Notes on Schubert Polynomials

I. G. Macdonald
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 [LSB].

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 thank the audience, especially Adriano Garsia and Jeff Remmel, for their support.

San Diego, 1991

I.C. Macdonald
Contents

I Permutations ................................................................. 1
II Divided differences ...................................................... 23
III Multi-Schur functions ................................................ 33
IV Schubert polynomials (1) ............................................. 43
      Appendix to Chapter IV (by Nantel Bergeron) .............. 65
V Orthogonality ............................................................ 72
VI Double Schubert polynomials ...................................... 85
VII Schubert polynomials (2) .......................................... 95
      Appendix: Schubert varieties .................................... 107
     Notes and References ............................................. 113
     Bibliography ......................................................... 115
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 $(3214)$ 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

$$s_i^2 = 1,$$

$$s_is_j = s_j s_i \text{ if } |i - j| > 1,$$

$$s_is_j s_i = s_is_j s_i \text{ if } 1 \leq i < j < n.$$

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 values of the square $\Sigma_n = [1, n] \times [1, n]$, and we shall adopt the convention of matrices, that in $\Sigma_n$ 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 $\Sigma_n : (s \times v)(i, j) = (s(i), v(j))$. 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 $w_r$ is the permutation

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

and it is clear that

$$I(w_r) = \begin{cases} s_r I(w) \cup \{(r, r + 1)\} & \text{if } w(r) < w(r + 1), \\ s_r I(w) \setminus \{(r + 1, r)\} & \text{if } w(r) > w(r + 1). \end{cases}$$
Proof: We shall show by induction on $f(u)$ that each $w \in S_n$ is a product of $s_i$'s. If $f(u) = 0$, then $w = 1$ and there is nothing to prove. If $f(u) > 0$ then $w(v) > v(r + 1)$ for some $r$, and hence $f(ww_1) = f(w) - 1$ by (1.3). Hence $w = s_{s_1} \ldots s_{s_r}$ say, and therefore (as $n^2 = 1$) $w = s_{s_1} \ldots s_{s_r}$.

For each $w \in S_n$, the length of $w$ is the minimal length of a sequence $(s_{e_1}, \ldots, s_{e_n})$ such that $w = s_{e_1} \ldots s_{e_n}$.

The length of $u \in S_n$ is equal to $f(u) = \text{Card } u$.

Proof: Let $f(u)$ temporarily denote the length of $u$. The proof of (1.4) shows that $w$ can be written as a word of length $f(u)$ in the $s_i$, so that $f(u) \leq f(u)$. Conversely, let $w = s_{e_1} \ldots s_{e_n}$ be any expression of $w$ as a product of $s_i$. To show that $f(u) \leq f(u)$ it is enough to show that $f(u) \leq p$. Let $u = s_{e_1} \ldots s_{e_n}$, from (1.3) we have $f(u) \leq f(u) + 1$ and hence

$$f(u) \leq p = 0.$$  

Hence the proof is completed by induction on $p$.

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

(i) $f(w) = 0$ if and only if $w = 1$.

(ii) $f(w) = 1$ if and only if $w = s_{e_1} \ldots s_{e_n}$ ($1 \leq r < n - 1$).

(iii) $f(w - 1) = f(u)$.

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

$$f(u) = f(u) = \frac{1}{2}(n(n - 1) - f(u)).$$

Proof: (i), (ii) require no comment. Also (iii) is clear, since $w = s_{e_1} \ldots s_{e_n}$ if and only if $w^{-1} = s_{e_1} \ldots s_{e_n}$.

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

$$u = (u(n), u(n - 1), \ldots, u(1))$$

so that $I(u)$ is the complement of $I(u)$ in $f(u)$, and therefore

$$f(u) = \frac{1}{2}(n(n - 1) - f(u)).$$

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

$$f(ww_1) = f(w - 1)$$

and hence

$$f(ww_1) = \frac{1}{2}(n(n - 1) - f(w)).$$

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 $p = f(w)$ such that $w = s_{e_1} \ldots s_{e_n}$.

