s_{f_r - 1} \cdots s_{m_1 + \cdots + m_r}),$$

and w_r has shape

$$\lambda^{(r)} = \lambda(w_r) = ((p_1 + a_1)^{m_1}, \ldots, (p_r + a_r)^{m_r}, p_{r+1}^{m_r+1}, \ldots, p_k^{m_k})$$

where $a_i = f_i - (m_1 + \cdots + m_i) \geq 0$ by (1.36). Notice also that

$$(p_{i-1} + a_{i-1}) - (p_i + a_i) = (m_i + p_{i-1} - p_i) - (f_i - f_{i-1}) \geq 0$$

by (1.37).

Finally we reach $w_k = w u_1 \cdots u_k$, which is dominant with shape (and code)

$$\mu = \lambda^{(k)} = ((p_1 + a_1)^{m_1}, \ldots, (p_k + a_k)^{m_k}).$$

We have

$$\ell(w) = |\lambda| = \sum m_i p_i,$$
$$\ell(w_k) = |\lambda^{(k)}| = \sum m_i (p_i + a_i),$$

and

$$\ell(u_r) = a_r m_r \qquad (1 \leq r \leq k)$$

so that

$$\ell(w_k) = \ell(w) + \sum_{r=1}^{k} \ell(u_r)$$

and therefore, since $w = w_k (u_1 \ldots u_k)^{-1}$,

$$\mathfrak{S}_w = \partial_{u_1} \cdots \partial_{u_k} \mathfrak{S}_{w_k}$$

by (4.2). Now by (4.6) and (3.5′) we have

$$\mathfrak{S}_{w_k} = x^\mu = s_\mu(X_1, \ldots, X_m)$$

where $m = m_1 + \cdots + m_k = \ell(\lambda)$. Hence by repeated use of (3.10) we obtain

$$\begin{aligned}\mathfrak{S}_{w_{k-1}} &= \partial_{u_k}\mathfrak{S}_{w_k} \\ &= s_{\lambda^{(k-1)}}(X_1, \ldots, X_{m_1+\cdots+m_{k-1}}, X_{f_k-m_k+1}, \ldots, X_{f_k-1}, X_{f_k}) \\ &= s_{\lambda^{(k-1)}}(X_1, \ldots, X_{m_1+\cdots+m_{k-1}}, (X_{f_k})^{m_k})\end{aligned}$$

by virtue of (3.6). If we now operate with $\partial_{u_{k-1}}$ we shall obtain in the same way

$$\mathfrak{S}_{w_{k-1}} = \partial_{u_{k-1}}\mathfrak{S}_{w_{k-1}} = s_{\lambda^{(k-2)}}(X_1, \ldots, X_{m_1+\cdots+m_{k-2}}, (X_{f_{k-1}})^{m_{k-1}}, (X_{f_k})^{m_k})$$

and so finally

$$\mathfrak{S}_w = s_\lambda((X_{f_1})^{m_1}, \ldots, (X_{f_k})^{m_k}). \ \|$$

Remarks. 1. As in Chapter I, let

$$\lambda' = (q_1^{n_1}, \ldots, q_k^{n_k})$$

be the conjugate partition, so that

$$m_1 + \cdots + m_i = q_{k+1-i} \qquad (1 \le i \le k)$$

and therefore

$$\begin{aligned}p_i + a_i &= p_i + f_i - q_{k+1-i} \\ &= g_{k+1-i}\end{aligned}$$

by (1.41), where $(g_1^{n_1}, \ldots, g_k^{n_k})$ is the flag of w^{-1}. Thus

(4.10) $$\mu = \lambda^{(k)} = (g_k^{m_1}, g_{k-1}^{m_2}, \ldots, g_1^{m_k}).$$

2. The result (4.9) admits a converse. If $\lambda = (p_1^{m_1}, \ldots, p_k^{m_k})$ as above, every non-zero multi-Schur function $s_\lambda((X_{f_1})^{m_1}, \ldots, (X_{f_k})^{m_k})$ that satisfies the conditions of the duality theorem (3.8″), namely

(1) $$0 \le f_{i+1} - f_i \le m_{i+1} + n_{k+1-i} \qquad (1 \le i \le k-1),$$

is the Schubert polynomial of a vexillary permutation, namely the permutation with shape λ and flag $\phi = (f_1^{m_1}, \ldots, f_k^{m_k})$. This follows from (1.38) and (4.9), since the conditions (1) on the flag ϕ coincide with those of (1.37). (The conditions (1.36), namely

$$f_i \ge m_1 + \cdots + m_i \qquad (1 \le i \le k)$$

ensure that the multi-Schur function does not vanish indentically.)

Let H_n denote the additive subgroup of $P_n = \mathbf{Z}[x_1, \ldots, x_n]$ spanned by the monomials $x^\alpha, \alpha \subset \delta_n = (n-1, n-2, \ldots, 1, 0)$.

(4.11) *The Schubert polynomials* $\mathfrak{S}_w, w \in S_n$ *form a* $\mathbf{Z}$*-basis of* H_n.

Proof: By (4.3) each $\mathfrak{S}_w$ lies in H_n. If

$$\sum a_w \mathfrak{S}_w = 0 \qquad (a_w \in \mathbf{Z})$$

is a linear dependence relation, then by homogeneity we have

$$(1) \qquad \sum_{\ell(w)=p} a_w \mathfrak{S}_w = 0$$

for each $p \geq 0$, and by operating on (1) with ∂_w we see that $a_w = 0$. Hence the $\mathfrak{S}_w$ are linearly independent and hence form a $\mathbf{Q}$-basis of $H_n \otimes \mathbf{Q}$. It follows that each monomial $x^\alpha, \alpha \subset \delta_n$, can be expressed in the form

$$(2) \qquad x^\alpha = \sum_{\ell(w)=|\alpha|} b_w \mathfrak{S}_w$$

with rational coefficients b_w; by operating on (2) with ∂_w we have $b_w = \partial_w x^\alpha$, and hence the b_w are integers. ‖

From (4.11) it follows that

(4.12) *The* $\mathfrak{S}_w, w \in S_\infty$, *form a* $\mathbf{Z}$*-basis of* $P_\infty = \mathbf{Z}[x_1, x_2, \ldots]$.

Proof: Let x^α be a monomial in P_∞. Then $\alpha \subset \delta_n$ for sufficiently large n, hence x^α is a linear combination of the $\mathfrak{S}_w$. ‖

For each $n \geq 1$, let $S^{(n)}$ denote the set of all permutations w such that $w(n+1) < w(n+2) < \cdots$, or equivalently such that the code of w has length $\leq n$.

(4.13) *The* $\mathfrak{S}_w, w \in S^{(n)}$, *form a* $\mathbf{Z}$*-basis of* P_n.

Proof: By (4.3)(iii) we have

$$\begin{aligned} \mathfrak{S}_w \in P_n &\iff \partial_m \mathfrak{S}_w = 0 \text{ for all } m > n \\ &\iff w \in S^{(n)}. \end{aligned}$$

Let $P_n' \subset P_n$ be the $\mathbf{Z}$-span of the $\mathfrak{S}_w, w \in S^{(n)}$. If $P_n' \neq P_n$, choose $f \in P_n - P_n'$; by (4.12) we can write f as a linear combination of Schubert polynomials, say

$$(1) \qquad f = \sum_w a_w \mathfrak{S}_w$$

where there is at least one term with $a_w \neq 0$ and $w \notin S^{(n)}$. Hence for some $m > n$ we have $\partial_m \mathfrak{S}_w = \mathfrak{S}_{ws_m}$, and since $\partial_m f = 0$ we obtain from (1) a nontrivial linear dependence relation among the Schubert polynomials, contradicting (4.12). Hence $P_n' = P_n$, which proves (4.13). ‖

Let $\eta : P_n \to \mathbf{Z}$ be the homomorphism defined by $\eta(x_i) = 0$ $(1 \leq i \leq n)$. In other words, $\eta(f)$ is the constant term of f, for each polynomial $f \in P_n$. The expression of f in terms of Schubert polynomials is then

(4.14)
$$f = \sum_{w \in S^{(n)}} \eta(\partial_w f) \mathfrak{S}_w.$$

Proof: By (4.13) and linearity, it is only necessary to verify this formula when f is a Schubert polynomial $\mathfrak{S}_v, v \in S^{(n)}$, and then it follows from (4.2) and (4.3)(ii) that $\eta(\partial_w \mathfrak{S}_v)$ is equal to 1 when $w = v$ and is zero otherwise. ‖

(4.15) *Let* $f = \sum \alpha_i x_i$ *be a homogeneous linear polynomial, and let* w *be a permutation. Then*

$$f \mathfrak{S}_w = \sum (\alpha_i - \alpha_j) \mathfrak{S}_{wt_{ij}},$$

where t_{ij} *is the transposition that interchanges* i *and* j, *and the sum is over all pairs* $i < j$ *such that* $\ell(wt_{ij}) = \ell(w) + 1$.

Proof: The polynomial $f\mathfrak{S}_w$ is homogeneous of degree $\ell(w) + 1$, and hence by (4.14) we have

$$f \mathfrak{S}_w = \sum_v \partial_v (f \mathfrak{S}_w) \cdot \mathfrak{S}_v$$

summed over v of length $\ell(w) + 1$. Now by (2.13)

$$\partial_v (f \mathfrak{S}_w) = v(f) \partial_v \mathfrak{S}_w + \sum (\alpha_i - \alpha_j) \partial_{vt_{ij}} \mathfrak{S}_w$$

summed over $i < j$ such that $\ell(vt_{ij}) = \ell(v) - 1 = \ell(w)$. It follows that $\partial_v(f\mathfrak{S}_w) = \alpha_i - \alpha_j$ if $w = vt_{ij}$, and is zero otherwise. ‖

In particular:

(4.15′)
$$x_r \mathfrak{S}_w = \sum \sigma(t) \mathfrak{S}_{wt}$$

summed over transpositions $t = t_{ir}$ *such that* $\ell(wt) = \ell(w) + 1$, *where* $\sigma(t) = -1$ or $+1$ *according as* $i < r$ *or* $i > r$. ‖

(4.15″) (Monk's formula) $\mathfrak{S}_{s_r} \mathfrak{S}_w = \sum \mathfrak{S}_{wt}$ *summed over transpositions* $t = t_{ij}$ *such that* $i \leq r < j$ *and* $\ell(wt) = \ell(w) + 1$. ‖

Remark. As pointed out by A. Lascoux, Monk's formula (4.15$''$) (which is the counterpart of Pieri's formula in the theory of Schur functions) characterizes the algebra of Schubert polynomials.

We shall apply (4.15$'$) in the following situation. Suppose that r is the last descent of w, so that $w(r) > w(r+1)$ and $w(r+1) < w(r+2) < \cdots$. Choose the largest $s > r$ such that $w(r) > w(s)$ and let $v = wt_{rs}$. Then from (4.15$'$) applied to v we have

$$(1) \qquad x_r\mathfrak{S}_v = \mathfrak{S}_w - \sum_{w'} \mathfrak{S}_{w'}$$

summed over all permutations $w' = vt_{qr}$ where $q < r$ and $\ell(w') = \ell(v) + 1 = \ell(w)$. Hence $w'(q) = v(r) > v(q) = w(q)$, and $w'(j) = w(j)$ for $j < q$.

Let us arrange the permutations of a given length p in reverse lexicographical ordering, so that if $\ell(w) = \ell(w') = p$ then w' precedes w if and only if for some $i \geq 1$ we have

$$w'(j) = w(j) \text{ for } j < i, \text{ and } w'(i) > w(i).$$

For this ordering there is a first element, namely the permutation $(p+1, 1, 2, \ldots, p)$.

We have proved

(4.16) *For each permutation* $w \neq 1$ *the Schubert polynomial* $\mathfrak{S}_w$ *can be expressed in the form*

$$\mathfrak{S}_w = x_r\mathfrak{S}_v + \sum_{w'} \mathfrak{S}_{w'}$$

where r *is the last descent of* w, $\ell(v) = \ell(v) - 1$ *and each* w' *in the sum precedes* w *in the reverse lexicographical ordering.* ‖

From (4.16) we deduce immediately that

(4.17) *For each permutation* w, S_w *is a polynomial in* $x_1, x_2, \ldots$ *with positive integral coefficients.*

For we may assume, as inductive hypothesis, that (4.17) is true for all permutations v such that either $\ell(v) - \ell(w)$, or $\ell(v) = \ell(w)$ and v precedes w in the reverse lexicographical ordering; and then (4.16) shows that the result is true for w. (The permutation $(p+1, 1, 2, \ldots, p)$ has code (p), hence is dominant with Schubert polynomial x_1^p by (4.7).) ‖

Now fix integers m, n such that $1 \leq m < n$, and let $w \in S^{(n)}$, so that $\mathfrak{S}_w \in P_n$. By (4.12) we can express $\mathfrak{S}_w$ uniquely in the form

$$(4.18) \qquad \mathfrak{S}_w(x_1, \ldots, x_n) = \sum_{u,v} d^w_{uv}\mathfrak{S}_u(x_1, \ldots, x_m)\mathfrak{S}_v(x_{m+1}, \ldots, x_n)$$

summed over $u \in S^{(m)}$ and $v \in S^{(n-m)}$.

(4.19) *The coefficients* d^w_{uv} *in* (4.18) *are non-negative integers.*

Proof: We proceed by induction on $\ell(v)$. Suppose first that $d^w_{uv} \neq 0$ and that $\ell(v) > 0$, so that $v \neq 1$. Then there exists $j > m$ such that $\partial_j \mathfrak{S}_v(x_{m+1}, \ldots, x_n) \neq 0$. From (4.18) we conclude that $\partial_j \mathfrak{S}_w \neq 0$, hence is equal to $\mathfrak{S}_{ws_j}$, and therefore we have $d^w_{u,v} = d^{ws_j}_{u,vs_{j-m}}$ and $\ell(vs_j) = \ell(v) - 1$. By the inductive hypothesis, we conclude that $d^w_{uv} \geq 0$ if $v \neq 1$.

It remains to consider the case $v = 1$. Let $\rho_m : P_n \to P_m$ be the homomorphism for which $\rho(x_i) = x_i$ if $i \leq m$, and $\rho(x_i) = 0$ if $i > m$. From (4.18) we have

$$\rho_m \mathfrak{S}_w = \sum_u d^w_{u,1} \mathfrak{S}_u. \tag{2}$$

Let r be the last descent of w. If $r \leq m$ then $\mathfrak{S}_w \in P_r$ and hence $\rho_m \mathfrak{S}_w = \mathfrak{S}_w$, so that $d^w_{u,1}$ is equal to 1 if $u = w$, and is zero otherwise. If $r > m$ we deduce from (4.16) that

$$\rho_m \mathfrak{S}_w = \sum_{w'} \rho_m \mathfrak{S}_{w'}. \tag{3}$$

Assume that the coefficients $d^{w'}_{u,1}$ are ≥ 0 whenever w' precedes w in the reverse lexicographical ordering. Then it follows from (2) and (3) that each $d^w_{u,1} \geq 0$. (As remarked before (4.16), the first element in this ordering (if $\ell(w) = p$) is the permutation $(p+1, 1, 2, \ldots, p)$, for which the last descent r is equal to 1.) ‖

Kohnert's algorithm

Let D be a "diagram", which for present purposes means any finite non empty set of lattice points (i, j) in the positive quadrant ($i \geq 1, j \geq 1$). Choose a point $p = (i, j) \in D$ which is rightmost in its row, and suppose that not all the points $(1, j), \ldots, (i-1, j)$ directly above p belong to D. If h is the largest integer less than i such that $(h, j) \notin D$, let D_1 denote the diagram obtained from D by replacing $p = (i, j)$ by (h, j). We can then repeat the process on D_1, by choosing the rightmost element in some row, and obtain a diagram D_2, and so on. Let $K(D)$ denote the set of all diagrams (including D itself) obtainable from D by a sequence of such moves.

Next, we associate with each diagram D a monomial

$$x^D = \prod_{i \geq 1} x_i^{a_i}$$

where a_i is the number of elements of D in the i^{th} row, i.e., the number of j such that $(i, j) \in D$.

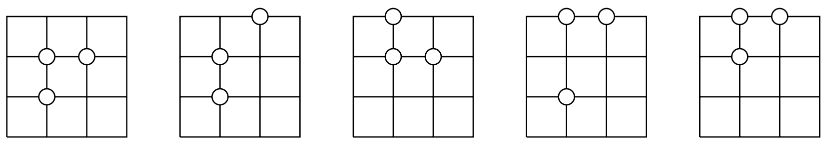
With this notation established, Kohnert's algorithm states that

(4.20) *For each permutation w we have*

$$\mathfrak{S}_w = \sum_{D \in K(D(w))} x^D$$

where $D(w)$ is the diagram (1.20) of w.

Example. If $w = (1432)$, $K(D(w))$ consists of the diagrams



and $\mathfrak{S}_w = x_2^2 x_3 + x_1 x_2 x_3 + x_1 x_2^2 + x_1^2 x_3 + x_1^2 x_2$.

A proof of a related algorithm by N. Bergeron is given in the Appendix to this chapter. The present status of (4.20) is that it is true for w vexillary [K], but open in general.

The shift operator

Let $f \in P_n$ and let $m \geq n$. Then

$$\tau f = \tau_m f = \partial_1 \cdots \partial_m (x_1 \cdots x_m f) \tag{4.21}$$
$$= \pi_1 \cdots \pi_m (f)$$

is independent of m, because $\pi_m f = f$ if f is symmetrical in x_m and x_{m+1}, and in particular if f does not contain x_m, x_{m+1}.

The operator $\tau : P_n \to P_{n+1}$ is called the *shift operator*. For example, we have

$$\tau x_1 = \partial_1 (x_1^2) = x_1 + x_2$$

and for $i \geq 2$,

$$\begin{aligned} \tau x_i &= \partial_1 \cdots \partial_i (x_1 \cdots x_{i-1} x_i^2) \\ &= \partial_1 \cdots \partial_{i-1} (x_1 \cdots x_{i-1} (x_i + x_{i+1})) \\ &= x_{i+1} \partial_1 \cdots \partial_{i-1} (x_1 \cdots x_{i-1}) \\ &= x_{i+1} \end{aligned}$$

so that by (4.4)

$$\tau \mathfrak{S}_{s_i} = \tau(x_1 + \cdots + x_i) = x_1 + \cdots + x_{i+1} = \mathfrak{S}_{s_{i+1}}.$$

More generally,

(4.22) *For all permutations w,*

$$\tau \mathfrak{S}_w = \mathfrak{S}_{1 \times w}$$

where $1 \times w$ *is the permutation* $(1, w(1)+1, w(2)+1, \ldots)$.

Proof: For each $r \geq 1$ let $w_0^{(r)}$ be the longest element of S_r, and let $\delta_r = (r-1, r-2, \ldots, 1)$. Then if $w \in S_n$ we have

$$\begin{aligned}\tau \mathfrak{S}_w &= \partial_1 \cdots \partial_n (x_1 \cdots x_n \partial_{w^{-1}w_0^{(n)}} x^{\delta_n}) \\ &= \partial_1 \cdots \partial_n \partial_{w^{-1}w_0^{(n)}} (x^{\delta_{n+1}}).\end{aligned}$$

Now $s_1 \cdots s_n$ is the cycle $1 \to 2 \to \cdots \to n+1 \to 1$, and hence

$$s_1 \cdots s_n w^{-1} w_0^{(n)} = (1 \times w)^{-1} w_0^{(n+1)}$$

so that

$$\ell(s_1 \cdots s_n w^{-1} w_0^{(n)}) = \ell(s_1 \cdots s_n) + \ell(w^{-1} w_0^{(n)})$$

and therefore by (2.7) we have

$$\tau \mathfrak{S}_w = \partial (1 \times w)^{-1} w_0^{(n+1)} (x^{\delta_{n+1}}) = \mathfrak{S}_{1 \times w}. \,\|$$

(4.23) *Let* $\alpha \in \mathbf{N}^n$ *and* $0 \leq p_1 \leq \cdots \leq p_n$. *Then*

$$\tau s_\alpha(X_{p_1}, \cdots, X_{p_n}) = s_\alpha(X_{p_1+1}, \ldots, X_{p_n+1}).$$

Proof: Since $\tau = \pi_1 \pi_2 \cdots \pi_{p_n}$, this follows from (3.10). ‖

(4.24) *We have*

$$\partial_i \tau^r = 0$$

for $1 \leq i \leq r$.

Proof: By (4.12) it is enough to show that $\partial_i \tau^r \mathfrak{S}_w = 0$ for all permutations w, and this follows from (4.22) and (4.2). ‖

For each $n \geq 1$ let $\rho_n : P_\infty \mapsto P_n$ be the homomorphism defined by

$$\rho_n(x_i) = \begin{cases} x_i & \text{if } i \leq n, \\ 0 & \text{if } i > n. \end{cases}$$

(4.25) *Let* $w_0^{(n)}$ *be the longest element of* S_n. *Then*

$$\pi_{w_0^{(n)}}(f) = \rho_n \tau^n (f)$$

for all $f \in P_n$.

Proof: By linearity we may assume that $f = x^\alpha$ where $\alpha \in \mathbf{N}^n$. Since $x^\alpha = s_\alpha(X_1, \ldots, X_n)$ by (3.5′), we have

$$\tau^n(x^\alpha) = s_\alpha(X_{n+1}, \ldots, X_{2n})$$

by (4.23), and hence

$$\rho_n \tau^n(x_\alpha) = s_\alpha(X_n, \dots, X_n)$$

which is equal to $\pi_{w_0^{(n)}}(x^\alpha)$ by (2.16′). ‖

Transitions

A *transition* is an equation of the form

$$(\mathbf{T}(\mathbf{w}, \mathbf{r})) \qquad \mathfrak{S}_w = x_r \mathfrak{S}_u + \sum_{v \in \Phi} \mathfrak{S}_v$$

where $r \geq 1$, w and u are permutations and Φ is a set of permutations. It exists only for certain values of r, depending on w. An example is (4.16), in which r is the last descent of w.

By (4.15′) we have

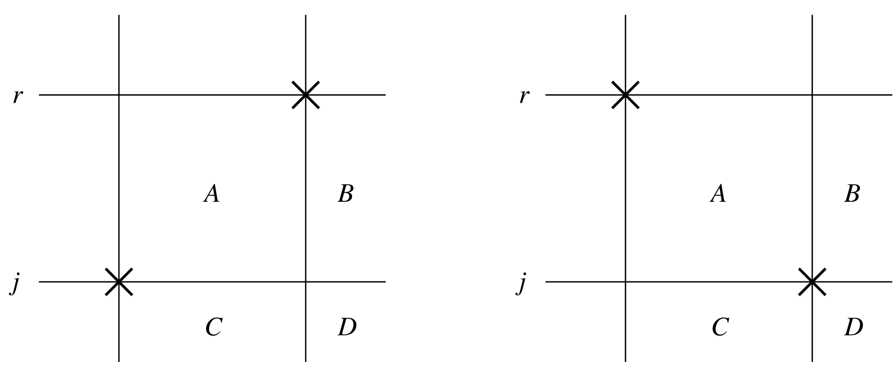
$$x_r \mathfrak{S}_u = \sum_t \sigma(t) \mathfrak{S}_{ut}$$

summed over transpositions $t = t_{ir}$ such that $\ell(ut) = \ell(u) + 1$, where $\sigma(t)$ is the sign of $i - r$. So for $T(w, r)$ to hold there must be exactly one $j > r$ such that

$$(1) \qquad \ell(ut_{rj}) = \ell(u) + 1,$$

$$(2) \qquad w = ut_{rj}.$$

Consider the graphs $G(w)$ and $G(u)$ of w and u. They differ only in rows r and j :



$G(w)$ $\qquad\qquad$ $G(u)$

By (1.10) the relation (1) above is equivalent to $A \cap G(u) = \emptyset$, where A is the open region indicated in the diagram. Moreover, j is the only integer $> r$ such that $u(j) > u(r)$ and $A \cap G(u) = \emptyset$, and this will be the case if and only if $(A \cup B \cup C) \cap G(u)$ is empty. Since $(A \cup B \cup C) \cap G(u) = (A \cup B \cup C) \cap G(w)$, it follows that

(4.26) *There is a transition* $T(w, r)$ *if and only if*

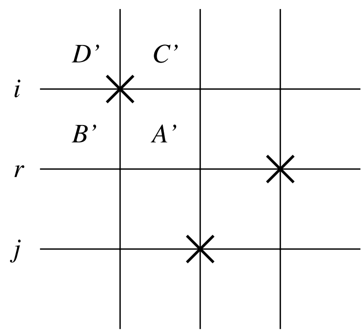
$$(A \cup B \cup C) \cap G(w) = \emptyset.$$

From (4.26) it follows that if $T(w,r)$ exists we must have $w(r) > w(r+1)$, i.e., r must be a *descent* of w. Hence

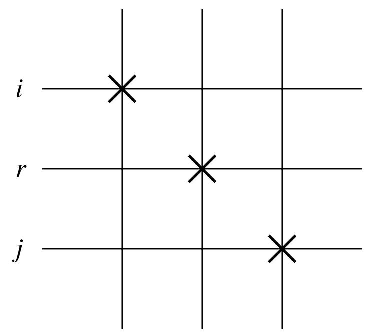
$$d_0(w) \le r \le d_1(w)$$

where $d_0(w)$ (resp. $d_1(w)$) is the first (resp. last) descent of w. (In terms of the code $c(w), d_0(w)$ is the first descent of the sequence $c(w)$, and $d_1(w)$ is the largest i such that $c_i(w) \ne 0$.) In general, not all descents of w will give rise to transitions, but the last descent always does, by (4.16).

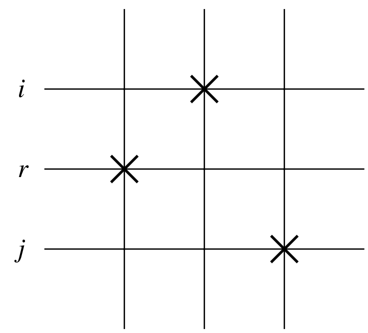
Consider next the set $\Phi = \Phi(w,r)$ of permutations that feature in $T(w,r)$. Each $v \in \Phi$ is of the form $v = ut_{ir}$ with $i < r$ and $\ell(v) = \ell(u) + 1$ $(= \ell(w))$. Again by (1.10), this means that



$G(w)$



$G(u)$



$G(v)$

$A' \cap G(w)$ is empty, where A' is the open region indicated in the diagram above.

The element $v = ut_{ir}$ of Φ for which i is maximal is called the *leader* of Φ. Thus $v \in \Phi$ is the leader if and only if

$$(A' \cup B') \cap G(w) = \emptyset. \tag{4.27}$$

(4.28) *Remark.* The set Φ will be empty if and only if there is no $i < r$ such that $w(i) < w(j)$. We can always avoid this possibility by replacing w by $1 \times w$. If $\Phi(w,r)$ is not empty, then $v \mapsto 1 \times v$ is a bijection of $\Phi(w,r)$ onto $\Phi(1 \times w, r+1)$.

The condition (4.26) is stable under reflection in the main diagonal, which interchanges $G(w)$ and $G(w^{-1})$. Hence

(4.29) *The transition $T(w,r)$ exists if and only if $T(w^{-1},s)$ exists, where $s = w(j)$. Moreover we have*

$$\Phi(w^{-1},s) = \Phi(w,r)^{-1}$$

so that $T(w^{-1},s)$ is the relation

$$\mathfrak{S}_{w^{-1}} = x_s \mathfrak{S}_{u^{-1}} + \sum_{v \in \Phi} \mathfrak{S}_{v^{-1}}.$$

We may notice directly one corollary of (4.29). Let

$$\mathfrak{S}_w(1) = \mathfrak{S}_w(1, 1, \ldots)$$

be the number of monomials in $\mathfrak{S}_w$, each counted with its multiplicity. (By (4.17), $\mathfrak{S}_w$ is a positive sum of monomials.) If $T(w,r)$ is a transition, we have

$$\mathfrak{S}_w(1) = \mathfrak{S}_u(1) + \sum_{v \in \Phi} \mathfrak{S}_v(1)$$

and also, by (4.29)

$$\mathfrak{S}_{w^{-1}}(1) = \mathfrak{S}_{u^{-1}}(1) + \sum_{v \in \Phi} \mathfrak{S}_{v^{-1}}(1).$$

From these two relations it follows, by induction on $\ell(w)$ and on the integer $\mathfrak{S}_w(1)$, that

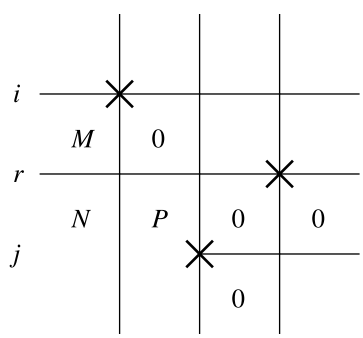
(4.30) $$\mathfrak{S}_w(1) = \mathfrak{S}_{w^{-1}}(1)$$

or in other words that $\mathfrak{S}_w$ and $\mathfrak{S}_{w^{-1}}$ each contain the same number of monomials. So if Kohnert's algorithm (4.20) is true, we should have

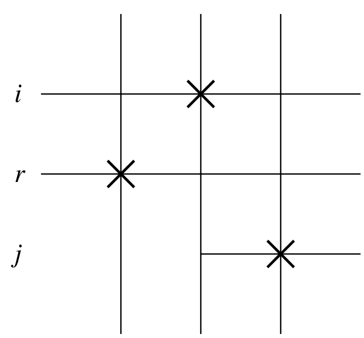
$$\text{Card } K(D(w)) = \text{Card } K(D(w^{-1})).$$

Doubtless the combinatorialists will seek a "bijective" proof of this fact.

Let $T(w,r)$ be a transition and let $v \in \Phi(w,r)$. Consider again the graphs of w and v :



$G(w)$



$G(v)$

Let m, n, p denote respectively the number of points of $G(w)$ (or equivalently $G(v)$) in the open regions of M, N, P. (The regions marked with a zero contain no graph points.) Then we have

(4.31) $$c_i(w) = m + n, \qquad c_r(w) = n + p + 1,$$
$$c_i(v) = m + n + p + 1, \qquad c_r(v) = n,$$

and $c_k(v) = c_k(w)$ *if* $k \neq i, r$. *In particular,* $c_r(w) > c_r(v)$ *for all* $v \in \Phi(w,r)$.

Proof: $c_i(w)$ is the number of positive integers $k > i$ such that $w(k) < w(i)$, hence is equal to $m+n$. Similarly for the other assertions. ‖

Suppose first that $m = 0$, i.e (by (4.27)) that v is the leader of Φ. Then from (4.31) we have $c_i(w) = c_r(v)$ and $c_r(w) = c_i(v)$. Hence in this case $c(v) = t_{ir}c(w)$ and therefore $\lambda(v) = \lambda(w)$.

If on the other hand $m > 0$, there are two possibilites :

either

$$c_i(v) > c_i(w) \geq c_r(w) > c_r(v),$$

or

$$c_i(v) > c_r(w) > c_i(w) > c_r(v).$$

In both cases it follows that $\lambda(v)$ is of the form $R^a\lambda(w)$, where R is a raising operator and $a \geq 1$. Hence $\lambda(v) > \lambda(w)$ (for the dominance partial ordering on partitions), and we have proved

(4.32) *If $T(w,r)$ is a transition, we have $\lambda(v) \geq \lambda(w)$ for all $v \in \Phi(w,r)$, with equality if and only if v is the leader of Φ.* ‖

Recall (1.26) that for any permutation w we have

$$\lambda(w)' \geq \lambda(w^{-1}).$$

Hence for $v \in \Phi(w,r)$ we have

$$\lambda(w)' \overset{(*)}{\geq} \lambda(v)' \geq \lambda(v^{-1}) \overset{(*)}{\geq} \lambda(w^{-1}) \tag{4.33}$$

by (4.29) and (4.32). Moreover, at least one of the inequalities $(*)$ is strict unless v is the leader of $\Phi(w,r)$ *and* v^{-1} is the leader of $\Phi(w^{-1},s)$ (in the notation of (4.29)). In the notation of the diagram preceding (4.27) this means that

$$(A' \cup B' \cup C') \cap G(w) = \emptyset$$

and hence, as in the proof of (4.26), that Card $\Phi \leq 1$.

(4.34) *If $T(w,r)$ is a transition with w vexillary, then $\Phi(w,r)$ is either empty or consists of one vexillary permutation.*

Proof: Suppose that Φ is not empty, and let $v \in \Phi$. By (1.27) we have $\lambda(w)' = \lambda(w^{-1})$, and hence all the inequalities in (4.33) are equalities. Thus v is vexillary, and by the remarks above it is the only member of Φ. ‖

(4.35) *Let* $T(w,r)$ *be a transition with* $r > d_0(w)$. *Then*

$$d_0(v) \geq d_0(w)$$

for all $v \in \Phi(w,r)$.

Proof: As before, let $v = ut_{ir}$ with $i < r$, and let $d_o(w) = d$. We have to show that

$$(*) \qquad c_1(v) \leq \cdots \leq c_d(v).$$

We distinguish three cases:

(a) $i > d$, so that $d \leq i - 1$ and therefore $c_k(v) = c_k(w)$ for $1 \leq k \leq d$.

(b) $i = d$. In this case we have $c_k(v) = c_k(w)$ for $1 \leq k \leq d-1$, and

$$c_{d-1}(v) = c_{d-1}(w) \leq c_d(w) < c_d(v)$$

by (4.31), so that $c_{d-1}(v) < c_d(v)$.

(c) $i > d$. Since $d < r$ we have $i+1 < r$ and $c_i(w) \leq c_{i+1}(w)$, hence $w(i+1) > w(i)$. The diagram on p. 58 shows that $w(i+1) > w(j)$, or equivalently $v(i+1) > v(i)$, so that $c_i(v) \leq c_{i+1}(v)$. Hence

$$c_{i-1}(v) = c_{i-1}(w) \leq c_i(w) < c_i(v) \leq c_{i+1}(v)$$

and therefore

$$c_{i-1}(v) < c_i(v) \leq c_{i+1}(v).$$

Since the sequences $(c_1(v), \ldots, c_d(v))$ and $(c_1(w), \ldots, c_d(w))$ differ only in the i^{th} place, we have $c_1(v) \leq \cdots \leq c_d(v)$ as required. ‖

The *maximal* transition for w is $T(w, d_1(w))$. Let us temporarily write $w \to v$ to mean that $v \in \Phi(w, d_1(w))$.

(4.36) *Suppose that*

$$w = w_o \to w_1 \to \cdots \to w_p$$

is a chain of maximal transitions in which none of the w_i *is Grassmannian. Then*

$$p < (d_1(w) - d_0(w))\ell(w).$$

Proof: For any permutation v, let $e(v) = d_1(v) - d_0(v) \geq 0$. Also let $f(v)$ denote the last nonzero term in the sequence $c(v)$, i.e. $f(v) = c_{d_1(v)}(v)$. Recall that v is Grassmannian if and only if it has only one descent, that is to say if and only if $e(v) = 0$.

From (4.35) we have

$$d_0(w_k) \geq d_0(w_{k-1})$$

for $1 \leq k \leq p$, and from (4.31) we have

$$(1) \qquad c_r(w_k) < c_r(w_{k-1})$$

where $r = d_1(w_{k-1})$. Hence $d_1(w_k) \leq d_1(w_{k-1})$ and therefore

$$e(w_k) \leq e(w_{k-1}).$$

Moreover, if $e(w_k) = e(w_{k-1})$ we must have $d_1(w_k) = d_1(w_{k-1})$ and hence by (1)

$$f(w_k) < f(w_{k-1}).$$

It follows that the $p+1$ points $(x_k, y_k) = (e(w_k), f(w_k))$ are all distinct. Since they all satisfy $1 \leq x_k \leq e(w)$ and $1 \leq y_k \leq \ell(w)$, we have $p+1 \leq e(w)\ell(w)$, as required. ‖

The rooted tree of a permutation

In what follows we shall when necessary replace a permutation w by $1 \times w$, in order to ensure that at each stage the set $\Phi(w,r)$ is not empty (4.28). Observe that this replacement does not change the bound $(d_1(w) - d_0(w))\ell(w)$ in (4.36).

The *rooted tree* T_w of a permutation w defined as follows :

(i) if w is vexillary, then $T_w = \{w\}$;

(ii) if w is not vexillary, take the maximal transition for w:

$$(*) \qquad \mathfrak{S}_w = x_r \mathfrak{S}_u + \sum_{v \in \Phi} \mathfrak{S}_v$$

where $r = d_1(w)$. (If Φ is empty, replace w by $1 \times w$ as explained above.) To obtain T_w, join w by an edge to each $v \in \Phi$, and attach to each $v \in \Phi$ its tree T_v.

By (4.36), T_w is a finite tree, and by construction all its endpoints are vexillary permutations of length $\ell(w)$. It follows from (4.28) that $v \mapsto 1 \times v$ is a bijection of T_w onto $T_{1\times w}$. Thus T_w depends (up to isomorphism) only on the diagonal equivalence class (Chapter I) of the permutation w.

Recall that $\rho_m : P_\infty \to P_m$ is the homomorphism defined by $\rho_m(x_i) = x_i$ if $1 \leq i \leq m$, and $\rho_m(x_i) = 0$ if $i > m$.

(4.37) *Let V be the set of endpoints of T_w. Then if $m \leq d_0(w)$ we have*

$$\rho_m(\mathfrak{S}_w) = \sum_{v \in V} s_{\lambda(v)}(X_m).$$

Proof: If w is vexillary we have $\rho_m(\mathfrak{S}_w) = s_{\lambda(w)}(X_m)$ by (4.4), since $\phi_1(w) = d_0(w) \geq m$. If w is not vexillary, it follows from the maximal transition $(*)$ above that

$$\rho_m(\mathfrak{S}_w) = \sum_{v \in \Phi} \rho_m(\mathfrak{S}_v)$$

since $r = d_1(w) > d_0(w) \geq m$. The result now follows by induction on $\mathrm{Card}(T_w)$. ‖

Multiplication of Schur functions

Let μ, ν be partitions and let $u \in S_n, u' \in S_p$ be Grassmannian permutations of shapes μ, ν respectively. Let $w = u \times u' \in S_{n+p}$, so that by (4.6) and (4.8)

$$\begin{aligned} \mathfrak{S}_w &= \mathfrak{S}_u \cdot \mathfrak{S}_{1_n \times u'} \\ &= s_\mu(X_r) s_\nu(X_s) \end{aligned}$$

where $r = d_0(u)$ and $s = n + d_0(u')$. Hence if $m \leq r$ we have

$$\rho_m(\mathfrak{S}_w) = s_\mu(X_m) s_\nu(X_m)$$

and so by (4.37)

$$s_\mu(X_m) s_\nu(X_m) = \sum_{v \in V} s_{\lambda(v)}(X_m)$$

where V is the set of endpoints of the tree T_w. Here the integer m can be aribtrarily large, because we can replace w by $1_k \times w$ for any positive integer k. Consequently we have

(4.38) $$s_\mu s_\nu = \sum_{v \in V} s_{\lambda(v)}$$

where V is the set of endpoints of the tree $T_{u \times u'}$, and u (resp. u') is Grassmannian of shape μ (resp. ν). ‖

The same argument evidently applies to the product of any number of Schur functions. If $\mu^{(1)}, \ldots, \mu^{(k)}$ are partitions, let $u_i \in S_{n_i}$ be a Grassmannian permutation of shape $\mu^{(i)}$, for each $i = 1, \ldots, k$ (so that $n_i \geq \ell(\mu^{(i)}) + \ell(\mu^{(i)\prime})$) and let $w = u_1 \times \cdots \times u_k$. Then

(4.38′) $$s_{\mu^{(1)}} \cdots s_{\mu^{(k)}} = \sum_{v \in V} s_{\lambda(v)}$$

where V is the set of endpoints of the tree T_w. ‖

In particular, suppose that each $\mu^{(i)}$ is one-part partition, say $\mu^{(i)} = (\mu_i)$, so that the left-hand side of (4.38′) becomes $h_{\mu_1} h_{\mu_2} \cdots = h_\mu$. Correspondingly, each u_i is a cycle of length $\mu_i + 1$, namely $u_i = (\mu_i + 1, 1, 2, \ldots, \mu_i)$. Now [M, Ch.I, §6] the coefficient of a Schur function s_λ in h_μ is the Kostka number $K_{\lambda\mu}$. Hence we have

(4.39) *$K_{\lambda\mu}$ is the number of endpoints of shape λ in the tree of $w = u_1 \times u_2 \times \cdots$.* ‖

Schubert polynomials for S_4

w	$\mathfrak{S}_w$
1234	1
1243	$x_1 + x_2 + x_3$
1324	$x_1 + x_2$
1342	$x_1x_2 + x_1x_3 + x_2x_3$
1423	$x_1^2 + x_1x_2 + x_2^2$
1432	$x_1^2x_2 + x_1^2x_3 + x_1x_2^2 + x_1x_2x_3 + x_2^2x_3$
2134	x_1
2143	$x_1^2 + x_1x_2 + x_1x_3$
2314	x_1x_2
2341	$x_1x_2x_3$
2413	$x_1^2x_2 + x_1x_2^2$
2431	$x_1^2x_2x_3 + x_1x_2^2x_3$
3124	x_1^2
3142	$x_1^2x_2 + x_1^2x_3$
3214	$x_1^2x_2$
3241	$x_1^2x_2x_3$
3412	$x_1^2x_2^2$
3421	$x_1^2x_2^2x_3$
4123	x_1^3
4132	$x_1^3x_2 + x_1^3x_3$
4213	$x_1^3x_2$
4231	$x_1^3x_2x_3$
4312	$x_1^3x_2^2$
4321	$x_1^3x_2^2x_3$

Appendix

A Combinatorial Construction of the Schubert Polynomials

by Nantel Bergeron

In this appendix, we shall give a combinatorial rule based on diagrams for the construction of the Schubert polynomials. A different algorithm had been conjectured (and proved in the case of vexillary permutations) by A. Kohnert. We shall give, at the end of this appendix, a sketch of how one can show the equivalence of the two rules. I wish to acknowledge my indebtedness to Mark Shimozono for the stimulating exchanges regarding this work.