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

$$f(u) = (s_{e_1}, \ldots, s_{e_n}, (a_i, a_p + 1) : 1 \leq r < n).$$

Proof: Let $w' = s_{e_0} \ldots s_{e_n}$. Then $f(w') = p - 1$ and hence by (1.2) and (1.3) we have

$$f(u) = s_{e_0} \ldots s_{e_n}.$$

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

$(b_1, \ldots, b_n, a_1, \ldots, a_p) \in R(w)$ for some $i = 1, 2, \ldots, p$.

Proof: By (1.7), applied to $w^{-1}$, we have $(b_1, a_1 + 1) \in R(w)$ and hence

$(b_1, a_1 + 1) = s_{e_0} \ldots s_{e_n}, (a_1, a_1 + 1)$

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

$$a_{e_0} = a_{e_1} = \ldots = (a_{e_n} = (s_{e_1} \ldots s_{e_n})^{-1}.$$
for some pair \((p,q)\) such that \(1 \leq p < q \leq r\).

Proof: Since \(f(s_a) = 1\) and \(f(t(s_a) \cdots s_1) < r\) there exists \(q \geq 2\) such that

\[ f(t(s_a) \cdots s_q) = q - 1, \quad f(t(s_a) \cdots s_q) < q \]

Let \( v = s_a \cdots s_q \), so that \( f(v) = q - 1 \) and \( f(tsv) \leq q - 1 \), whence by (1.3) we have \( f(tsv) = q - 2 \). Let \((b_1, \ldots , b_k, a)\) be a reduced word for \( tsvt \), then \((b_1, \ldots , b_k, a, b_1, \ldots , b_k)\) are reduced words for \( v \). By (1.8) (applied to \( v^{-1} \)) it follows that \( v = s_a \cdots s_q \), for some \( p = 1, 2, \ldots, q - 1 \), and hence

\[ w = tvs_a \cdots s_q = s_a \cdots s_q s_a \cdots s_q \]

If \( i < j \), let \( \tau_{ij} \) denote the transposition that interchanges \( i \) and \( j \) and fixes each \( k \neq i, j \). For each permutation \( w \), let \( s_j(w) \) denote the number of \( k \) such that \( i < k < j \) and \( w(k) \) lies between \( w(i) \) and \( w(j) \). Consideration of \( f(w) \) and \( f(ttw) \) shows that

\[ f(ttw) = \begin{cases} f(w) - 2s_j(w) - 1 & \text{if } f(w) > w(j), \\ f(w) + 2s_j(w) + 1 & \text{if } f(w) < w(j). \end{cases} \]

In particular, \( f(ttw) = f(w) + 1 \) if and only if \( s_j = 0 \).

(1.11) Let \( v, w \) be permutations and let \( s_1, \ldots, s_q \) be a reduced word for \( w \). Then the following conditions are equivalent:

- \( (i) \ f(v) < f(w) \) and \( v^{-1}w \) is a transposition,
- \( (ii) \ w = s_1 \cdots s_q \) for some \( r = 1, 2, \ldots, p \).

Proof: \((i) \rightarrow (ii)\). Suppose that \( v^{-1}w = t_{ij} \), so that \( v = tw_{ij} \). Then (1.10) shows that \( f(w) > w(j) \), so that \( (i,j) \notin f(w) \). Hence by (1.7) we have \( (i,j) = s_a \cdots s_q (or, s_a a_1) \) for some \( r = 1, 2, \ldots, p \), and therefore

\[ s_q = (s_1 \cdots s_q s_a) \cdots (s_1 \cdots s_q s_a)^{r-1} \]

Consequently

\[ w = tw_{ij} = (s_1 \cdots s_q) (s_1 \cdots s_q) \cdots (s_1 \cdots s_q) \]

\[ = s_1 \cdots s_q \]