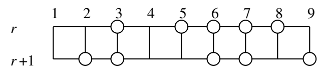
Combinatorial construction†

Here a *"diagram"* will be any finite non empty set of lattice points (i, j) in the positive quadrant $(i \geq 1,\ j \geq 1)$. For example the diagram $D(w)$ of a permutation w is a diagram in the above sense. Let D be any diagram. We denote by $D_{(r,r+1)}$ the diagram D restricted to the row r and $r+1$. Let $j(r, D) = (j_1, j_2, \ldots, j_k)$ be the columns of D in which there is exactly one element of $D_{(r,r+1)}$ per column. Choose a column $j_i \in j(r, D)$. Assume first that $(r+1, j_i) \in D_{(r,r+1)}$. If $i = k$ or if $(r, j_{i+1}) \in D_{(r,r+1)}$, let D_1 be the diagram obtained from D by replacing the element $(r+1, j_i)$ by (r, j_i). Now suppose instead that $(r, j_i) \in D_{(r,r+1)}$. We say that the point (r, j_i) is *r-fixed* with respect to $D(w)$ if the number of elements of D in the column j_i and in the rows $r' > r$ is equal to the number of elements of $D(w)$ in the same area. Now if $i = 1$ (and if there is no r-fixed element with respect to $D(w)$ in D) or if $(r+1, j_{i-1}) \in D_{(r,r+1)}$, let D_1 be the diagram obtained from D by

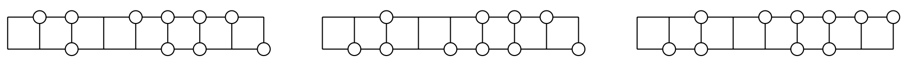
† *(Footnote added in February 2025.)* After the initial publication in 1991, it was realized that the set $\Omega(w)$ as it is defined here includes too many diagrams. Instead, one needs to work with the following definition, and adapt the proofs acccordingly. Full details can be found in the paper: *N. Bergeron, A combinatorial construction of Schubert polynomials, JCTA 60 (1992) 168-182.*

For a D diagram with label $e\colon D \to \mathbf{N}$, choose a column $j_i \in j(r, D)$. When $(r+1, j_i) \in D_{(r,r+1)}$, if $i = k$ or if $(r, j_{i+1}) \in D_{(r,r+1)}$, let D_1 be the diagram obtained from D by replacing the element $(r+1, j_i)$ by (r, j_i) keeping the same labels as in D. When $(r, j_i) \in D_{(r,r+1)}$, if $i = 1$ or $(r+1, j_{i-1}) \in D_{(r,r+1)}$ and if $e_{r,j_i} > r$, let D_1 be the diagram obtained from D by replacing the element (r, j_i) by $(r+1, j_i)$ and labeled as follow. Let $e = e_{r,j_i}$ in D, we put in D_1 the same labels than in D except if $e_{r+1,j} > e$ with $j \leq j_i$ in D then $e_{r+1,j}$ is re-labeled by e in D_1. Let $\Omega(w)$ denote the set of all diagrams (including $D(w)$) obtainable from $D(w)$, with labels $e_{i,j} = i$, by any sequence of B-moves.

replacing the element (r, j_i) by $(r+1, j_i)$. In both cases we say that the diagram D_1 is obtained from D by a *"B-move"* (with respect to $D(w)$). For example let D be such that $D_{(r,r+1)}$ is the following:



For this case $j(r, D) = (2, 5, 8, 9)$. We can perform on this diagram a B-move in column 2, 5 or 9 and obtain, respectively, the following diagrams:

 .

The element in column 8 is not allowed to move since $(r+1, 5) \notin D_{(r,r+1)}$. Let $\Omega(w)$ denote the set of all diagrams (including $D(w)$) obtainable from $D(w)$ by any sequence of B-moves.

Next for $D \in \Omega(w)$ let x^D denote the monomial $x_1^{a_1} x_2^{a_2} x_3^{a_3} \cdots$ where a_i is the number of elements of D in the ith row. For any permutation w we shall have the following theorem:

(B.1)
$$\mathfrak{S}_w = \sum_{D \in \Omega(w)} x^D.$$

To prove this we will proceed by reverse induction on $\ell(w)$. If $w = w_0$ (the longest element of S_n) then (B.1) holds since $\Omega(w_0)$ contains only the element $D(w_0)$ and $x^{D(w_0)} = x^\delta$. On the other hand from (4.3), $\mathfrak{S}_{w_0} = x^\delta$. Now if $w \neq w_0$ then let $r = min\{i : w(i) < w(i+1)\}$. From (4.2) we have

(B.2)
$$\mathfrak{S}_w = \partial_r \mathfrak{S}_{ws_r}.$$

Let $v = ws_r$. By the induction hypothesis equation (B.1) holds for $\mathfrak{S}_v$. The induction step will be to *"apply"* the operator ∂_r to the diagrams in $\Omega(v)$. To this end we need more tools.

For the moment let us fix $D \in \Omega(v)$. Let $a = a_r(D)$ and $b = a_{r+1}(D)$ be respectively the number of elements of D in the rth and $r+1$st rows. We have

(B.3)
$$\partial_r x^D = \partial_r \cdots x_r^a x_{r+1}^b \cdots = \begin{cases} \sum_{k=0}^{a-b-1} \cdots x_r^{a-r-1} x_{r+1}^{b+r} \cdots & \text{if } a > b, \\ 0 & \text{if } a = b, \\ -\sum_{k=0}^{b-a-1} \cdots x_r^{a+r} x_{r+1}^{b-r-1} \cdots & \text{if } a < b. \end{cases}$$

This suggests we define the operator ∂_r directly on the diagram D. For this we need only to concentrate our attention on the rows r and $r+1$ of D. Let $j(r, D) = (j_1, j_2, \ldots, j_p)$. Notice that in all columns $j < w(r)$ of $D_{(r,r+1)}$ there are exactly two elements and in column $w(r) = j_1$ of $D_{(r,r+1)}$ there is exactly one element in position (r, j_1). We shall now reduce the sequence of indices $j(r,D)$ according to the following rule. Let $J_{(0)} = (j_2, j_3, \ldots, j_p)$. Remove from $J_{(0)}$ all pairs j_k, j_{k+1} for which $(r, j_k) \in D$ and $(r+1, j_{k+1}) \in D$. Let us denote the resulting sequence by $J_{(1)}$. Repeat recursively this process on $J_{(1)}$ until no such pair can be found. Let us denote by $f(r,D) = (f_1, f_2, \ldots, f_q)$ the final sequence. From construction, the sequence $f(r,D)$ is such that if $(r, f_k) \in D$ then $(r, f_{k+1}) \in D$. Let $up(r,D)$ be the minimal k such that $(r, f_k) \in D$. If $(r+1, f_q) \in D$ then set $up(r,D) = q+1$. We are now in a position to define the operation of ∂_r on the diagram D. To this end let us first assume that $a > b$. This means that we have $a-b$ more elements in row r then in row $r+1$. Hence $q - up(r,D) + 1 \geq a - b - 1$ for q the length of $f(r,D)$. The equality holds if and only if $up(r,D) = 1$. In the case $a > b$ the operator ∂_r on the diagram D is defined by the map

$$(\mathbf{B.4a}) \qquad \partial_r D \to \{D_0, D_1, D_2, \ldots, D_{a-b-1}\}$$

where D_0 is identical to D except that we remove the element in position $(r, w(r))$ and for $k = 1, 2, .., a-b-1$ we successively set D_k to be identical to D_{k-1} except that the element $(r, f_{up(r,D)+k-1})$ is replaced by $(r+1, f_{up(r,D)+k-1})$. Now if $a < b$ we have $up(r,D) - 1 \geq b - a + 1$ (with equality iff $up(r,D) = q+1$). So $up(r,D) - 1 > b - a$. In this case the operator ∂_r on the diagram D is defined by the map

$$(\mathbf{B.4b}) \qquad \partial_r D \to \{D_0, D_1, D_2, \ldots, D_{b-a-1}\}$$

where D_0 is identical to D except that we remove the element in position (r, w_r) and the element $(r+1, f_{up(r,D)-1})$ is replaced by $(r, f_{up(r,D)-1})$. For $k = 1, 2, .., b-a-1$ we successively set D_k to be identical to D_{k-1} except that the element $(r+1, f_{up(r,D)-k-1})$ is replaced by $(r, f_{up(r,D)-k-1})$. Finally if $a = b$ then

$$(\mathbf{B.4c}) \qquad \partial_r D \to \{\}.$$

With this definition of ∂_r we have that

$$(\mathbf{B.5}) \qquad \partial_r x^D = \pm \sum_{D_i \in \partial_r D} x^{D_i},$$

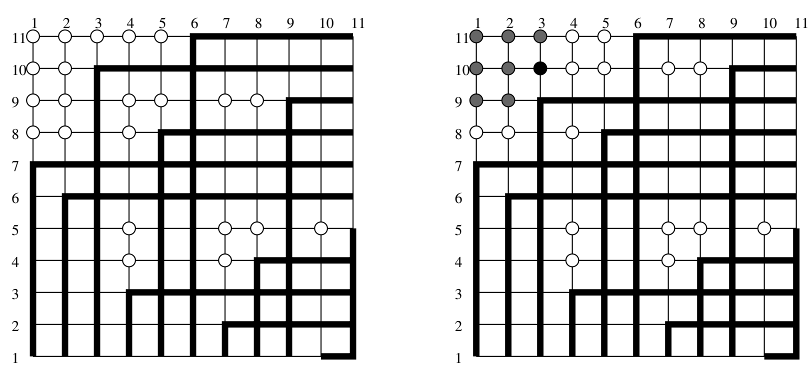
with the positive sign in case (B.4a) and the negative sign in case (B.4b). For (B.4c) the result of (B.5) is zero.

We shall now show that

$$(\mathbf{B.6}) \qquad \partial_r \textit{ maps } \Omega(v) \textit{ into } \Omega(w).$$

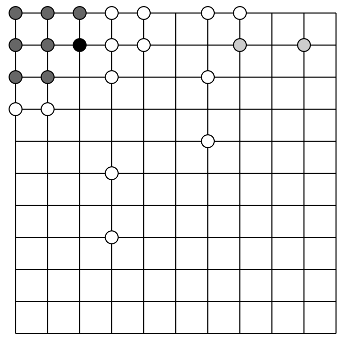
Proof: The reader will notice that in $D(v)$ the rectangle defined by the rows $1, 2, \ldots, r+1$ and the columns $1, 2, \ldots, w(r) - 1$ is filled with elements. None of these elements can B-move. Hence these elements are fixed in any diagram $D \in \Omega(v)$. The same applies to all elements in column $w(r)$; they are packed in the smallest rows and there are no elements in the rows strictly greater than r. Now let D be a diagram in $\Omega(v)$ and assume that $\partial_r D = \{D_0, D_1, \ldots, D_m\}$ is non-empty. The remark above implies that the element in position $(r, w(r))$ does not affect the sequence of B-moves from $D(v)$ to D. Hence we can apply the same sequence of B-moves to $D(v) - \{(r, w(r))\}$ and obtain D_0. Moreover $D(v) - \{(r, w(r))\}$ is obtainable from $D(w)$ by a simple sequence of B-moves in rows $r, r+1$, for this one successively B-moves all the elements in row $r+1$ and columns given by $j(r, D(w))$. This gives that D_0 is obtainable from $D(w)$ by a sequence of B-moves, that is $D_0 \in \Omega(w)$. Now from the construction of $\partial_r D$, D_k $(k > 0)$ is obtained from D_{k-1} by exactly one B-move. Hence $\partial_r D \subset \Omega(w)$. $\|$

It is appropriate at this point to give an example. Let $w = (6, 3, 9, 5, 1, 2, 11, 8, 4, 7, 10)$. Hence $r = 2$ and $v = (6, 9, 3, 5, 1, 2, 11, 8, 4, 7, 10)$. We have depicted below the diagrams $D(w)$ and $D(v)$. In our example the fixed elements described above are colored in grey and the element in position $(r, w(r))$ is colored black.

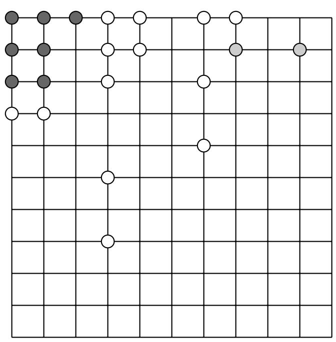


$D(w)$ $D(v)$

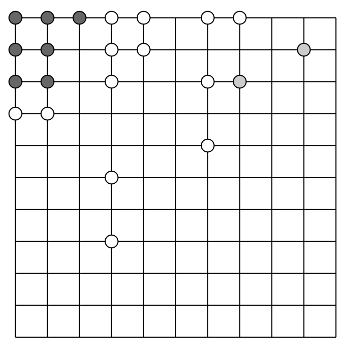
Now let D be the following diagram of $\Omega(v)$.

$D \quad =$ 

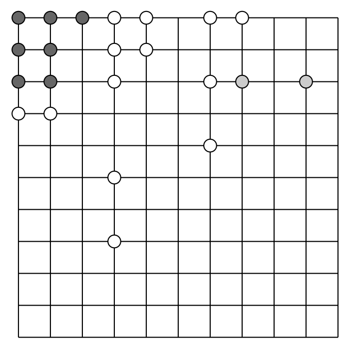
Here, $a_r(D) = 7$, $a_{r+1}(D) = 4$ and $j(r, D) = (3, 5, 7, 8, 10)$. The reduced sequence $f(r, D)$ is $(8, 10)$ and $up(r, D) = 1$. Hence $\partial_r D = \{D_0, D_1, D_2\}$ where



D_0



D_1



D_2

To prove (B.1) the first step is to find a subset of $\Omega(v)$ such that when we operate with ∂_r we obtain $\Omega(w)$. To this end let

$$\Omega_0(v) = \{D \in \Omega(v) : a_r(D) > a_{r+1}(D) \ \ and \ \ up(r, D) = 1\}.$$

We have

$$\textbf{(B.7)} \qquad \Omega(w) = \bigcup_{D \in \Omega_0(v)} \partial_r D \qquad (disjoint \ \ union).$$

Proof: It is clear from construction that the subsets $\partial_r D$ are disjoint when $D \in \Omega_0(v)$. From (B.6) we only have to prove that for any $D' \in \Omega(w)$ there is a $D \in \Omega_0(v)$ such that $D' \in \partial_r D$. To see that, reduce the sequence $j(r, D') = (j_1, \ldots, j_p)$ by removing recursively all pairs j_k, j_{k+1} for which $(r, j_k) \in D'$ and $(r + 1, j_{k+1}) \in D'$. Denote the final sequence by $f'(r, D')$. Let D be the bubble diagram obtained from D' by adding an element in position $(r, w(r))$ and successively B-moving all elements in positions $(r + 1, f_i) \in D'$. We have that $D \in \Omega(v)$. To see this one applies to $D(v)$ the sequence of B-moves from $D(w)$ to $D - \{(r, w(r))\}$. Of course one should ignore any B-move in rows

$r, r+1$ performed on the original elements of $D(v)$ in row r. But by the choice of r, the other B-moves apply almost directly and the resulting diagram is precisely D. Moreover since $f(r, D) = f'(r, D')$ and $up(r, D) = 1$ we have $D \in \Omega_0(v)$ and $D' \in \partial_r D$. $\|$

We shall now investigate the effect of ∂_r on $\Omega_1(v) = \Omega(v) - \Omega_0(v)$. More precisely we have

(B.8)
$$\sum_{D \in \Omega_1(v)} \partial_r x^D = 0.$$

Proof: There are two classes of diagrams in $\Omega_1(v)$. The first class contains the diagrams D for which $a_r(D) = a_{r+1}(D)$. In this case it is trivial that $\partial_r x^D = 0$. The other class is formed by the diagrams D such that $a_r(D) \neq a_{r+1}(D)$ and $up(r, D) > 1$. In this case we shall construct an involution, $D \to D'$, such that $\partial_r x^D + \partial_r x^{D'} = 0$. Let $f(r, D) = (f_1, f_2, \ldots, f_q)$, $a = a_r(D)$ and $b = a_{r+1}(D)$. We first define the involution for the case $a > b$. Since $up(r, D) > 1$ we must have $q - up(r, D) + 1 \geq a - b$. So let D' be identical to D except that the elements in positions $(r, f_{up(r,D)})$, $(r, f_{up(r,D)+1})$, $\ldots$, $(r, f_{up(r,D)+a-b-1})$ are B-moved to the positions $(r+1, f_{up(r,D)})$, $(r+1, f_{up(r,D)+1})$, $\ldots$, $(r+1, f_{up(r,D)+a-b-1})$. It is clear that $D' \in \Omega(v)$. But $f(r, D') = f(r, D)$ and $up(r, D') > up(r, D) > 1$, hence $D' \in \Omega_1(v)$. Moreover we have $a_r(D) = b$ and $a_{r+1}(D) = a$, hence $\partial_r x^D + \partial_r x^{D'} = 0$. The case $a < b$ is similar to the previous one. $\|$

A proof of (B.1) is now completed combining (B.2), (B.5), (B.7) and (B.8). More precisely using the induction hypothesis, we have

$$\begin{aligned}
\mathfrak{S}_w &= \partial_r \mathfrak{S}_v && (B.2)\\
&= \sum_{D \in \Omega(v)} \partial_r x^D \\
&= \sum_{D \in \Omega_0(v)} \partial_r x^D && (B.8)\\
&= \sum_{D \in \Omega_0(v)} \sum_{D_i \in \partial_r D} x^{D_i} && (B.5)\\
&= \sum_{D' \in \Omega(w)} x^{D'}. \ \| && (B.7)
\end{aligned}$$

Kohnert's construction

Let D be any diagram. Choose $(i, j) \in D$ such that $(i, j') \notin D$ for all $j' > j$. Let us suppose that there is a point $(i', j) \notin D$ with $i' < i$. Then let $h < i$ be the largest integer such that $(h, j) \notin D$ and let D_1 denote the diagram obtained from D by replacing (i, j) by (h, j). We say that D_1 is obtained from D by a *"K-move"*. Now let $K(D(w))$ denote the set of all diagrams (including D

itself) obtainable from D by any sequence of K-moves. Kohnert's conjecture states that for any permutation w we have

$$\mathfrak{S}_w = \sum_{D \in K(D(w))} x^D. \tag{B.9}$$

A. Kohnert has proved (B.9) for the case where w is a vexillary permutation but the general case was still open. For the interested reader here is a sketch of how one may prove (B.9).

We have noticed by computer that $\Omega(w) = K(D(w))$. The idea then is to show both inclusions by induction. The inclusion $K(D(w)) \subset \Omega(w)$ is the easiest one. We only have to show that any K-move of an element (i, j) to (h, j) can be simulated using B-moves. For this we proceed by induction on $i - h$. If $i - h = 1$ then the K-move is simply one B-move. Now if $i - h > 1$, we first perform the sequence of B-moves in row $h, h+1$ necessary to B-move the element $(h+1, j)$ to (h, j). Then using the induction hypothesis we can K-move (i, j) to $(h + 1, j)$. Finally we reverse the first sequence of B-moves in rows $h, h + 1$. That shows $K(D(w)) \subset \Omega(w)$.

The other inclusion needs a lot more work. For $D \in K(D(w))$ and i any row of D let $B_i(D)$ denote the set of all diagrams (including D) obtainable from D by any sequence of B-moves in the rows $i, i + 1$ only. It is clear that if i is big enough then $B_i(D) \subset K(D(w))$. We may then proceed by reverse induction on i. Now for a fixed i, notice that $B_i(D(w))$ is obtainable from $D(w)$ using only K-moves. Let Ω_0 denote the set of all diagrams obtainable from $B_i(D(w))$ by any sequence of K-moves for which no elements crosses the border between the rows $i + 1$ and $i + 2$. A simple inductive algorithm may be used here to show that for any $D \in \Omega_0$ we have $B_i(D) \subset \Omega_0$. Next let Ω_k denote the set of all diagrams of $K(D(w))$ which have k more elements than $D(w)$ in the rows $1, 2, \ldots, i + 1$. For almost all the cases it is fairly easy to show (using induction on k and the induction hypothesis on i) that for $D \in \Omega_k$ we have $B_i(D) \subset \Omega_k$. But some of the cases are really hard to formalize! Now this completed would show that $\Omega(w) \subset K(D(w))$ since $K(D(w)) = \cup\Omega_k$.

Chapter V

Orthogonality

Recall that

$$P_n = \mathbf{Z}[x_1, \ldots, x_n],$$

$$\Lambda_n = \mathbf{Z}[x_1, \ldots, x_n]^{S_n}$$

where $x_1, \ldots, x_n$ are independent indeterminates.

(5.1) *P_n is a free Λ_n-module of rank $n!$ with basis*

$$B_n = \{x^\alpha : 0 \le \alpha_i \le i-1,\ 1 \le i \le n\}.$$

Proof: by induction on n. The result is trivially true when $n = 1$, so assume that $n > 1$ and that P_{n-1} is a free Λ_{n-1}-module with basis B_{n-1}. Since $P_n = P_{n-1}[x_n]$, it follows that P_n is a free $\Lambda_{n-1}[x_n]$-module with basis B_{n-1}. Now

$$\Lambda_{n-1}[x_n] = \Lambda_n[x_n],$$

because the identities

$$e_r(x_1, \ldots, x_n) = \sum_{s=0}^{r} (-x_n)^s e_{r-s}(x_1, \ldots, x_n)$$

show that $\Lambda_{n-1} \subset \Lambda_n[x_n]$, and on the other hand it is clear that $\Lambda_n \subset \Lambda_{n-1}[x_n]$. Hence P_n is a free $\Lambda_n[x_n]$-module with basis B_{n-1}.

To complete the proof it remains to show that $\Lambda_n[x_n]$ is a free Λ_n-module with basis $1, x_n, \ldots, x_n^{n-1}$. Since $\prod_{i=1}^{n} (x_n - x_i) = 0$, we have

$$x_n^n = e_1 x_n^{n-1} - e_2 x_n^{n-2} + \ldots + (-1)^{n-1} e_n,$$

from which it follws that the x_n^{n-i} $(1 \le i \le n)$ generate $\Lambda_n[x_n]$ as a Λ_n-module. On the other hand, if we have a relation of linear dependence

$$\sum_{i=1}^{n} f_i x_n^{n-i} = 0$$

with coefficients $f_i \in \Lambda_n$, then we have also

$$\sum_{i=1}^{n} f_i x_j^{n-i} = 0$$

for $j = 1, 2, \ldots, n$, and since

$$det(x_j^{n-i}) = \prod_{i<j}(x_i - x_j) \neq 0,$$

it follows that $f_1 = \cdots = f_n = 0$. $\|$

As before, let $\delta = (n-1, n-2, \cdots, 1, 0)$. By reversing the order of $x_1, \ldots, x_n$ in (5.1) it follows that

(5.1′) *The monomials* $x^\alpha, \alpha \subset \delta$*(i.e.,* $0 \le \alpha_i \le n-i$ *for* $1 \le i \le n$*) form a* Λ_n*-basis of* P_n. $\|$

We define a scalar product on P_n, with values in Λ_n, by the rule

(5.2) $$< f, g >= \partial_{w_0}(fg) \qquad (f, g \in P_n)$$

where w_0 is the longest element of S_n. Since ∂_{w_0} is Λ_n-linear, so is the scalar product.

(5.3) *Let* $w \in S_n$ *and* $f, g \in P_n$. *Then*

(i) $< \partial_w f, g >=< f, \partial_{w^{-1}} g >$

(ii) $< wf, g >= \epsilon(w) < f, w^{-1} g >$.

where $\epsilon(w) = (-1)^{\ell(w)}$ *is the sign of* w.

Proof: (i) It is enough to show that $< \partial_i f, g >=< f, \partial_i g >$ for $i \le i \le n-1$. We have

$$< \partial_i f, g > = \partial_{w_0}((\partial_i f)g) = \partial_{w_0 s_i} \partial_i((\partial_i f)g)$$
$$= \partial_{w_0 s_i}((\partial_i f)(\partial_i g))$$

because $\partial_i f$ is symmetrical in x_i and x_{i+1}. The last expression is symmetrical in f and g, hence $< \partial_i f, g >=< \partial_i g, f >=< f, \partial_i g >$ as required.

(ii) Again it is enough to show that $< s_i f, g >= - < f, s_i g >$. We have

$$< s_i f, g >= \partial_{w_0}((s_i f)g) = \partial_{w_0 s_i} \partial_i s_i(f(s_i g))$$

and since $\partial_i s_i = -\partial_i$ this is equal to

$$-\partial_{w_0 s_i} \partial_i(f(s_i g)) = -\partial_{w_0}(f(s_i g)) = - < f, s_i g > . \ \|$$

(5.4) *Let* $u, v \in S_n$ *be such that* $\ell(u) + \ell(v) = \binom{n}{2}$. *Then*

$$< \mathfrak{S}_u, \mathfrak{S}_v >= \begin{cases} 1 & \textit{if } v = w_0 u, \\ 0 & \textit{otherwise.} \end{cases}$$

Proof: We have

$$\begin{aligned} <\mathfrak{S}_u, \mathfrak{S}_v> &= <\partial_{u^{-1}w_0} x^\delta, \mathfrak{S}_v> \\ &= <x^\delta, \partial_{w_0 u}\mathfrak{S}_v> \end{aligned}$$

by (5.3). Also $\ell(w_0 u) = \ell(w_0) - \ell(u) = \ell(v)$, hence

$$\partial_{w_0 u}\mathfrak{S}_v = \begin{cases} 1 & \text{if } v = w_0 u, \\ 0 & \text{otherwise.} \end{cases}$$

It follows that

$$<\mathfrak{S}_u, \mathfrak{S}_v> = \begin{cases} 0 & \text{if } v \neq w_0 u, \\ <x^\delta, 1> = \partial_{w_0}(x^\delta) = 1 & \text{if } v = w_0 u. \ \| \end{cases}$$

(5.5) *Let* $u, v \in S_n$. *Then*

$$<w_0\mathfrak{S}_u, \mathfrak{S}_{vw_0}> = \epsilon(v)\delta_{uv}.$$

Proof: We have

$$\begin{aligned} <w_0\mathfrak{S}_u, \mathfrak{S}_{vw_0}> &= <w_0\mathfrak{S}_u, \partial_{w_0 v^{-1} w_0} x^\delta> \\ &= <\partial_{w_0 v w_0}(w_0\mathfrak{S}_u), x^\delta> \\ &= \epsilon(v) <w_0 \partial_v \mathfrak{S}_u, x^\delta> \end{aligned}$$

by (5.3) and (2.12). By (4.2) the scalar product is therefore zero unless $\ell(u) - \ell(v) = \ell(uv^{-1})$, and then it is equal to $\epsilon(v) <w_0\mathfrak{S}_{uv^{-1}}, x^\delta>$. Now $\mathfrak{S}_{uv^{-1}}$ is a linear combination of monomials x^α such that $\alpha \subset \delta$ and $|\alpha| = \ell(u) - \ell(v)$. Hence $w_0(\mathfrak{S}_{uv^{-1}})x^\delta$ is a sum of monomials x^β where

$$\beta = w_0\alpha + \delta \subset w_0\delta + \delta = (n-1, \dots, n-1).$$

Now $\partial_{w_0} x^\beta = 0$ unless all the components β_i of β are distinct; since $0 \le \beta_i \le n-1$ for each i, it follows that $\partial_{w_0} x^\beta = 0$ unless $\beta = w\delta$ for some $w \in S_n$, and in that case

$$w_0\alpha = \beta - \delta = w\delta - \delta$$

must have all its components ≥ 0. So the only possibility that gives a nonzero scalar product is $w = 1, \alpha = 0, u = v$, and in that case

$$\begin{aligned} <w_0\mathfrak{S}_u, \mathfrak{S}_{vw_0}> &= \epsilon(v) <1, x^\delta> \\ &= \epsilon(v)\partial_{w_0}(x^\delta) = \epsilon(v). \ \| \end{aligned}$$

(5.6) *The Schubert polynomials* $\mathfrak{S}_w, w \in S_n$, *form a* Λ_n-*basis of* P_n.

Proof: Let $u, v \in S_n$ and let

$$\textbf{(1)} \qquad w_0\mathfrak{S}_u = \sum_{\alpha \subset \delta} a_{u\alpha} x^\alpha,$$

$$(\mathbf{2}) \qquad \epsilon(v)\mathfrak{S}_{vw_0} = \sum_{\beta\subset\delta} b_{v\beta}x^\beta,$$

with coefficients $a_{u\alpha}, b_{v\beta} \in \Lambda_n$. Let $c_{\alpha\beta} =< x^\alpha, x^\beta >$. Then from (5.5) we have

$$\sum_{\alpha,\beta} a_{u\alpha}c_{\alpha\beta}b_{v\beta} = \delta_{uv},$$

or in matrix terms

$$(\mathbf{3}) \qquad ACB^t = 1$$

where $A = (a_{u\alpha}), B = (b_{v\beta})$ and $C = (c_{\alpha\beta})$ are square matrices of size $n!$, with coefficients in Λ_n. From (3) it follows that each of A, B, C has determinant ± 1; hence the equations (2) can be solved for $x^\beta, \beta \subset \delta$, as Λ_n-linear combinations of the Schubert polynomials $\mathfrak{S}_w, w \in S_n$. Since by (5.1′) the x^β from a Λ_n-basis of P_n, so also do the $\mathfrak{S}_w$. ‖

We have

$$(\mathbf{5.7}) \qquad < f, g >= \sum_{w\in S_n} \epsilon(w)\partial_w(w_0 f)\partial_{ww_0}(g)$$

for all $f, g \in P_n$.

Proof: Let $\Phi(f, g)$ denote the right-hand side of (5.7). We claim first that

$$(\mathbf{1}) \qquad \Phi(f, g) \in \Lambda_n.$$

For this it is enough to show that $\partial_i\Phi = 0$ for $1 \le i \le n-1$. Let

$$A_i = \{w \in S_n : \ell(s_iw) > \ell(w)\},$$

then S_n is the disjoint union of A and s_iA, and $s_iA = Aw_0$. Hence

$$\Phi(f, g) = \sum_{w\in A_i} \epsilon(w)\{\partial_w(w_0f)\partial_i(\partial_{s_iww_0}g) - \partial_i\partial_w(w_0f)(\partial_{s_iww_0}g)\}.$$

Since for all $\phi, \psi \in P_n$ we have

$$\partial_i(\phi\partial_i\psi - (\partial_i\phi)\psi) = (\partial_i\phi)(\partial_i\psi) - (\partial_i\phi)(\partial_i\psi) = 0,$$

it follows that $\partial_i\Phi(f, g) = 0$ for all i as required.

Next, since each operator ∂_w is Λ_n-linear, it follows that $\Phi(f, g)$ is Λ_n-linear in each argument. By (5.6) it is therefore enough to verify (5.7) when $f = w_0\mathfrak{S}_u$ and $g = \mathfrak{S}_{vw_0}$, where $u, v \in S_n$. We have then

$$\Phi(w_0\mathfrak{S}_u, \mathfrak{S}_{vw_0}) = \sum_{w\in S_n} \epsilon(w)\partial_{w^{-1}}(\mathfrak{S}_u)\partial_{w^{-1}w_0}(\mathfrak{S}_{vw_0})$$

which by (4.2) is equal to

(2)
$$\sum_w \epsilon(w)\mathfrak{S}_{uw}\mathfrak{S}_{vw}$$

summed over $w \in S_n$ such that

$$\ell(uw) = \ell(u) - \ell(w^{-1}) = \ell(u) - \ell(w)$$

and

$$\ell(vw) = \ell(vw_0) - \ell(w^{-1}w_0) = \ell(w) - \ell(v).$$

Hence the polynomial (2) is (i) symmetric in $x_1, ..., x_n$ (by (1) above), (ii) independent of x_n, (iii) homogeneous of degree $\ell(u) - \ell(v)$. Hence it vanishes unless $\ell(u) = \ell(v)$ and $u = w^{-1} = v$, in which case it is equal to $\epsilon(w) = \epsilon(v)$. Hence

$$\Phi(w_0\mathfrak{S}_u, \mathfrak{S}_{vw_0}) = \epsilon(v)\delta_{uv} = < w_0\mathfrak{S}_u, \mathfrak{S}_{vw_0} >$$

by (5.5). This completes the proof of (5.7). ‖

Now let $x = (x_1, \dots x_n)$ and $y = (y_1, \dots, y_n)$ be two sequences of independent variables, and let

(5.8)
$$\Delta = \Delta(x, y) = \prod_{i+j \le n} (x_i - y_j)$$

(the "semiresultant"). We have

(5.9)
$$\Delta(wx, x) = \begin{cases} 0 & \text{if } w \ne w_0, \\ \epsilon(w_0)a_\delta(x) & \text{if } w = w_0. \end{cases}$$

For

$$\Delta(wx, x) = \prod_{i+j \le n} (x_{w(i)} - x_j)$$

is non-zero if and only if $w(i) \ne j$ whenever $i + j \le n$, that is to say if and only if $w \ne w_0$; and

$$\begin{aligned} \Delta(w_0x, x) &= \prod_{i+j \le n} (x_{n+1-i} - x_j) \\ &= \prod_{j<k} (x_k - x_j) = \epsilon(w_0)a_\delta(x). \ \| \end{aligned}$$

The polynomial $\Delta(x, y)$ is a linear combination of the monomials $x^\alpha, \alpha \subset \delta$, with coefficients in $\mathbf{Z}[y_1, \dots, y_n] = P_n(y)$, hence by (4.11) can be written uniquely in the form

$$\Delta(x, y) = \sum_{w \in S_n} \mathfrak{S}_w(x)T_w(y)$$

with $T_w(y) \in P_n(y)$. By (5.5) we have

$$T_w(y) = < \Delta(x,y), w_0\mathfrak{S}_{ww_0}(-x) >_x$$

where the suffix x means that the scalar product is taken in the x variables. Hence

$$\begin{aligned} T_w(y) &= \partial_{w_0}(\Delta(x,y)w_0(\mathfrak{S}_{ww_0}(-x))) \\ &= a_\delta(x)^{-1}\sum_{v\in S_n}\epsilon(v)\Delta(vx,y)vw_0(\mathfrak{S}_{ww_0}(-x)) \end{aligned} \tag{1}$$

by (2.10), where $v \in S_n$ acts by permuting the x_i.

Now this expression (1) must be independent of $x_1,\dots,x_n$. Hence we may set $x_i = y_i$ ($1 \le i \le n$). But then (5.9) shows that the only non-zero term in the sum (1) is that corresponding to $v = w_0$, and we obtain

$$T_w(y) = \mathfrak{S}_{ww_0}(-y).$$

Hence we have proved

(5.10) ("Cauchy formula")

$$\Delta(x,y) = \sum_{w\in S_n}\mathfrak{S}_w(x)\mathfrak{S}_{ww_0}(-y). \qquad \|$$

Remark. Let $n = r+s$ where $r,s \ge 1$, and regard $S_r \times S_s$ as a subgroup of S_n, with S_r permuting $1,2,\dots,r$ and S_s permuting $r+1,\dots,r+s$. Let $w_0^{(r)}, w_0^{(s)}$ be the longest elements of S_r, S_s respectively, and let $u = w_0^{(r)} \times w_0^{(s)}$. If $w \in S_n$, we have $\partial_u\mathfrak{S}_w = \mathfrak{S}_{wu}$ if $\ell(wu) = \ell(w) - \ell(u)$, that is to say if wu is Grassmannian (with its only descent at r), and $\partial_u\mathfrak{S}_w = 0$ otherwise. Hence by applying ∂_u to the x-variables in (5.10) we obtain

$$\partial_u\Delta(x,y) = \sum_{v\in G_{rs}}\mathfrak{S}_v(x)\mathfrak{S}_{vuw_0}(-y)$$

where $G_{r,s} \subset S_n$ is the set of Grassmannian permutations v with descent at r (i.e. $v(i) < v(i+i)$ if $i \ne r$). On the other hand, it is easily verified that

$$\partial_u\Delta(x,y) = \prod_{i=1}^{r}\prod_{j=1}^{s}(x_i - y_j)$$

and that $v' = vuw_0$ is the permutation

$$(v(r+1),\dots,v(r+s),v(1),\dots,v(r))$$

hence is also Grassmannian, with descent at s.

The shape of v is

$$\lambda = \lambda(v) = (v(r) - r, \dots, v(2) - 2, v(1) - 1)$$

and the shape of v' is say

$$\mu' = \lambda(v') = (v(r+s) - s, \dots, v(r+2) - 2, v(r+1) - 1).$$

The relation between these two partitions is

$$\mu_i = s - \lambda_{r+1-i} \qquad (1 \le i \le r)$$

that is to say λ is the complement, say $\hat{\mu}$, of μ in the rectangle (s^r) with r rows and s columns. Hence, replacing each y_j by $-y_j$, we obtain from (5.10) by operating with ∂_u on both sides and using (4.8)

$$\textbf{(5.11)} \qquad \prod_{i=1}^{r} \prod_{j=1}^{s} (x_i + y_j) = \sum s_{\hat{\mu}}(x) s_{\mu'}(y)$$

summed over all $\mu \subset (s^r)$, where $\hat{\mu}$ is the complement of μ in (s^r). This is one version of the usual Cauchy identity [M, Chapter I ,(4.3)$'$].

Let $(\mathfrak{S}^w)_{w \in S_n}$ be the Λ_n-basis of P_n dual to the basis $(\mathfrak{S}_w)$ relative to the scalar product (5.2). By (5.3) and (5.5) we have

$$< \mathfrak{S}_u, w_0 \mathfrak{S}_{vw_0} >= \epsilon(vw_0) \delta_{uv}$$

or equivalently

$$< \mathfrak{S}_u(x), w_0 \mathfrak{S}_{vw_0}(-x) >= \delta_{uv}$$

which shows that

$$\textbf{(5.12)} \qquad \mathfrak{S}^w(x) = w_0 \mathfrak{S}_{ww_0}(-x)$$

for all $w \in S_n$. From (5.10) it follows that

$$\Delta(x, y) = \sum_{w \in S_n} \mathfrak{S}_w(x) w_0 \mathfrak{S}^w(y)$$

or equivalently

$$\textbf{(5.13)} \qquad \prod_{1 \le i < j \le n} (x_i - y_j) = \sum_{w \in S_n} \mathfrak{S}_w(x) \mathfrak{S}^w(y).$$

Let $(x_\beta)_{\beta\subset\delta}$ be the basis dual to $(x^\alpha)_{\alpha\subset\delta}$. If

$$\mathfrak{S}_u = \sum a_{u\alpha} x^\alpha,$$
$$\mathfrak{S}^v = \sum b_{v\beta} x_\beta,$$

then by taking scalar products we have

$$\sum_\alpha a_{u\alpha} b_{v\beta} = \delta_{uv}$$

and therefore also

$$\sum_w a_{w\alpha} b_{w\beta} = \delta_{\alpha\beta},$$

so that

$$\sum_{w\in S_n} \mathfrak{S}_w(x)\mathfrak{S}^w(y) = \sum_{\alpha,\beta}(\sum_w a_{w\alpha}b_{w\beta})x^\alpha y_\beta$$
$$= \sum_\alpha x^\alpha y_\alpha.$$

From (5.13) it follows that y_α is the coefficient of x^α in $\prod_{i<j}(x_i - y_j)$, and hence we find

(5.14) $$x_\alpha = (-1)^{|\beta|}\prod_{i=1}^{n-1} e_{\beta_i}(x_{i+1},\dots,x_n)$$

where $\beta = \delta - \alpha$. ‖

Let

$$C(x,y) = \epsilon(w_0)\Delta(w_0x,y) = \prod_{i<j}(y_i - x_j).$$

If $f(x)\in H_n$ (4.11), let $f(y)$ denote the polynomial in $y_1,\dots,y_n$ obtained by replacing each x_i by y_i. Then we have

(5.15) $$< f(x), C(x,y) >_x = f(y),$$

where as before the suffix x means that the scalar product is taken in the x variables. In other words, $C(x,y)$ is a “reproducing kernel” for the scalar product.

Proof: From (5.13) we have

$$C(x,y) = \sum_{w\in S_n}\epsilon(w_0)\mathfrak{S}_w(w_0x)\mathfrak{S}_{ww_0}(-y).$$

Hence by (5.5)

$$< C(x,y), \mathfrak{S}_{ww_0}(x) >_x = \epsilon(ww_0)\mathfrak{S}_{ww_0}(-y)$$
$$= \mathfrak{S}_{ww_0}(y).$$

Hence (5.15) is true for all Schubert polynomials $\mathfrak{S}_u, u \in S_n$. Since the scalar product is Λ_n-linear it follows from (5.6) that (5.15) is true for all $f \in H_n$. ‖

Let θ_{yx} be the homomorphism that replaces each y_i by x_i. Then (5.15) can be restated in the form

(5.15′) $$\theta_{yx} < f(x), C(x,y) >_x = f(x)$$

for all $f \in H_n$.

Now let $z = [z_1, \ldots, z_n]$ be a third set of variables and consider

(1) $$< C(x,y), \partial_u v^{-1} C(x,z) >_x$$

for $u, v \in S_n$, where ∂_u and v^{-1} act on the x variables. By (5.3) this is equal to

(2) $$\epsilon(v) < C(x,z), v\partial_{u^{-1}} C(x,y) >_x$$

and by (5.15′) we have

(3) $$\theta_{yx} < C(x,y), \partial_u v^{-1} C(x,z) >_x = \partial_u v^{-1} C(x,z),$$

(4) $$\theta_{zx} < C(x,z), v\partial_{u^{-1}} C(x,y) >_x = v\partial_{u^{-1}} C(x,y).$$

Since θ_{yx} and θ_{zx} commute, it follows from (1)-(4) that

$$\begin{aligned} \theta_{yx} v\partial_{u^{-1}} C(x,y) &= \epsilon(v)\theta_{zx}\partial_u v^{-1} C(x,z) \\ &= \epsilon(v)\theta_{yx}\partial_u v^{-1} C(x,y). \end{aligned}$$

Hence we have

(5.16) $$\theta(v\partial_{u^{-1}} w_0 \Delta) = \epsilon(v)\theta(\partial_u v^{-1} w_0 \Delta)$$

for all $u, v \in S_n$, where $\Delta = \Delta(x,y)$ and $\theta = \theta_{yx}$.

Let E_n denote the algebra of operators ϕ of the form

$$\phi = \sum_{w \in S_n} \phi_w w,$$

with coefficients $\phi_w \in Q_n = \mathbf{Q}(x_1, \ldots, x_n)$. For such a ϕ we have

(5.17) $$\phi_w = \epsilon(w_0) a_\delta^{-1} \theta(\phi(w^{-1} w_0 \Delta))$$

for all $w \in S_n$, where ϕ and $w^{-1}w_0$ act on the x variables in Δ.

For $\theta(\phi(w^{-1}w_0\Delta)) = \sum_{u\in S_n} \phi_u\theta(uw^{-1}w_0\Delta)$, and by (5.8) $\theta(uw^{-1}w_0\Delta) = \Delta(uw^{-1}w_0x, x)$ is zero if $u \neq w$, and is equal to $\epsilon(w_0)a_\delta$ if $u = w$.

Let $u \in S_n$, and let $(a_1, \ldots, a_p)$ be a reduced word for u, so that $\partial_u = \partial_{a_1} \cdots \partial_{a_p}$. Since $\partial_a = (x_a - x_{a+1})^{-1}(1 - s_a)$ for each $a \geq 1$, it follows that we may write

(5.18)
$$\partial_u = \epsilon(w_0)a_\delta^{-1} \sum_{v\leq u} \alpha_{uv}v,$$

where $v \leq u$ means that v is of the form $s_{b_1} \ldots s_{b_q}$, where $(b_1, \ldots, b_q)$ is a subword of $(a_1, \ldots, a_p)$.

The coefficients α_{uv} in (5.18) are polynomials, for it follows from (5.16) and (5.17) that

(5.19)
$$\begin{aligned}\alpha_{uv} &= \theta(\partial_u(v^{-1}w_0\Delta)) \\ &= \epsilon(v)\theta(v\partial_{u^{-1}}w_0\Delta).\end{aligned}$$

(5.20) *For all* $f \in P_n$ *we have*

$$\theta(\partial_u(\Delta f)) = \begin{cases} w_0 f & \text{if } u = w_0, \\ 0 & \text{otherwise.} \end{cases}$$

Proof: From (5.18) we have

$$\theta(\partial_u(\Delta f)) = a_\delta^{-1} \sum_{v\leq u} \alpha_{uv}v(f)\theta(v\Delta).$$

By (5.9) this is zero if $u \neq w_0$, and if $u = w_0$ then by (2.10)

$$\begin{aligned}\theta(\partial_{w_0}(\Delta f)) &= a_\delta^{-1} \sum_{w\in S_n} \epsilon(w)w(f)\theta(w\Delta) \\ &= a_\delta^{-1}\epsilon(w_0)w_0(f)\epsilon(w_0)a_\delta = w_0(f)\end{aligned}$$

by (5.9) again. ‖

The matrix of coefficients (α_{uv}) in (5.18) is triangular with respect to the ordering $\leq$, and one sees easily that the diagonal entries α_{uu} are non-zero (they are products in which each factor is of the form $x_i - x_j$). Hence we may invert the equations (5.18), say

(5.21)
$$u = \sum_{v\leq u} \beta_{uv}\partial_v$$

and thus we can express any $\phi \in E_n$ as a linear combination of the operators ∂_w. Explicitly, we have

(5.22)
$$\phi = \sum_{w\in S_n} \theta(\phi(\partial_{w^{-1}w_0}\Delta))\partial_w.$$

Proof: By linearity we may assume that $\phi = f\partial_u$ with $f \in Q_n$. Then

$$\theta(\phi(\partial_{w^{-1}w_0}\Delta)) = f\theta(\partial_u\partial_{w^{-1}w_0}\Delta).$$

Now by (4.2) $\partial_u\partial_{w^{-1}w_0}$ is either zero or equal to $\partial_{uw^{-1}w_0}$, and by (5.20) $\theta(\partial_{uw^{-1}w_0}\Delta)$ is zero if $w \neq u$, and is equal to 1 if $w = u$. Hence the right-hand side of (5.22) is equal to $f\partial_u = \phi$, as required. ‖

In particular, it follows from (5.22) and (5.21) that

(5.23) $$\beta_{uv} = \theta(u\partial_{v^{-1}w_0}\Delta),$$

hence is a polynomial.

The coefficients α_{uv}, β_{uv} in (5.18) and (5.23) satisfy the following relations:

(5.24) (i) $\beta_{uv} = \epsilon(uv)\alpha_{vw_0,uw_0}$,

(ii) $\alpha_{u^{-1}v^{-1}} = v^{-1}(\alpha_{uv})$,

(iii) $\alpha_{\overline{u},\overline{v}} = \epsilon(uw_0)w_0(\alpha_{uv})$,

for all $u, v \in S_n$, *where* $\overline{u} = w_0uw_0, \overline{v} = w_0vw_0$.

Proof: (i) By (5.23) and (2.12) we have

$$\begin{aligned}\beta_{uv} &= \epsilon(v^{-1}w_0)\theta(uw_0\partial_{w_0v^{-1}}w_0\Delta) \\ &= \epsilon(v^{-1}w_0)\epsilon(uw_0)\theta(\partial_{vw_0}w_0u^{-1}w_0\Delta) && \text{by (5.16)} \\ &= \epsilon(uv)\alpha_{vw_0,uw_0}. && \text{by (5.19).}\end{aligned}$$

(ii) From (5.18) we have

$$\begin{aligned}\theta(v\partial_{u^{-1}}w_0\Delta) &= \epsilon(w_0)v(a_\delta^{-1})\sum_w v(\alpha_{u^{-1},w^{-1}})\theta(vw^{-1}w_0\Delta) \\ &= \epsilon(v)v(\alpha_{u^{-1},v^{-1}}) && \text{by (5.9),}\end{aligned}$$

and likewise

$$\begin{aligned}\theta(\partial_u v^{-1}w_0\Delta) &= \frac{\epsilon(w_0)}{a_\delta}\sum_w \alpha_{uw}\theta(wv^{-1}w_0\Delta) \\ &= \alpha_{uv}\end{aligned}$$

again by (5.9). Hence (ii) follows from (5.16).

(iii) Since $\partial_{\overline{u}} = \epsilon(u)w_0\partial_u w_0$ (2.12) we have

$$\begin{aligned}\sum_v \alpha_{\overline{uv}}\overline{v} &= \epsilon(uw_0)w_0(\sum_v \alpha_{uv}v)w_0 \\ &= \epsilon(uw_0)\sum_v w_0(\alpha_{uv})\overline{v}\end{aligned}$$

and hence $\alpha_{\overline{uv}} = \epsilon(uw_0)w_0(\alpha_{uv})$. ‖

(5.25) *Let* E'_n *be the subalgebra of operators* $\phi \in E_n$ *such that* $\phi(P_n) \subset P_n$. *Then* E'_n *is a free* P_n*-module with basis* $(\partial_w)_{w\in S_n}$.

Proof: If $\phi = \sum_{w\in S_n} \phi_w \partial_w \in E'_n$, then by (5.22)

$$\phi_w = \theta(\phi(\partial_{w^{-1}w_0}\Delta)) \in P_n.$$

On the other hand, the ∂_w are a Q_n-basis of E_n, and hence are linearly independent over P_n. ‖

Chapter VI

Double Schubert Polynomials

Let $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$ be two sequences of independent indeterminates, and recall (5.8) that

$$\Delta(x, y) = \prod_{i+j \leq n} (x_i - y_j).$$

For each $w \in S_n$, we define the *double Schubert polynomial* $\mathfrak{S}_w(x, y)$ to be

$$\textbf{(6.1)} \qquad \mathfrak{S}_w(x, y) = \partial_{w^{-1}w_0} \Delta(x, y)$$

where $\partial_{w^{-1}w_0}$ acts on the x variables.

Since $\Delta(x, 0) = x^\delta$ we have

$$\textbf{(6.2)} \qquad \mathfrak{S}_w(x, 0) = \mathfrak{S}_w(x),$$

the (single) Schubert polynomial indexed by w.

From the Cauchy formula (5.10) we have

$$\mathfrak{S}_w(x, y) = \sum_{v \in S_n} \partial_{w^{-1}w_0} \mathfrak{S}_{vw_0}(x) \mathfrak{S}_v(-y)$$

and by (4.2)

$$\partial_{w^{-1}w_0} \mathfrak{S}_{vw_0}(x) = \mathfrak{S}_{vw}(x)$$

if $\ell(vw) = \ell(vw_0) - \ell(w^{-1}w_0)$, i.e. if $\ell(vw) = \ell(w) - \ell(v)$, and

$$\partial_{w^{-1}w_0} \mathfrak{S}_{vw_0}(x) = 0$$

otherwise. Hence

$$\textbf{(6.3)} \qquad \mathfrak{S}_w(x, y) = \sum_{u,v} \mathfrak{S}_u(x) \mathfrak{S}_v(-y)$$

summed over all $u, v \in S_n$ *such that* $w = v^{-1}u$ *and* $\ell(w) = \ell(u) + \ell(v)$.

From (6.3) it follows that $\mathfrak{S}_w(x, y)$ is a homogeneous polynomial of degree $\ell(w)$ in $x_1, \dots, x_{n-1}$, $y_1, \dots, y_{n-1}$. We have

(6.4) (i) $\mathfrak{S}_{w_0}(x, y) = \Delta(x, y)$,

(ii) $\mathfrak{S}_1(x, y) = 1$,

(iii) $\mathfrak{S}_{w^{-1}}(x, y) = \mathfrak{S}_w(-y, -x) = \epsilon(w)\mathfrak{S}_w(y, x)$ *for all* $w \in S_n$,

(iv) $\mathfrak{S}_w(x, x) = 0$ *for all* $w \in S_n$ *except* $w = 1$.

Proof: (i) is immediate from the definition (6.1).

(ii) and (iii) follow from (6.3).

(iv) follows from (5.20), since $\mathfrak{S}_w(x, x) = \theta(\partial_{w^{-1}w_0}\Delta) = 0$ if $w \neq 1$. ‖

(6.5) (Stability) *If* $m > n$ *and* i *is the embedding of* S_n in S_m, *then*

$$\mathfrak{S}_{i(w)}(x, y) = \mathfrak{S}_w(x, y)$$

for all $w \in S_n$.

Proof: This again follows from (6.3) and the stability of the single Schubert polynomials (4.5). ‖

From (6.5) it follows that the double Schubert polynomials $\mathfrak{S}_w(x, y)$ are well defined for all permutations $w \in S_\infty$.

For any commutative ring K, let $K(S_\infty)$ denote the K-module of all functions on S_∞ with values in K. We define a multiplication in $K(S_\infty)$ as follows: for $f, g \in K(S_\infty)$,

$$(fg)(w) = \sum_{u,v} f(u)g(v)$$

summed over all $u, v \in S_\infty$ such that $uv = w$ and $\ell(u) + \ell(v) = \ell(w)$. For this multiplication, $K(S_\infty)$ is an associative (but not commutative) ring, with identity element $\underline{\underline{1}}$, the characteristic function of the identity permutation 1. It carries an involution $f \mapsto f^*$, defined by

$$f^*(w) = f(w^{-1})$$

which satisfies

$$(fg)^* = g^* f^*$$

for all $f, g \in K(S_\infty)$.

(6.6) Let $f, g \in K(S_\infty)$.

(i) *If* $fg = f$ *and* $f(1)$ *is not a zero divisor in* K*, then* $g = \underline{\underline{1}}$.

(ii) *If* $fg = \underline{\underline{1}}$*, then* $gf = \underline{\underline{1}}$.

(iii) f *is a unit (i.e. invertible) in* $K(S_\infty)$ *if and only if* $f(1)$ *is a unit in* K.

Proof: (i) We have $f(1) = f(1)g(1)$ and hence $g(1) = 1$. We shall show by induction on $\ell(w)$ that $g(w) = 0$ for all $w \neq 1$. So let $r > 0$ and assume that $g(v) = 0$ for all $v \in S_\infty$ such that $1 \leq \ell(v) \leq r - 1$. Let w be a permutation of length r. We have

$$f(w) = (fg)(w) = f(w)g(1) + f(1)g(w) + \sum_{u,v} f(u)g(v) \tag{1}$$

where the sum on the right is over $u, v \in S_\infty$ such that $u \neq 1, v \neq 1, uv = w$ and $\ell(u) + \ell(v) = \ell(w)$, so that $1 \leq \ell(v) \leq r - 1$ and therefore $g(v) = 0$. Hence (1) reduces to $f(1)g(w) = 0$ and therefore $g(w) = 0$ as required.

(ii) We have $f(1)g(1) = 1$ so that $f(1)$ is a unit in K. Also $f(gf) = (fg)f = f$, whence $gf = \underline{\underline{1}}$ by (i) above.

(iii) Suppose f is a unit in $K(S_\infty)$, with inverse g. Since $fg = \underline{\underline{1}}$ we have $f(1)g(1) = 1$, whence $f(1)$ is an unit in K.

Conversely, if $f(1)$ is an unit in K we construct an inverse g of f as follows. We define $g(1) = f(1)^{-1}$ and proceed to define $g(w)$ by induction on $\ell(w)$. Assume that $g(v)$ has been defined for all v such that $\ell(v) < \ell(w)$ and set

$$g(w) = -f(1)^{-1} \sum_{u,v} f(u)g(v)$$

summed over u, v such that $uv = w, v \neq w$ and $\ell(u) + \ell(v) = \ell(w)$. This definition gives $(fg)(w) = 0$ as required. ‖

Now let $\mathfrak{S}(x)$ (resp. $\mathfrak{S}(x, y)$) be the function on S_∞ whose value at a permutation w is $\mathfrak{S}_w(x)$ (resp. $\mathfrak{S}_w(x, y)$). (The coefficient ring K is now the ring $\mathbf{Z}[x, y]$ of polynomials in the $x's$ and $y's$.) Since $\mathfrak{S}_1(x) = \mathfrak{S}_1(x, y) = 1$, it follows from (6.6)(iii) that $\mathfrak{S}(x)$ and $\mathfrak{S}(x, y)$ are units in $K(S_\infty)$.

(6.7) (i) $\mathfrak{S}(x, 0) = \mathfrak{S}(x)$,

(ii) $\mathfrak{S}(x, x) = \underline{\underline{1}}$,

(iii) $\mathfrak{S}(x, y)^* = \mathfrak{S}(-y, -x)$,

(iv) $\mathfrak{S}(x)^{-1} = \mathfrak{S}(0, x)$,

(v) $\mathfrak{S}(x)^* = \mathfrak{S}(-x)^{-1}$,

(vi) $\mathfrak{S}(x, y) = \mathfrak{S}(y)^{-1}\mathfrak{S}(x) = \mathfrak{S}(y, x)^{-1}$.

Proof: (i)-(iii) follow directly from (6.2) and (6.4).

From (6.3) and (6.4) we have

$$\mathfrak{S}_w(x,y) = \sum_{u,v} \mathfrak{S}_{u^{-1}}(-y)\mathfrak{S}_v(x) = \sum_{u,v} \mathfrak{S}_u(0,y)\mathfrak{S}_v(x)$$

summed over $u, v \in S_\infty$ such that $uv = w$ and $\ell(u) + \ell(v) = \ell(w)$. In other words,

(**1**) $$\mathfrak{S}(x,y) = \mathfrak{S}(0,y)\mathfrak{S}(x).$$

In particular, when $y = x$ we obtain $\mathfrak{S}(0,x)\mathfrak{S}(x) = \mathfrak{S}(x,x) = \underline{\underline{1}}$ by (ii) above, and hence $\mathfrak{S}(0,x) = \mathfrak{S}(x)^{-1}$. This establishes (iv); part(v) now follows from (iv) and (iii), and (vi) from (iv) and (1) above. ‖

From (6.7) (vi) we have

$$\mathfrak{S}(x) = \mathfrak{S}(y)\mathfrak{S}(x,y)$$

or explicitly

$$\mathfrak{S}_w(x) = \sum_{u,v} \mathfrak{S}_u(y)\mathfrak{S}_v(x,y)$$

summed over u, v such that $uv = w$ and $\ell(u) + \ell(v) = \ell(w)$, so that $u = wv^{-1}$ and $\mathfrak{S}_u = \partial_v \mathfrak{S}_w$ by (4.2). Hence

$$\mathfrak{S}_w(x) = \sum_v \mathfrak{S}_v(x,y)\partial_v \mathfrak{S}_w(y)$$

(where the operators ∂_v act on the y variables). The sum here may be taken over all permutations v, since $\partial_v \mathfrak{S}_w = 0$ unless $\ell(wv^{-1}) = \ell(w) - \ell(v)$. By linearity and (4.13) it follows that

(6.8) (Interpolation Formula) *For all* $f \in P_n = \mathbf{Z}[x_1, \ldots, x_n]$ *we have*

$$f(x) = \sum_w \mathfrak{S}_w(x,y)\partial_w f(y)$$

summed over permutations $w \in S^{(n)}$. ‖

(The reason for the restriction to $S^{(n)}$ in the summation is that if $w \notin S^{(n)}$ we shall have $w(m) > w(m+1)$ for some $m > n$, and hence $\partial_w = \partial_v \partial_m$ where $v = ws_m$; but $\partial_m f = 0$ for all $f \in P_n$, since $m > n$, and therefore $\partial_w f = 0$.)

Remarks. 1. By setting each $y_i = 0$ in (6.8) we regain (4.14).

2. When $n = 1$, the sum is over $S^{(1)}$, which consists of the permutations $w_p = s_p s_{p-1} \ldots s_1$ $(p \geq 0)$; w_p is dominant, of shape (p), so that (see (6.15) below) $\mathfrak{S}_{w_p}(x,y) = (x - y_1) \cdots (x - y_p)$. Hence the case $n = 1$ of (6.8) is *Newton's interpolation formula*

$$f(x) = \sum_{p \geq 0} (x - y_1) \cdots (x - y_p) f_p(y_1, \ldots, y_{p+1})$$

where $f_p = \partial_p \partial_{p-1} \cdots \partial_1 f$, or explicitly

$$f_p(y_1, \ldots, y_{p+1}) = \sum_{i=1}^{p+1} \frac{f(y_i)}{\prod_{j \neq i}(y_i - y_j)}.$$

For any integer r, let $\mathfrak{S}_w(x, r)$ denote the polynomial obtained from $\mathfrak{S}_w(x, y)$ by setting $y_1 = y_2 = \cdots = r$. Since

$$\begin{aligned} \mathfrak{S}_{w_0}(x, r) = \Delta(x, r) &= \prod_{i=1}^{n-1} (x_i - r)^{n-i} \\ &= \mathfrak{S}_{w_0}(x - r) \end{aligned}$$

where $x - r$ means $(x_1 - r, x_2 - r, \ldots)$, it follows from the definitions (6.1) and (4.1) that

$$\mathfrak{S}_w(x, r) = \mathfrak{S}_w(x - r)$$

for all permutations w. Hence, by (6.7)(vi),

$$\mathfrak{S}(x - r) = \mathfrak{S}(r)^{-1} \mathfrak{S}(x)$$

and in particular, for all integers q,

$$\mathfrak{S}(q - r) = \mathfrak{S}(r)^{-1} \mathfrak{S}(q)$$

from which it follows that

(6.9) $$\mathfrak{S}(r) = \mathfrak{S}(1)^r$$

for all $r \in \mathbf{Z}$.

Since $\mathfrak{S}_w(x)$ is a sum of monomials with positive integral coefficients (4.17), $\mathfrak{S}_w(1)$ is the number of monomials in $\mathfrak{S}_w(x)$ (each monomial counted the number of times it occurs). By homogeneity, we have

(6.10) $$\mathfrak{S}_w(r) = r^{\ell(w)} \mathfrak{S}_w(1).$$

From (6.7)(v) and (6.9) we obtain

$$\mathfrak{S}(1)^* = \mathfrak{S}(-1)^{-1} = \mathfrak{S}(1)$$

so that we have another proof of the fact (4.30) that $\mathfrak{S}_w(1) = \mathfrak{S}_{w^{-1}}(1)$.

Now consider the function $F = \mathfrak{S}(1) - \underline{\underline{1}}$, whose value at $w \in S_\infty$ is

$$F(w) = \begin{cases} \text{number of monomials in } \mathfrak{S}_w, & \text{if } w \neq 1, \\ 0, & \text{if } w = 1. \end{cases}$$

For each positive integer p we have

$$F^p = (\mathfrak{S}(1) - \underline{1})^p$$
$$= \sum_{r=0}^{p} (-1)^r \binom{p}{r} \mathfrak{S}(1)^r$$
$$(\mathbf{1}) \qquad = \sum_{r=0}^{p} (-1)^r \binom{p}{r} \mathfrak{S}(r)$$

by (6.9). The value of (1) at a permutation w of length p is by (6.10) equal to

$$\left(\sum_{r=0}^{p} (-1)^r \binom{p}{r} r^p \right) \mathfrak{S}_w(1)$$

which is equal to $p!\mathfrak{S}_w(1)$ (consider the coefficient of t^p in $(e^t - 1)^p$). On the other hand, $F^p(w)$ is by definition equal to

$$(\mathbf{2}) \qquad \sum_{w_1, \ldots, w_p} F(w_1) \cdots F(w_p)$$

summed over all sequences $(w_1, \ldots, w_p)$ of permutations such that $w_1 \ldots w_p = w, \ell(w_1) + \cdots + \ell(w_p) = \ell(w) = p$, and $w_i \neq 1$ for $1 \leq i \leq p$. It follows that each w_i has length 1, hence $w_i = s_{a_i}$ say, and that $(a_1, \ldots, a_p)$ is a reduced word for w. Since

$$\mathfrak{S}_{s_a} = x_1 + \cdots + x_a$$

by (4.4), we have $F(w_i) = \mathfrak{S}_{s_{a_i}}(1) = a_i$, and hence the sum (2) is equal to $\sum a_1 a_2 \cdots a_p$ summed over all $(a_1, \ldots, a_p) \in R(w)$.

We have therefore proved that

(6.11) *The number of monomials in* $\mathfrak{S}_w$ *is*

$$\mathfrak{S}_w(1) = \frac{1}{p!} \sum a_1 a_2 \cdots a_p$$

summed over all $(a_1, \ldots, a_p) \in R(w)$, *where* $p = \ell(w)$. ‖

Remarks. 1. The reduced words for $1_m \times w (m \geq 1)$ are $(m + a_1, \ldots, m + a_p)$ where $(a_1, \ldots, a_p) \in R(w)$. Hence from (6.11) and homogeneity we have

$$\mathfrak{S}_{1_m \times w} \left(\frac{1}{m} \right) = \frac{1}{p!} \sum \left(1 + \frac{a_1}{m} \right) \cdots \left(1 + \frac{a_p}{m} \right)$$

summed over $R(w)$ as before. Letting $m \to \infty$, we deduce that

$$(\mathbf{6.12}) \qquad \operatorname{Card} R(w) = p! \lim_{m \to \infty} \mathfrak{S}_{1_m \times w} \left(\frac{1}{m} \right).$$

2. If w is dominant of length p, then $\mathfrak{S}_w$ is a monomial by (4.7), and hence in this case

$$\sum_{R(w)} a_1 \dots a_p = p!$$

3. Suppose that w is vexillary of length p. Then by (4.9) we have

$$\mathfrak{S}_w = s_\lambda(X_{\phi_1}, \dots, X_{\phi_r})$$

where λ is the shape of w and $\phi = (\phi_1, \dots, \phi_r)$ the flag of w. Hence

$$\mathfrak{S}_{1_m \times w} = s_\lambda(X_{\phi_1+m}, \dots, X_{\phi_r+m})$$

for each $m \geq 1$. If we now set each $x_i = \frac{1}{m}$ and then let $m \to \infty$, we shall obtain in the limit the Schur function s_λ for the series e^t ([M], Ch. I, §3, Ex. 5), which is equal to $h(\lambda)^{-1}$, where $h(\lambda)$ is the product of the hook-lengths of λ. Hence it follows from (6.12) that if w is vexillary of length p, then

(6.13)
$$\text{Card } R(w) = \frac{p!}{h(\lambda)}$$

where λ is the shape of w. In other words, the number of reduced words for a vexillary permutation of length p and shape $\lambda \vdash p$ is equal to the degree of the irreducible representation of S_p indexed by λ.

4. It seems likely that there is a q-analogue of (6.11). Some experimental evidence suggests the following conjecture:

($\mathbf{6.11_q}$?)
$$\mathfrak{S}_w(1, q, q^2, \dots) = \sum q^{\phi(\boldsymbol{a})} \frac{(1-q^{a_1}) \cdots (1-q^{a_p})}{(1-q) \cdots (1-q^p)}$$

summed as in (6.11) over all reduced words $\boldsymbol{a} = (a_1, \dots, a_p)$ for w, where

$$\phi(\boldsymbol{a}) = \sum \{i : a_i < a_{i+1}\}.$$

When w is vexillary the double Schubert polynomial $\mathfrak{S}_w(x,y)$ can be expressed as a multi-Schur function, just as in the case of (single) Schubert polynomials (Chap. IV). We consider first the case of a dominant permutation:

(6.14) *If w is dominant of shape λ, then*

$$\begin{aligned}\mathfrak{S}_w(x,y) &= \prod_{(i,j) \in \lambda} (x_i - y_j) \\ &= s_\lambda(X_1 - Y_{\lambda_1}, \dots, X_m - Y_{\lambda_m})\end{aligned}$$

where $m = \ell(\lambda)$ *and* $X_i = x_1 + \cdots + x_i, Y_i = y_1 + \cdots + y_i$ for all $i \geq 1$.

Proof: As in (4.6) we proceed by descending induction on $\ell(w), w \in S_n$. The result is true for $w = w_0$, since w_0 is dominant of shape δ and

$$\mathfrak{S}_{w_0}(x,y) = \Delta(x,y) = \prod_{(i,j)\in\delta}(x_i - y_j).$$

Suppose $w \neq w_0$ is dominant of shape λ. Then $\lambda \subset \delta$ (and $\lambda \neq \delta$). Let $r \geq 0$ be the largest integer such that $\lambda'_i = n - i$ for $1 \leq i \leq r$, and let $a = \lambda'_{r+1} + 1 \leq n - r - 1$. Then ws_a is dominant, $\ell(ws_a) = \ell(w) + 1$, and $\lambda(ws_a) = \lambda(w) + \epsilon_a$, and therefore

$$\begin{aligned}\mathfrak{S}_w(x,y) &= \partial_a \mathfrak{S}_{ws_a}(x,y) \\ &= \partial_a((x_a - y_{r+1}) \prod_{(i,j)\in\lambda}(x_i - y_j))\end{aligned}$$

by the inductive hypothesis; since $\lambda_a = \lambda_{a+1}$ it follows that

$$\mathfrak{S}_w(x,y) = \prod_{(i,j)\in\lambda}(x_i - y_j)$$

which is equal to $s_\lambda(X_1 - Y_{\lambda_1}, \ldots, X_m - Y_{\lambda_m})$ by (3.5). ∥

(6.15) *If* w *is Grassmannian of shape* λ *then*

$$\mathfrak{S}_w(x,y) = s_\lambda(X_m - Y_{\lambda_1+m-1}, \cdots, X_m - Y_{\lambda_m}).$$

Proof: This follows from (6.14) just as (4.8) follows from (4.7). ∥

Finally, let w be vexillary with shape

$$\lambda(w) = (p_1^{m_1}, \ldots, p_k^{m_k})$$

and flag

$$\phi(w) = (f_1^{m_1}, \ldots, f_k^{m_k})$$

as in Chapter IV. Then w^{-1} is also vexillary, with shape

$$\lambda(w^{-1}) = \lambda(w)' = (q_1^{n_1}, \ldots, q_k^{n_k})$$

the conjugate of $\lambda(w)$, and flag

$$\phi(w^{-1}) = (g_1^{n_1}, \ldots, g_k^{n_k})$$

where by (1.41)

$$g_i + q_i = f_{k+1-i} + p_{k+1-i} \qquad (1 \leq i \leq k).$$

With this notation recalled, we have

(6.16) $$\mathfrak{S}_w(x,y) = s_\lambda((X_{f_1} - Y_{g_k})^{m_1}, \ldots, (X_{f_k} - Y_{g_1})^{m_k}).$$

Proof: The proof is essentially the same as that of (4.9) (which is the case $y = 0$). By (4.10) the dominant permutation w_k constructed from w in the proof of (4.9) has shape

$$\mu = (g_k^{m_1}, g_{k-1}^{m_2}, \ldots, g_1^{m_k})$$

and therefore by (6.15) we have

$$\mathfrak{S}_{w_k}(x,y) = s_\mu(X_1', \ldots, X_m')$$

where $m = m_1 + \cdots + m_k = \ell(\lambda)$ and the sequence $(X_1', \ldots, X_m')$ is obtained by subtracting the sequence $((Y_{g_k})^{m_1}, \ldots, (Y_{g_1})^{m_k})$ term by term from the sequence $(X_1, \ldots, X_m)$. Hence the same argument as in (4.9) establishes (6.17). ‖

Remark. From (6.16) and (6.4)(iii) we obtain

$$s_\lambda(Z_1^{m_1}, \ldots, Z_k^{m_k}) = (-1)^{|\lambda|} s_{\lambda'}((-Z_k)^{n_1}, \ldots, (-Z_1)^{n_k})$$

where $Z_i = X_{f_i} - Y_{g_{k+i-1}}$ so that (if $rk\ (x_i) =\ rk\ (y_i) = 1$ for each $i \geq 1$)

$$\begin{aligned} rk\ (Z_{i+1} - Z_i) &= f_{i+1} - f_i + g_{k+1-i} - g_{k-i} \\ &= m_{i+1} - n_{k+1-i} \end{aligned}$$

by (1.41). Hence (6.4)(iii) reduces to the duality theorm (3.8″) (with $\mu = 0$) when w is vexillary.

Let τ_x (resp. τ_y) be the shift operator (4.21) acting on the x (resp. y) variables. Then we have

(6.17) $$\tau_x^r \tau_y^r \mathfrak{S}_w(x,y) = \mathfrak{S}_{1_r \times w}(x,y)$$

for all $r \geq 1$ and all permutations w.

Proof: By (6.3) and (4.21) we have

$$\tau_x^r \tau_y^r \mathfrak{S}_w(x,y) = \sum_{u,v} \epsilon(v) \mathfrak{S}_{1_r \times u}(x) \mathfrak{S}_{1_r \times v}(y)$$

summed over u, v such that $v^{-1}u = w$ and $\ell(u) + \ell(v) = \ell(w)$. By (6.3) again, the right-hand side is equal to $\mathfrak{S}_{1_r \times w}(x,y)$. ‖

In particular, suppose that w is vexillary. With the notation of (6.16), the flag of $1_r \times w$ (resp. $1_r \times w^{-1}$) is obtained from that of w (resp. w^{-1}) by replacing each f_i by $f_i + r$ (resp. each g_i by $g_i + r$). Hence by (6.16) we have

$$\mathfrak{S}_{1_r \times w}(x, y) = s_\lambda((X_{f_1+r} - Y_{g_k+r})^{m_1}, \ldots, (X_{f_k+r} - Y_{g_1+r})^{m_k})$$

and hence

$$\textbf{(6.18)} \qquad \rho_r^{(x)} \rho_r^{(y)} \mathfrak{S}_{1_r \times w}(x, y) = s_\lambda(X_r - Y_r)$$

for all $r \geq 1$, where $\rho_r^{(x)}$ (resp. $\rho_r^{(y)}$) is the homomorphism ρ_r of (4.25) acting on the x (resp. y) variables.

(6.19) *Let π_x (resp. π_y) denote $\pi_{w_0^{(r)}}$ acting on the x (resp. y) variables. Then if w is vexillary of shape λ, we have*

$$\pi_x \pi_y \mathfrak{S}_w(x, y) = s_\lambda(X_r - Y_r).$$

Proof: By (4.24) we have $\pi_x = \rho_r^{(x)} \tau_x^r$ and $\pi_y = \rho_y^{(r)} \tau_y^r$. Hence (6.19) follows from (6.17) and (6.18). ‖

In particular, suppose that w is dominant of shape λ, so that by (6.14)

$$\mathfrak{S}_w(x, y) = \prod_{(i,j) \in \lambda} (x_i - y_j) = f_\lambda(x, y) \text{ say.}$$

In this case (6.19) gives

$$\pi_x \pi_y f_\lambda(x, y) = s_\lambda(X_r - Y_r)$$

for all $r \geq 1$, which is Sergeev's formula (3.12′).

Chapter VII

Schubert Polynomials (2)

Recall the decomposition (4.17) of a Schubert polynomial $\mathfrak{S}_w$:

$$\mathfrak{S}_w(x_1, x_2, \dots) = \sum_{u,v} d^w_{uv} \mathfrak{S}_u(x_1, \dots, x_m)\mathfrak{S}_v(x_{m+1}, x_{m+2}, \dots)$$

Our first aim in this Chapter will be to give a method for calculating the coefficients d^w_{uv}. We shall then apply our results to the calculation of Card $(R(w))$, the number of reduced decompositions $w = s_{a_1} \cdots s_{a_p}$ (where $p = \ell(w)$) of a permutation w.

For this purpose, we introduce the operators $\partial_i^*,\ i \geq 1$, defined by

(7.1)
$$\partial_i^* \mathfrak{S}_w = \begin{cases} \mathfrak{S}_{s_i w} & \text{if } \ell(s_i w) < \ell(w), \\ 0 & \text{otherwise.} \end{cases}$$

Remarks. 1. If ω is the (linear) involution defined by $\omega(\mathfrak{S}_w) = \mathfrak{S}_{w^{-1}}$ for each permutation w, it follows from (4.2) that $\partial_i^* = \omega \partial_i \omega$. Hence we may define $\partial_w^* = \omega \partial_w \omega$ for any permutation w, and we have $\partial_w^* = \partial_{a_1}^* \cdots \partial_{a_p}^*$ whenever $(a_1, \dots, a_p)$ is a reduced word for w.

2. If $w \in S_n$ we have $\partial_i^* \mathfrak{S}_w = 0$ for all $i > n$, because $\partial_i^* \mathfrak{S}_w = \omega \partial_i \mathfrak{S}_{w^{-1}}$, which is zero because $w^{-1}(i) < w^{-1}(i+1)$.

(7.2) *∂_i^* commutes with ∂_j for all $i, j \geq 1$.*

Proof: We have

$$\partial_i^* \partial_j \mathfrak{S}_w = \begin{cases} \partial_i^* \mathfrak{S}_{w s_j} = \mathfrak{S}_{s_i w s_j} & \text{if } \ell(s_i w s_j) = \ell(w) - 2, \\ 0 & \text{otherwise.} \end{cases}$$

Likewise

$$\partial_j \partial_i^* \mathfrak{S}_w = \begin{cases} \partial_j \mathfrak{S}_{s_i w} = \mathfrak{S}_{s_i w s_j} & \text{if } \ell(s_i w s_j) = \ell(w) - 2, \\ 0 & \text{otherwise.} \end{cases}$$

Hence $\partial_i^* \partial_j - \partial_j \partial_i^*$ vanishes on each Schubert polynomial $\mathfrak{S}_w$, and therefore vanishes identically. $\|$

(7.3) *Let* $w_0 = w_0^{(n)}$ *be the longest element of* S_n. *Then for* $r = 1, 2, \ldots, n-1$ *we have*

$$(1 + t\partial^*_{n-r}) \cdots (1 + t\partial^*_{n-1})\mathfrak{S}_{w_0} = (1 + t\partial_1) \cdots (1 + t\partial_r)\mathfrak{S}_{w_0}$$

as polynomials in $t, x_1, x_2, \ldots$.

Proof: The coefficient of t^p $(1 \le p \le r)$ on the left-hand side is

$$\textbf{(1)} \qquad \sum \partial^*_{a_1} \cdots \partial^*_{a_p} \mathfrak{S}_{w_0}$$

summed over all reduced sequences $(a_1, \ldots, a_p)$ satisfying

$$n - r \le a_1 \le \ldots \le a_p \le n - 1.$$

Let $b_i = n - a_{p+1-i}$ for all $1 \le i \le p$, so that

$$\textbf{(2)} \qquad 1 \le b_1 < \ldots < b_p \le r.$$

Let $w = s_{a_p} \cdots s_{a_1}$, so that $w_0 w w_0 = s_{b_1} \cdots s_{b_p}$. Then

$$\begin{aligned}\partial^*_{a_1} \cdots \partial^*_{a_p} \mathfrak{S}_{w_0} &= \mathfrak{S}_{w^{-1}w_0} = \partial_{w_0 w w_0} \mathfrak{S}_{w_0} \\ &= \partial_{b_1} \cdots \partial_{b_p} \mathfrak{S}_{w_0}.\end{aligned}$$

Hence (1) is equal to

$$\sum \partial_{b_1} \cdots \partial_{b_p} \mathfrak{S}_{w_0}$$

summed over all reduced sequences $(b_1, \ldots, b_p)$ satisfying (2), which is the coefficient of t^p on the right hand side of (7.2). ‖

Next, we have

$$\textbf{(7.4)} \qquad \mathfrak{S}_{1 \times w_0}(t, x_1, \ldots, x_{n-1}) = (1 + t\partial_1) \cdots (1 + t\partial_{n-1})\mathfrak{S}_{w_0}(x_1, \ldots, x_{n-1}).$$

Proof: By (4.22) we have to show that

$$(1 + t\partial_1) \cdots (1 + t\partial_{n-1}) s_\delta(X_1, \ldots, X_{n-1}) = s_\delta(t + X_1, \ldots, t + X_{n-1})$$

where $X_i = x_1 + \cdots + x_i$ for each $i \ge 1$, and $\delta = \delta_n$. For this it is enough to show that

$$\textbf{(1)} \qquad (1 + t\partial_i) s_\delta(X_1, \ldots, X_i, t + X_{i+1}, \ldots, t + X_{n-1}) = s_\delta(X_1, \ldots, X_{i-1}, t + X_i, \ldots, t + X_{n-1})$$

for $i = 1, 2, \ldots, n-1$.

Both sides of (1) are determinants with $n-1$ rows and columns which agree in all rows except the i^{th} row. On the left-hand side, the elements of the i^{th} row are by (3.10)

$$h_k(X_i) + th_{k-1}(X_{i+1})$$

and on the right-hand side they are $h_k(t+X_i)$, where k runs form $n-2i+1$ to $2n-2i-1$ in each case.

Now we have

$$\begin{aligned} h_k(X_i) + th_{k-1}(X_{i+1}) &= h_k(t+X_i) - th_{k-1}(t+X_i) + th_{k-1}(t+X_{i+1}) - t^2h_{k-2}(t+X_{i+1}) \\ &= h_k(t+X_i) - t(t-x_{i+1})h_{k-2}(t+X_{i+1}) \end{aligned}$$

Hence if we add $t(t-x_{i+1})$ times the $(i+1)^{\text{th}}$ row to the i^{th} row in the determinant on the left-hand side, we shall obtain the right-hand side of (1). ‖

For each $r \geq 1$, let

$$\Phi_r(t) = t^r(1+t\partial^*_{r+1})(1+t\partial^*_{r+2})\cdots$$

For each permutation w, we have $(1+t\partial^*_j)\mathfrak{S}_w = \mathfrak{S}_w$ for all sufficiently large j by (7.1), so that $\Phi_r(t)\mathfrak{S}_w$ is a polynomial in t (and $x_1, x_2, \ldots$). With this notation, we have

(7.5) $$\partial_1\partial_2\cdots\partial_{n-r+1}(x_1^n x_2^{n-1}\cdots x_n) = \Phi_{r-1}(x_1)\mathfrak{S}_{w_0^{(n)}}(x_2, x_3, \ldots)$$

Proof: Let $s = n-r+1$ and

$$a = x_2^{s-1}x_3^{s-2}\cdots x_s, \quad b = x_{s+2}^{r-2}x_{s+3}^{r-3}\cdots x_n, \quad c = (x_2\cdots x_{s+1})^{r-1}$$

so that $abc = x_2^{n-1}x_3^{n-2}\cdots x_n$. Hence

$$\begin{aligned} \partial_1\partial_2\cdots\partial_n(x_1^n x_2^{n-1}\cdots x_n) &= x_1^{r-1}bc\partial_1\cdots\partial_s(x_1^s x_2^{s-1}\cdots x_s) \\ &= x_1^{r-1}bc\mathfrak{S}_{1\times w_0^{(s)}}(x_1,\ldots,x_s) && \text{by (4.21)} \\ &= x_1^{r-1}bc(1+x_1\partial_2)\cdots(1+x_1\partial_s)a && \text{by (7.4)} \\ &= x_1^{r-1}(1+x_1\partial_2)\cdots(1+x_1\partial_s)abc \\ &= x_1^{r-1}(1+x_1\partial_2)\cdots(1+x_1\partial_s)\mathfrak{S}_{w_0^{(n)}}(x_1,\ldots,x_n) \\ &= x_1^{r-1}(1+x_1\partial^*_r)\cdots(1+x_1\partial^*_{n-1})\mathfrak{S}_{w_0^{(n)}}(x_2,\ldots,x_n) && \text{by (7.3).} \ \| \end{aligned}$$

Let w be any permutation. If $w(1) = r$, then $s_1\cdots s_{r-1}w(1) = 1$, so that we may write

$$s_1\cdots s_{r-1}w = 1\times w_1$$

where w_1 is defined by

$$w_1(i) = \begin{cases} w(i+1) & \text{if } w(i+1) < r, \\ w(i+1) - 1 & \text{if } w(i+1) > r. \end{cases}$$

If the code of w is $(c_1, c_2, \ldots)$ (so that $c_1 = r - 1$), the code of w_1 is $(c_2, c_3, \ldots)$. With this notation we have

$$\textbf{(7.6)} \qquad \mathfrak{S}_w(x_1, x_2, \ldots) = \Phi_{r-1}(x_1)\mathfrak{S}_{w_1}(x_2, x_3, \ldots)$$

Proof: Suppose that $w \in S_{n+1}$. Then

$$\begin{aligned} w_0^{(n+1)} w &= w_0^{(n+1)} s_{r-1} \cdots s_1 (1 \times w_1) \\ &= s_{n-r+2} \cdots s_n w_0^{(n+1)} (1 \times w_0^{(n)})(1 \times w_0^{(n)} w_1) \\ &= s_{n-r+1} \cdots s_1 (1 \times w_0^{(n)} w_1) \end{aligned}$$

since $w_0^{(n+1)}(1 \times w_0^{(n)}) = s_n s_{n-1} \cdots s_1$. Hence

$$\begin{aligned} \mathfrak{S}_w(x_1, \ldots, x_n) &= \partial_{w^{-1} w_0^{(n+1)}} (x_1^n x_2^{n-1} \cdots x_n) \\ &= \partial_{1 \times w_1^{-1} w_0^{(n)}} \partial_1 \cdots \partial_{n-r+1} (x_1^n \cdots x_n) \\ &= \partial_{1 \times w_1^{-1} w_0^{(n)}} \Phi_{r-1}(x_1) \mathfrak{S}_{w_0^{(n)}}(x_2, x_3, \ldots, x_n) && \text{by (7.5)} \\ &= \Phi_{r-1}(x_1) \partial_{1 \times w_1^{-1} w_0^{(n)}} \mathfrak{S}_{w_0^{(n)}}(x_2, x_3, \ldots, x_n) && \text{by (7.2)} \\ &= \Phi_{r-1}(x_1) \mathfrak{S}_{w_1}(x_2, x_3, \ldots). \ \| \end{aligned}$$

Remark. The right-hand side of (7.6) is a sum of terms of the form $x_1^p \mathfrak{S}_u(x_2, x_3, \ldots)$. By applying (7.6) to each $\mathfrak{S}_u$, and so on, we can decompose $\mathfrak{S}_w$ into a sum of monomials, and thus we have another proof of the fact (4.17) that $\mathfrak{S}_w$ is a polynomial in $x_1, x_2, \ldots$ with positive integer coefficients.

Next, let $m \geq 1$ and assume that the permutation w statisfies

$$w(1) > w(2) > \cdots > w(m).$$

Define a partition $\mu = \mu(w, m)$ of length $\leq m$ by

$$\mu_i = w(i) - (m + 1 - i) \qquad (1 \leq i \leq m).$$

If $w \in S_{m+n}$ we have $\mu_1 \leq n$, hence $\mu \subset (n^m)$.

Also let

$$\Phi_\mu(x_1, \ldots, x_m) = \Phi_{\mu_m}(x_m) \cdots \Phi_{\mu_2}(x_2) \Phi_{\mu_1}(x_1)$$

and let w_m be the permutation whose code is $(c_{m+1}, c_{m+2}, \ldots)$, where $(c_1, c_2, \ldots)$ is the code of w.

With this notation established, we have

$$\mathfrak{S}_w(x) = x^{\delta_m} \Phi_\mu(x_1, \ldots, x_m) \mathfrak{S}_{w_m}(x_{m+1}, x_{m+2}, \ldots). \tag{7.7}$$

Proof: We proceed by induction on m; the case $m = 1$ is (7.6). From (7.6) we have

$$\begin{aligned}\mathfrak{S}_w(x) &= \Phi_{\mu_1+m-1}(x_1)\mathfrak{S}_{w_1}(x_2, x_3, \ldots) \\ &= \sum_u x_1^{\mu_1+m+p-1} \mathfrak{S}_{uw_1}(x_2, x_3, \ldots)\end{aligned}$$

summed over all $u = s_{a_1} \cdots s_{a_p}$, where

$$c_1(w) + 1 = \mu_1 + m \le a_1 < \cdots < a_p$$

and $\ell(uw_1) = \ell(w_1) - p$. The code of uw_1 satisfies $c_i(uw_1) = c_i(w_1)$ for $1 \le i \le m-1$, and hence

$$(uw_1)_{m-1} = s_{a_1-m+1} \cdots s_{a_p-m+1} w_m.$$

It follows that

$$\sum_u x_1^{\mu_1+m+p-1} \mathfrak{S}_{(uw_1)_{m-1}}(x_{m+1}, x_{m+2}, \ldots) = x_1^{m-1} \Phi_{\mu_1}(x_1) \mathfrak{S}_{w_m}(x_{m+1}, x_{m+2}, \ldots)$$

and therefore, by the inductive hypothesis,

$$\begin{aligned}\mathfrak{S}_w(x) &= \sum_u x_1^{\mu_1+m+p-1} x_2^{m-2} \cdots x_{m-1} \Phi_{\mu_m}(x_m) \cdots \Phi_{\mu_2}(x_2) \mathfrak{S}_{(uw_1)_{m-1}}(x_{m+1}, x_{m+2}, \ldots) \\ &= x_1^{m-1} x_2^{m-2} \cdots x_{m-1} \Phi_{\mu_m}(x_m) \cdots \Phi_{\mu_1}(x_1) \mathfrak{S}_{w_m}(x_{m+1}, x_{m+2}, \ldots). \ \|\end{aligned}$$

Finally, for any permutation w, let v be the unique element of S_m such that $wv(1) > \cdots > wv(m)$, and let $\mu = \mu(wv, m)$. We have $\ell(wv) = \ell(w) + \ell(v)$ and $(wv)_m = w_m$, so that by (7.7)

$$\mathfrak{S}_{wv}(x) = x^{\delta_m} \Phi_\mu(x_1, \ldots, x_m) \mathfrak{S}_{w_m}(x_{m+1}, x_{m+2}, \cdots).$$

Hence

$$\begin{aligned}\mathfrak{S}_w(x) &= \partial_v \mathfrak{S}_{wv}(x) \\ &= \partial_v(x^{\delta_m} \Phi_\mu(x_1, \ldots, x_m)) \mathfrak{S}_{w_m}(x_{m+1}, x_{m+2}, \ldots).\end{aligned} \tag{7.8}$$

Now by (4.14), for any polynomial $f \in P_m$, we have

$$f = \sum_{u \in S^{(m)}} \eta(\partial_u f) S_u$$

where $S^{(m)}$ consists of the permutations whose codes have length $\leq m$, and $\eta(\partial_u f)$ is the constant term of the polynomial $\partial_u f$. Applying this to (7.8), we obtain our final result:

(7.9) $$\mathfrak{S}_w(x) = \sum_u \mathfrak{S}_u(x_1, \ldots, x_m)\eta(\partial_{uv}(x^{\delta_m}\Phi_\mu(x_1, \ldots, x_m)))\mathfrak{S}_{w_m}(x_{m+1}, x_{m+2}, \ldots)$$

summed over all $u \in S^{(m)}$ *such that* $\ell(uv) = \ell(u) + \ell(v)$. ‖

For each such u, the constant term $\eta(\partial_{uv}(x^{\delta_m}\Phi(x_1, \ldots, x_m)))$ is a polynomial in the (non-commuting) operators ∂_i^* with integer coefficients. Hence (7.9) gives a decomposition of the Schubert polynomial $\mathfrak{S}_w(x)$ of the form

(7.10) $$\mathfrak{S}_w(x) = \sum_{u,v} d_{uv}^w \mathfrak{S}_u(y)\mathfrak{S}_v(z),$$

where $y = (x_1, \ldots, x_m)$ and $z = (x_{m+1}, x_{m+2}, \ldots)$. If $w \in S^{(m+n)}$, so that $\mathfrak{S}_w(x) \in P_{m+n}$, then $u \in S^{(m)}$ and $v \in S^{(n)}$ in this sum. From (4.18) we know that the coefficients d_{uv}^w in (7.10) are ≥ 0.

In particular, if we apply (7.7) to a permutation of the form $w_0^{(m)} \times w$, we shall obtain

(1) $$\mathfrak{S}_{w_0^{(m)} \times w}(x) = x^{\delta_m}\Phi_0(x_1, \ldots, x_m)\mathfrak{S}_w(x_{m+1}, x_{m+2}, \ldots).$$

On the other hand, by (4.6) we have

(2) $$\mathfrak{S}_{w_0^{(m)} \times w} = \mathfrak{S}_{w_0^{(m)}}\mathfrak{S}_{1_m \times w}$$

and comparison of (1) and (2) gives

(7.12) $$\mathfrak{S}_{1_m \times w}(x) = \Phi_0(x_1, \ldots, x_m)\mathfrak{S}_w(x_{m+1}, x_{m+2}, \ldots).$$

By (4.3), $\mathfrak{S}_{1_m \times w}$ is symmetrical in $x_1, \ldots, x_m$. Hence so is the operator $\Phi_0(x_1, \ldots, x_m)$, and we may therefore write Φ_0 in the form

(7.13) $$\Phi_0(x_1, \ldots, x_m) = \sum_{\lambda, v} \alpha_m(\lambda, v)s_\lambda(x_1, \ldots, x_m)\partial_v^*$$

summed over partitions λ of length $\leq m$ and permutations v, with integral coefficients $\alpha_m(\lambda, v)$. From (7.12) and (7.13) we have

(7.14) $$\mathfrak{S}_{1_m \times w} = \sum_{\lambda, v} \alpha_m(\lambda, v)s_\lambda(x_1, \ldots, x_m)\mathfrak{S}_{vw}(x_{m+1}, x_{m+2}, \ldots)$$

summed over λ of length $\leq m$ and v such that $\ell(vw) = \ell(w) - \ell(v)$. The Schur functions occuring here are precisely the Schubert polynomials $\mathfrak{S}_u$, where u is Grassmannian with descent at m. Hence, by (4.18),

(7.15) *The coefficients* $\alpha_m(\lambda, v)$ *in* (7.13) *are* ≥ 0. ‖

Since $\Phi_0(x_1, \dots, x_m, 0) = \Phi_0(x_1, \dots, x_m)$ and $s_\lambda(x_1, \dots, x_m, 0) = s_\lambda(x_1, \dots, x_m)$ if $\ell(\lambda) \leq m$, it follows from (7.13) that

$$\alpha_{m+1}(\lambda, v) = \alpha_m(\lambda, v) = \alpha(\lambda, v) \text{ say} \tag{7.16}$$

for all partitions λ such that $\ell(\lambda) \leq m$.

We may also calculate the operator $\Phi_0(x_1, \dots, x_m)$ as follows. For each integer $p \geq 1$ and each subset D of $\{1, 2, \dots, p-1\}$ let

$$Q_{D,p}(x_1, \dots, x_m) = \sum x_{u_1} \cdots x_{u_p}$$

summed over all sequences $(u_1, \dots, u_p)$ such that $1 \leq u_1 \leq \cdots \leq u_p \leq m$ and $u_i < u_{i+1}$ whenever $i \in D$. Then $Q_{D,p}(x_1, \dots, x_m)$ is a homogeneous polynomial of degree p, and is zero if $m \leq \mathrm{Card}(D)$.

Now let $\mathbf{a} = (a_1, \dots, a_p)$ be a reduced word, so that $\ell(s_{a_1} \cdots s_{a_p}) = p$. The *descent set* of $\mathbf{a}$ is

$$D(\mathbf{a}) = \{i : a_i > a_{i+1}\}.$$

We now define, for each permutation w,

$$F_w(x_1, \dots, x_m) = \sum_{\boldsymbol{a} \in R(w)} Q_{D(\boldsymbol{a}), \ell(w)}(x_1, \dots, x_m),$$

a homogeneous polynomial of degree $\ell(w)$.

With these definitions we have

$$\Phi_0(x_1, \dots, x_m) = \sum_w F_w(x_1, \dots, x_m) \partial_w^*. \tag{7.17}$$

Proof: Let $\mathbf{a} = (a_1, \dots, a_p)$ be a reduced word. Since

$$\Phi_0(x_i) = (1 + x_i \partial_1^*)(1 + x_i \partial_2^*) \cdots$$

it is clear from the definitions that the coefficient of $\partial_a^* = \partial_{a_1}^* \cdots \partial_{a_p}^*$ in $\Phi_0(x_1, \dots, x_m) = \prod_{i=1}^m \Phi_0(x_i)$ is just $Q_{D(\boldsymbol{a}),p}(x_1, \dots, x_m)$. Hence

$$\begin{aligned} \Phi_0(x_1, \dots, x_m) &= \sum_a Q_{D(\boldsymbol{a}),p}(x_1, \dots, x_m) \partial_{\boldsymbol{a}}^* \\ &= \sum_w F_w(x_1, \dots, x_m) \partial_w^*. \ \| \end{aligned}$$

Comparison of (7.17) and (7.13) now shows that $F_w(x_1, \dots, x_m)$ is a symmetric polynomial in $x_1, \dots, x_m$, and that

(7.18)
$$\begin{aligned} F_w(x_1, \dots, x_m) &= \sum_\lambda \alpha_m(\lambda, w) s_\lambda(x_1, \dots, x_m) \\ &= \rho_m(\mathfrak{S}_{1_m \times w^{-1}}). \end{aligned}$$

The sum in (7.18) is over partitions λ such that $\ell(\lambda) \le m$ and $|\lambda| = \ell(w)$. By (7.16) we have

$$F_w(x_1, \dots, x_m, 0) = F_w(x_1, \dots, x_m)$$

and therefore we have a well defined symmetric function $F_w \in \Lambda$, such that $\rho_m(F_w) = F_w(x_1, \dots, x_m)$ for all $m \ge 0$: namely

(7.19)
$$F_w = \sum_\lambda \alpha(\lambda, w) s_\lambda$$

where the sum is over partitions λ of $\ell(w)$, and $\alpha(\lambda, w) = \alpha_m(\lambda, w)$ for any $m \ge \ell(\lambda)$.

Since the coefficient of $x_1 \cdots x_p$ in $Q_{D,p}(x_1, \dots, x_m)$ is 1 if $m \ge p$, it follows that the coefficient of $x_1 \cdots x_p$ (where $p = \ell(w)$) in $F_w(x_1, \dots, x_m)$ is equal to $\mathrm{Card}(R(w))$ whenever $m \ge \ell(w)$. On the other hand, the coefficient of $x_1 \cdots x_p$ in a Schur function s_λ, where $|\lambda| = p$, is equal to f^λ, the number of standard tableaux of shape λ, or equivalently the degree of the irreducible representation χ^λ of S_p indexed by the partition λ ([M], Ch.I, §7). It follows therefore from (7.19) that

(7.20)
$$\mathrm{Card}\ R(w) = \sum_{|\lambda| = \ell(w)} \alpha(\lambda, w) f^\lambda. \ \|$$

Remark. Since the coefficients $\alpha(\lambda, w)$ are ≥ 0 by (7.15), the number of reduced words for w is always equal to the degree of an (in general reducible) representation of the symmetric group $S_{\ell(w)}$. It is therefore natural to ask whether there is a "natural" action of this symmetric group on the **Z**-span (or perhaps **Q**-span) of the set $R(w)$, with character $\sum_\lambda \alpha(\lambda, w) \chi^\lambda$.

We shall conclude with some properties of the symmetric functions F_w and the coefficients $\alpha(\lambda, w)$.

(7.21) *Let* $u \in S_m, v \in S_n$. *Then*

$$F_{u \times v}(x) = F_u(x) F_v(x).$$

Proof: By (7.18), we have for any N,

$$\begin{aligned} F_{u \times v}(x_1, \dots, x_N) &= \rho_N(S_{1_N \times u^{-1} \times v^{-1}}) \\ &= \rho_N(S_{1_N \times u^{-1}} S_{1_{m+N} \times v^{-1}}) && \text{by (4.6)} \\ &= \rho_N(S_{1_N \times u^{-1}}) \rho_N(\rho_{m+N}(S_{1_{m+N} \times v^{-1}})) \\ &= F_u(x_1, \dots, x_N) F_v(x_1, \dots, x_N). \ \| \end{aligned}$$

(7.22) *Let* $w \in S_n$ *and let* $\overline{w} = w_0 w w_0$, *where* w_0 *is the longest element of* S_n. *Then*

$$F_{w^{-1}} = F_{\overline{w}} = \omega F_w$$

where ω *is the involution that interchanges* s_λ and $s_{\lambda'}$. *In other words*

$$\alpha(\lambda, w^{-1}) = \alpha(\lambda, \overline{w}) = \alpha(\lambda', w)$$

for all partitions λ.

For the proof of (7.22) we require a lemma. If t is a standard tableau of shape λ, the *descent set* $D(t)$ of t is the set of i such that $i+1$ lies in a lower row than i in the tableau t. We have

(7.23) $$s_\lambda = \sum_t Q_{D(t),p}$$

where the sum is over the standard tableaux of shape λ, *and* $p = |\lambda|$.

Proof: In the notation of [M, Ch. I, §5], s_λ is the sum of monomials x^T where T runs through the (column-strict) tableaux of shape λ. Each such tableau T determines a standard tableau t, as follows. If a square in the j^{th} column of the diagram of λ is occupied by the number i, replace i by the pair (i, j). Since T is column-strict the pairs (i, j) so obtained are all distinct. If we now order them lexiographically, (so that (i, j) precedes $\lambda(i', j')$ if and only if either $i < i'$, or $i = i'$ and $j < j'$) and relabel them as $1, 2, \ldots, p$, we have a standard tableau t : say $T \to t$. It follows easily that $\sum_{T \to t} x^T = Q_{D(t),p}$, which proves the lemma. ∥

If D is any subset of $\{1, 2, \ldots, p-1\}$, let $\overline{D}$ denote the complementary subset, and let $D^* = \{p - i : i \in D\}$. From the definition of $Q_{D,p}$ we have

(1) $$Q_{D,p}(x_m, x_{m-1}, \ldots, x_1) = Q_{D^*,p}(x_1, \ldots, x_m).$$

If $\mathbf{a} = (a_1, \ldots, a_p) \in R(w)$, let $\overline{\mathbf{a}} = (n - a_1, \ldots, n - a_p)$ and $\mathbf{a}^* = (n - a_p, \ldots, n - a_1)$. Then we have

(2) $$\overline{\mathbf{a}} \in R(\overline{w}), \qquad \mathbf{a}^* \in R(w^*),$$

where $w^* = (\overline{w})^{-1} = w_0 w^{-1} w_0$. Also

(3) $$D(\overline{\mathbf{a}}) = \overline{D(\mathbf{a})}, \qquad D(\mathbf{a}^*) = D(\mathbf{a})^*.$$

Moreover, it t is a standard tableau we have

(4) $$D(t') = \overline{D(t)}$$

where t' is the transpose of t, obtained by reflecting t in the main diagonal. For $i \in D(t)$ if and only if $i+1$ does *not* lie in a later column than i in the tableau t, that is to say if and only if $i \notin D(t')$.

Since F_w is symmetric, it follows from (1),(2), and (3) that

$$F_w(x_1, \ldots, x_m) = F_w(x_m, \ldots, x_1) = F_{w^*}(x_1, \ldots, x_m)$$

and hence by (7.16) that $F_w = F_{w^*}$.

From (7.23) and (4) above we have

$$\omega s_\lambda = s_{\lambda'} = \sum_{t \in St(\lambda)} Q_{\overline{D(t)},p}$$

for all partitions λ of p, where $St(\lambda)$ is the set of standard tableaux of shape λ, and hence it follows from (2) and (3) and the definition of F_w that $\omega F_w = F_{\overline{w}}$. Hence

$$\omega F_{w^{-1}} = F_{w^*} = F_w,$$

which completes the proof of (7.22). ‖

(7.24) (i) $\alpha(\mu, w) = 0$ *unless* $\lambda(w^{-1}) \leq \mu \leq \lambda(w)'$.

(ii) $\alpha(\mu, w) = 1$ *if* $\mu = \lambda(w^{-1})$ *or* $\mu = \lambda(w)'$.

(iii) *w is vexillary if and only if F_w is a Schur function.*

Proof: (i) Suppose $\alpha(\mu, w) \neq 0$. Then the monomial x^μ occurs in F_w, and hence there is a reduced word $(a_1, \ldots, a_p)$ for w such that

(**1**) $$a_1 < \cdots < a_{\mu_1}, \quad a_{\mu_1+1} < \cdots < a_{\mu_1+\mu_2}, \cdots.$$

By (1.14) the code of w is

(**2**) $$c(w) = \sum_{i=1}^{p} s_{a_p} \cdots s_{a_{i+1}}(\epsilon_{a_i}).$$

If $w^{(1)} = s_{a_p} \cdots s_{a_{\mu_1}+1}$, the sum of the first μ_1 terms of this series is

$$w^{(1)}(\epsilon_{a_{\mu_1}} + s_{a_{\mu_1}}(\epsilon_{a_{\mu_1-1}}) + \cdots + s_{a_{\mu_1}} \cdots s_{a_2}(\epsilon_{a_1})),$$

and since $a_1 < \ldots < a_{\mu_1}$ this is equal to

(**3**) $$w^{(1)}(\epsilon_{a_{\mu_1}} + \epsilon_{a_{\mu_1-1}} + \cdots + \epsilon_{a_1}) = V_1 \text{ say,}$$

where V_1 is a (0,1) vector (i.e., a vector with each component 0 or 1) of weight μ_1. Likewise the sum of the next block of μ_2 terms of the series (2) is a (0,1) vector V_2 of weight μ_2, and so on. Hence

$$c(w) = V_1 + \cdots + V_m$$

where $m = \ell(\lambda)$, and each V_i is a (0,1) vector of weight μ_i. Let V be the (0,1) matrix whose i^{th} row is V_i, for $i = 1, 2, \ldots, m$. Then V has row sums $\mu_1, \ldots, \mu_m$ and column sums $c_1(w), c_2(w), \ldots$ As in the proof of (1.26) it follows that $\mu \leq \lambda(w)'$. Since $\alpha(\mu, w) = \alpha(\mu', w^{-1})$ by (7.23), the same argument applied to μ' and w^{-1} gives $\mu' \leq \lambda(w^{-1})'$ i.e., $\lambda(w^{-1}) \leq \mu$.

(ii) Suppose now that $\mu = \lambda(w)'$. Then there is only one (0,1) matrix V with row sums μ_i and column sums c_i. Its first row V_1 is $\sum \epsilon_j$ summed over j such that $c_j \neq 0$, i.e. such that there exists $k > j$ with $w(k) < w(j)$. From (3) it follows that

$$wV_1 = \sum_{i=1}^{\mu_1} \epsilon_{a_i+1}$$

and therefore $a_1 + 1, \ldots, a_{\mu_1} + 1$ are the terms of the sequence w that have a smaller element somewhere to the right, in increasing order of magnitude. Hence a_1 has no smaller elements to the right of it, and therefore lies to the right of $a_1 + 1$, so that $\ell(s_{a_1} w) = \ell(w) - 1$. The same argument shows that $\ell(s_{a_2} s_{a_1} w) = \ell(s_{a_1} w) - 1$ and so on. Hence if $w_1 = s_{a_{\mu_1}} \cdots s_{a_1} w$ we have $\ell(w_1) = \ell(w) - \mu_1$, and $\lambda(w_1') = (\mu_2, \mu_3, \ldots)$. It follows by induction on $\ell(\mu)$ that the word $(a_1, \ldots, a_p)$ determined by the matrix V is reduced, and hence $\alpha(\mu, w) = 1$ when $\mu = \lambda(w)'$. By (7.23) it follows that $\alpha(\mu, w) = 1$ when $\mu = \lambda(w^{-1})$.

(iii) This follows immediately from (i) and (ii), and the characterization (1.27) of vexillary permutations. ‖

Appendix

Schubert varieties

Let V be a vector space of dimension n over a field K, and let $(e_1, \ldots, e_n)$ be a basis of V, fixed once and for all. A *flag* in V is a sequence $\mathbf{U} = (U_i)_{0 \leq i \leq n}$ of subspaces of V such that

$$0 = U_0 \subset U_1 \subset \cdots \subset U_n = V$$

with strict inclusions at each stage, so that dim $U_i = i$ for each i. In particular, if V_i is the subspace of V spanned by $e_1, \ldots, e_i$, then $\mathbf{V} = (V_i)_{0 \leq i \leq n}$ is a flag in V, called the *standard* flag.

The set $F = F(V)$ of flags in V is called the *flag manifold* of V.

Let G be the group of all automorphisims of the vector space V. Since we have fixed a basis of V, we may identify G with the general linear group $GL_n(k)$: if $g \in G$ and

$$ge_j = \sum_{i=1}^{n} g_{ij} e_i \qquad (1 \leq j \leq n)$$

then g is identified with the matrix (g_{ij}).

The group G acts on F: if $\mathbf{U} = (U_i)$ and $g \in G$, then $g\mathbf{U}$ is the flag (gU_i). Let B be the subgroup of G that fixes the standard flag $\mathbf{V}$. Then $g \in B$ if and only if ge_j is a linear combination of $e_1, \ldots, e_j$, for $1 \leq j \leq n$, that is to say if and only if $g_{ij} = 0$ whenever $i > j$, so that B is the group of upper triangular matrices in $GL_n(k)$.

A *basis* of a flag $\mathbf{U} = (U_i)$ is a sequence $(u_1, \ldots, u_n)$ in V such that $u_i \in U_i - U_{i-1}$ for $1 \leq i \leq n$, or equivalently such that $u_1, \ldots, u_i$ is a basis of U_i for each i. Given such a basis of $\mathbf{U}$, there is a unique $g \in G$ such that $ge_i = u_i$ for each i, and we have $\mathbf{U} = g\mathbf{V}$. Hence G acts transitively on the flag manifold F, and the mapping $g\mathbf{V} \mapsto gB$ is a bijection of F onto the coset space G/B.

For a flag $\mathbf{U} = (U_i)$, let

$$E_i = E_i(\mathbf{U}) = \{j : 1 \leq j \leq n \text{ and } U_i \cap V_j \neq U_i \cap V_{j-1}\}$$

for $0 \le i \le n$. Then $(E_0, \dots, E_n)$ is a 'flag of sets', i.e. we have

(A.1) (i) $\mathrm{Card}(E_i) = i$ *for* $0 \le i \le n$,

(ii)$E_{i-1} \subset E_i$ *for* $1 \le i \le n$.

Proof: (i) Fix i and let $d_j = \dim\,(U_i \cap V_j)$. Since

$$\frac{U_i \cap V_j}{U_i \cap V_{j-1}} = \frac{U_i \cap V_j}{(U_i \cap V_j) \cap V_{j-1}} \cong \frac{(U_i \cap V_j) + V_{j-1}}{V_{j-1}} \subset \frac{V_j}{V_{j-1}}$$

it follows that $d_j - d_{j-1} = 0$ or 1. Since $d_0 = 0$ and $d_n = i$, there are therefore i jumps in the sequence $(d_0, d_1, \dots, d_n)$, which proves (i).

(ii) Suppose that $j \notin E_i$, so that $U_i \cap V_j = U_i \cap V_{j-1}$. Intersecting with U_{i-1}, we see that $j \notin E_{i-1}$. Hence $E_{i-1} \subset E_i$. $\|$

From (A.1) it follows that that each $\mathbf{U} \in F$ determines a permutation $w \in S_n$ as follows : $w(i)$ is the unique element of $E_i - E_{i-1}$, for $i = 1, 2, \dots, n$. Let $\phi : F \to S_n$ denote the mapping so defined.

The symmetric group acts on V by permuting the basis elements e_i:

$$w(e_i) = e_{w(i)}$$

for $w \in S_n$ and $1 \le i \le n$. Hence we may regard S_n as a subgroup of G.

(A.2) *Let* $\mathbf{U} \in F, w \in S_n$. *Then* $\phi(\mathbf{U}) = w$ *if and only if* $\mathbf{U} = bw\mathbf{V}$ *for some* $b \in B$.

Proof: Suppose $\phi(\mathbf{U}) = w$. Then for $i = 1, \dots, n$ we have

$$U_i \cap V_{w(i)} \supset U_i \cap V_{w(i)-1} \tag{1}$$

and

$$U_{i-1} \cap V_{w(i)} = U_{i-1} \cap V_{w(i)-1} \tag{2}$$

By virtue of (1) we can choose $u_i \in U_i$ of the form

$$u_i = e_{w(i)} + \text{lower terms} \tag{3}$$

where by 'lower terms' is meant a linear combination of $e_1, \dots, e_{w(i)-1}$; and $u_i \notin U_{i-1}$ by virtue of (2).

By rewriting (3) in the form

$$u_{w^{-1}(j)} = e_j + \text{lower terms} \qquad (1 \le j \le n)$$

we see that there exists $b \in B$ such that $u_{w^{-1}(j)} = be_j$ for all j, or equivalently

$$u_i = be_{w(i)} = bwe_i.$$

Hence $\mathbf{U} = bw\mathbf{V}$ as required.

For the converse it is enough to show that (i) $\phi(w\mathbf{V}) = w$ and (ii) $\phi(b\mathbf{U}) = \phi(\mathbf{U})$ for all $b \in B$ and $\mathbf{U} \in F$. As to (i), $wV_i \cap V_j$ is spanned by the basis vectors $e_{w(k)}$ such that $k \leq i$ and $w(k) \leq j$, and therefore $wV_i \cap V_j \neq wV_i \cap V_{j-1}$ if and only if $j = w(k)$ for some $k \leq i$. Thus the set $E_i(w\mathbf{V})$ consists of $w(1), \ldots, w(i)$, which establishes (i). Finally as to (ii), we have $bU_i \cap V_j = b(U_i \cap V_j)$ if $b \in B$, so that $E_i(b\mathbf{U}) = E_i(\mathbf{U})$ and hence $\phi(b\mathbf{V}) = \phi(\mathbf{U})$ as required. ‖

From (A.2) we have immediately

(A.3) (Bruhat decomposition) *G is the disjoint union of the double cosets BwB, $w \in S_n$.* ‖

For each $w \in S_n$, let

$$C_w = (BwB)/B \subset G/B = F.$$

The subsets C_w are the *Schubert cells* in the flag manifold F. By (A.3), F is the disjoint union of the C_w.

Let $\mathbf{U} \in F$. Then $\mathbf{U} \in C_w$ if and only if $\mathbf{U}$ has a basis $(u_1, \ldots, u_n)$ such that $u_i \in V_{w(i)} - V_{w(i)-1}$ for each i. We may normalize the u_i by taking

$$u_i = e_{w(i)} + \text{lower terms}.$$

We can then subtract from u_i suitable multiples of the u_k for which $k < i$ and $w(k) < w(i)$, so as to make the coefficient of $e_{w(k)}$ in u_i zero for each such k. Then u_i is replaced by a vector of the form

$$e_{w(i)} + \sum_j a_{ij} e_j$$

where the sum is over $j < w(i)$ such that $j \neq w(k)$ for any $k < i$, i.e., such that $j < w(i)$ and $w^{-1}(j) > i$, or equivalently $(i,j) \in D(w)$, the diagram of w.

(A.4) *Let $\mathbf{U} \in F$. Then $\mathbf{U} \in C_w$ if and only if $\mathbf{U}$ has a basis $(u_1, \ldots, u_n)$ of the form*

$$u_i = e_{w(i)} + \sum_j a_{ij} e_j$$

where the sum is over all j in the i^{th} row of the diagram of w, and the coefficients a_{ij} are arbitrary elements of the field K. Moreover, the a_{ij} are uniquely determined by the flag $\mathbf{U}$, and the mapping $C_w \to K^{D(w)}$ so defined is a bijection.

Proof: Clearly each "matrix" $a = (a_{ij})$ of shape $D(w)$ determines a basis $(u_1, \dots, u_n)$ of V as above, and hence a flag $\mathbf{U} \in C_w$. If $a^* = (a^*_{ij})$ determines $(u^*_1, \dots, u^*_n)$ and the same flag $\mathbf{U}$, then each u^*_i must be expressible as

$$u^*_i = u_i + \sum_{j<i} c_{ij} u_j,$$

and from the form of u^*_i and the u_j it follows that $u^*_i = u_i$ for each i, and hence $a^* = a$. $\|$

Since Card $D(w) = \ell(w)$ it follows from (A.4) that the Schubert cell C_w is isomorphic to affine space of dimension $\ell(w)$.

Let $\mathbf{U} \in F$ and let $(u_1, \dots, u_n)$ be any basis of $\mathbf{U}$. Since $u_1, \dots, u_i$ is a basis for U_i for each $i = 1, \dots, n-1$, the flag $\mathbf{U}$ determines each of the exterior products $u_1 \wedge \cdots \wedge u_i \in \Lambda^i(V)$ up to a nonzero scalar multiple, and hence $\mathbf{U}$ determines the vector

$$(\mathbf{1}) \qquad u_1 \otimes (u_1 \wedge u_2) \otimes \cdots \otimes (u_1 \wedge \cdots \wedge u_{n-1}) \in E$$

up to a nonzero scalar multiple, where $E = V \otimes \Lambda^2 V \otimes \cdots \otimes \Lambda^{n-1} V$. If $P(E)$ denotes the projective space of E (i.e. the space whose points are the lines in E), we have an injective mapping

$$\pi : F \mapsto P(E)$$

(the *Plücker embedding*) for which $\pi(\mathbf{U})$ is the line in E generated by the vector (1).

Assume from now on that the field K is the field of complex numbers. Then the embedding π realizes the flag manifold F as a complex projective algebraic variety, which is smooth because F has a transitive group of automorphisms (namely G). Each Schubert cell C_w is a locally closed subvariety of F, isomorphic to affine space of dimension $\ell(w)$.

For each $w \in S_n$ let

$$X_w = \overline{C}_w$$

be the closure of C_w in F. The X_w are the *Schubert varieties* in F, and a flag $\mathbf{U}$ lies in X_w if and only if $\mathbf{U}$ has a basis $(u_1, \dots, u_n)$ such that $u_i \in V_{w(i)}$ for each i. Each X_w is in fact a union of Schubert cells C_v : if $(a_1, \dots, a_p)$ is a reduced word for w, then $C_v \subset X_w$ if and only if v is of the form $s_{b_1} \cdots s_{b_q}$ where $(b_1, \dots, b_q)$ is a subsequence of $(a_1, \dots, a_p)$, that is to say if and only if $v \leq w$ in the Bruhat order. In particular, $X_1 = C_1$ is the single point $\mathbf{V} \in F$. At the other extreme, if w_0 is the longest element of S_n, then X_{w_0} is the whole of F, and the dimension of F is $\ell(w_0) = \frac{1}{2}n(n-1)$.

Let $H^*(F; \mathbf{Z})$ be the cohomology ring (with integral coefficients) of the projective variety F. Each closed subvariety X of F determines an element $[X] \in H^*(F; \mathbf{Z})$, and cup-product in $H^*(F; \mathbf{Z})$

corresponds, roughly speaking, to intersection of subvarieties. In particular, for each $w \in S_n$, we have a cohomology class $[X_w] \in H^*(F; \mathbf{Z})$, and it is a consequence of the cell decomposition (A.3) of F that the $[X_w]$ form a $\mathbf{Z}$-basis of $H^*(F; \mathbf{Z})$. In particular, $[X_{w_0}]$ is the identity element.

The connection between the classes $[X_w]$ and the Schubert polynomials $\mathfrak{S}_w (w \in S_n)$ is given by

(A.5) *There is a surjective ring homomorphism*

$$\alpha : \mathbf{Z}[x_1, \ldots, x_n] \to H^*(F; \mathbf{Z})$$

such that

$$\alpha(\mathfrak{S}_w) = [X_{w_0 w}]$$

for each $w \in S_n$.

Proof: Let us temporarily write

$$\sigma_w = [X_{w_0 w}]$$

for $w \in S_n$. Monk [Mo] proved that for all $w \in S_n$ and $r = 1, \ldots, n-1$

$$\textbf{(1)} \qquad \sigma_w \cdot \sigma_{s_r} = \sum_t \sigma_{wt}$$

where the sum on the right hand side is over all transpositions $t = t_{ij}$ such that $i \le r < j \le n$ and $\ell(wt) = \ell(w) + 1$, as in (4.15″).

Define $\xi, \ldots, \xi_n \in H^*(F; \mathbf{Z})$ by

$$\begin{aligned} \xi_1 &= \sigma_1 \\ \xi_i &= \sigma_i - \sigma_{i-1} \qquad (2 \le i \le n-1) \\ \xi_n &= -\sigma_{n-1} \end{aligned}$$

From (1) we deduce the counterpart of (4.16): if r is the last descent of w (so that $r \le n-1$), then we have

$$\textbf{(2)} \qquad \sigma_w = \sigma_v \xi_r + \sum_{w'} \sigma_{w'}$$

where v, w' are as in (4.16). Now iteration of (4.16) will ultimately express $\mathfrak{S}_w$ as a sum of monomials, i.e. as a polynomial in $x_1, \ldots, x_{n-1}$; and iteration of (2) will express σ_w as the *same* polynomial in $\xi_1, \ldots, \xi_{n-1}$. Hence if we define $\alpha : P_n \mapsto H^*(F; \mathbf{Z})$ by $\alpha(x_i) = \xi_i$ $(1 \le i \le n)$, we have $\sigma_w = \alpha(\mathfrak{S}_w)$ for all $w \in S_n$, and the proof of (A.5) is complete. $\|$

In fact the kernel of the homomorphism α is generated by the elementary symmetric functions $e_1, \dots, e_n$ of the x's.

We shall draw one consequence of (A.5) that we have not succeeded in deriving directly from the definition (4.1) of the Schubert polynomials. Since the $\sigma_w, w \in S_n$, form a $\mathbf{Z}$-basis of $H^*(F;\mathbf{Z})$, any product $\sigma_u\sigma_v(u, v \in S_n)$ is uniquely a linear combination of the σ_w, and it follows from intersection theory on F that the coefficient of σ_w in $\sigma_u\sigma_v$ is a non-negative integer. From this we deduce

(A.6) *Let u, v be permutations, and write $\mathfrak{S}_u\mathfrak{S}_v$ as an integral linear combination of the $\mathfrak{S}_w$, say*

$$\mathfrak{S}_u\mathfrak{S}_v = \sum_w c_{uv}^w \mathfrak{S}_w. \tag{1}$$

Then the coefficients c_{uv}^w are non-negative.

We have only to choose n sufficiently large so that u, v and all the permutations w such that $c_{uv}^w \neq 0$ lie in S_n, and then apply the homomorphism α of (A.5).

Remark. The coefficients c_{uv}^w in (A.6) are zero unless

(a) $\ell(w) = \ell(u) + \ell(v)$,

(b) $u \leq w$ and $v \leq w$.

For $\mathfrak{S}_u\mathfrak{S}_v$ is homogeneous of degree $\ell(u) + \ell(v)$, which gives condition (a). Also we have

$$\begin{aligned} c_{uv}^w &= \partial_w(\mathfrak{S}_u\mathfrak{S}_v) \\ &= \sum_{v_1 \leq w} v_1\partial_{w/v_1}(\mathfrak{S}_u)\partial_{v_1}(\mathfrak{S}_v) \end{aligned}$$

by (2.17), and the only possible nonzero term in this sum is that corresponding to $v_1 = v$. Hence if $c_{uv}^w \neq 0$ we must have $v \leq w$, and by symmetry also $u \leq w$.

Notes and References

Chapter I. The notion of the *diagram* of a permutation w is ascribed to J. Riguet in [LS1]. The *code* of w is the Lehmer code, familiar to computer scientists. Vexillary permutations were introduced in [LS1] and enumerated in [LS4], though from a somewhat different point of view from that in the text.

Chapter II. Divided differences, in the context of an arbitrary root system, were introduced independently by Bernstein, Gelfand and Gelfand [BGG] and Demazure [D]. Both these papers establish (2.5), (2.10) and (2.13) in this more general context.

Chapter III. Multi-Schur functions were introduced, and the duality theorm (3.8) proved, by Lascoux [L1]. The proof of Sergeev's formula (3.12) is also due to Lascoux (private communication).

Chapter IV. Schubert polynomials, like divided differences, are defined in the context of an arbitrary root system in [BGG] and in [D]. What is special to the root systems of type A is the stability property (4.5), which ensures that the Schubert polynomial $\mathfrak{S}_w$ is well-defined for all permutations $w \in S_\infty$. Propositions (4.7), (4.8) and (4.9) are stated without proof in various places in [LS1]-[LS7] but as far as I am aware the only published proof of (4.9) is that of M. Wachs [W], which is different from the proof in the text. Proposition (4.15), appropriately modified, is valid for any root system, and in this more general form will be found in [BGG] and [D].

Chapter V. The scalar product (5.2) is introduced in [LS7]. The symmetry properties (5.23) of the coefficient matrices $(\alpha_{uv}), (\beta_{uv})$ are indicated in [LS6].

Chapter VI. Double Schubert polynomials were introduced in [L2]. For the interpolation forumla (6.8), see [LS5]. The generalization (6.20) of Sergeev's formula (3.12) is due to Lascoux (private communication).

Chapter VII. This chapter is mostly an amplification of [LS2]. Propositions (7.21)-(7.24) are due to Stanley [S].

[BGG] I.N. Bernstein, I.M. Gelfand, S.I. Gelfand, “Schubert cells and cohomology of the spaces G/P”, *Russian Math. Surveys* **28** (1973), 1–26.

[D] M. Demazure,“ Désingularization des variétés de Schubert généralisées”, *Ann. Sc. E.N.S.* **(4) 7** (1974) 53–88.

[K] A. Kohnert, Thesis, Bayreuth (1990).

[KP] W. Kráskiewicz and P. Pragacz, “Schubert functors and Schubert polynomials”, *preprint* (1986).

[L1] A. Lascoux, “Puissances extérieures, déterminants et cycles de Schubert”, *Bull. Soc. Math. France* **102** (1974) 161–179.

[L2] A. Lascoux, “Classes de Chern des variétés de drapeaux”, *C.R. Acad. Sci. Paris* **295** (1982) 393–398.

[L3] A. Lascoux, “Anneau de Grothendieck de la variété des drapeaux”, *preprint* (1988).

[LS1] A. Lascoux and M.-P. Schützenberger, “Polynômes de Schubert”, *C.R. Acad. Sci. Paris* **294** (1982) 447–450.

[LS2] A. Lascoux and M.-P. Schützenberger, “Structure de Hopf de l’anneau de cohomologie et de l’anneau de Grothendieck d’une variété de drapeaux”, *C.R. Acad. Sci. Paris* **295** (1982) 629–633.

[LS3] A. Lascoux and M.-P. Schützenberger, “Symmetry and flag manifolds”, *Springer Lecture Notes* **996** (1983) 118-144.

[LS4] A. Lascoux and M.-P. Schützenberger, “Schubert polynomials and the Littlewood-Richardson rule”, *Letters in Math. Physics* **10** (1985) 111–124.

[LS5] A. Lascoux and M.-P. Schützenberger, “Interpolation de Newton à plusieurs variables”, Sém. d’Algèbre M.P. Malliavin 1983-84, *Springer Lecture Notes* **1146** (1985) 161-175.

[LS6] A. Lascoux and M.-P. Schützenberger, “Symmetrization operators on polynomial rings”, *Funkt. Anal.* **21** (1987) 77–78.

[LS7] A. Lascoux and M.-P. Schützenberger, “Schubert and Grothendieck polynomials”, *preprint* (1988).

[LS8] A. Lascoux and M.-P. Schützenberger, “Tableaux and non-commutative Schubert polynomials”, *to appear in Funkt. Anal.* (1989).

[M] I.G. Macdonald, "Symmetric functions and Hall polynomials," *O.U.P.* (1979).

[Mo] D. Monk, "The geometry of flag manifolds", *Proc. London Math. Soc.* **(3) 9** (1959) 253–286.

[S] R.P. Stanley, "On the number of reduced decompositions of elements of Coxeter groups", *Eur. J. Comb.* **5** (1984) 359–372.

[W] M.L. Wachs, "Flagged Schur functions, Schubert polynomials and symmetrizing operators", *J. Comb. Theory (A)* **40** (1985) 276–289.