\( (ii) \rightarrow (i) \). Clearly \( f(v) < f(w) \), and the calculation shows that \( v^{-1}w \) is the transposition \((11)\)
Notes on Schubert Polynomials

Let \( v = s_i s_{i+1} \cdots s_{r-1} \) for some \( r = 1, 2, \ldots, p \). Hence \( v = s_1 \cdots s_{r-1} \) and so \( v = w \) 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 \(-\). In other words, if \( v \) and \( w \) are permutations, \( v \leq w \) means that there exists \( r \geq 0 \) and \( \nu, \nu_1, \ldots, \nu_r, \) in \( S_n \) such that

\[
v = \nu_0 = \nu_1 = \cdots = \nu_r = w
\]

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

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

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

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

(ii) \( \Rightarrow \) (i). By definition there exist \( \nu_0, \nu_1, \ldots, \nu_m \), where \( m \geq 1 \), such that

\[
\nu_0 = \nu_1 = \cdots = \nu_m = w.
\]

We have \( v = \nu_{m+1} \) and \( \nu_{m+1} = w \). Hence there exists \( k = 1, 2, \ldots, m \) such that \( v_j = s_i v_j \) for \( 0 \leq j \leq k - 1 \), and \( v_k = v \).

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

\[
s_i v_{j+1} = s_i v_j = s_i v_{j+1} \quad (1 \leq j \leq k - 1).
\]

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

Hence

\[
v_k = s_i v_{k+1}.
\]

From (1) and (2) it follows that

\[
v = s_i v_{m+1} = s_i v_{m+2} = \cdots = s_i v_{2} = v_0 = w
\]

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 \( \nu \) be the longest element of \( S_n \), and replace \( v, w \) respectively by \( s_i v w \) and \( s_i w v \). Then we have

\[
\nu \leq \nu \Rightarrow s_i v w \leq s_i w v \quad (by \ (1.12))
\]

\[
\nu w v < s_i v w \quad (by \ (ii) \Rightarrow (i))
\]

\[
\nu i w < s_i w v \quad (by \ (1.12) \ again)
\]

and the proof is complete.]

(1.17) Let \( v, w \) be permutations and let \( w = (a_1, a_2, \ldots, a_p) \) be a reduced word for \( w \). Then the following conditions are equivalent:

(i) \( v \leq w \);

(ii) there exists a subsequence \( b = (b_1, \ldots, b_k) \) of \( a_1, \ldots, a_p \) such that \( v = a_1 \cdots a_k \);

(iii) there exists a reduced subsequence \( b = (b_1, \ldots, b_k) \) of \( a_1 \cdots a_p \).

Proof: It 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 \( r = p + q = f(v) + f(w) \). If \( r = 0 \), we have \( v = w = 1 \), so assume that \( r \geq 1 \). We distinguish two cases:

(a) \( v < w \). In this case we have \( b_1 \neq a_1 \), hence \( (b_1, \ldots, b_k) \) is a subsequence of \( (a_1, \ldots, a_p) \), which is a reduced word for \( v \). By the inductive hypothesis we have \( \nu \leq s_i w < w \), hence \( v \leq w \).

(b) \( v = w \). In this case \( f(v) + f(w) = p + q = r - 1 \), and \( s_i v = s_i w = v \). If \( a_i = b_i \) we have \( v_i = w_i = \cdots = w_{p-k} \), and if \( a_i \neq b_i \) then \( a_1, \ldots, a_i, b_{i+1}, \ldots, b_k \) is a non-reduced subsequence of \( (a_1, \ldots, a_p) \). Hence the inductive hypothesis implies that \( s_i v \leq w \). But also \( s_i w \leq w \), hence \( v \leq w \) by (1.16).}

(1.18) Let \( w \in S_n \), and let \( \nu \) be a transposition. Then

\[
f(\nu w) < f(\nu) \Rightarrow w < \nu.
\]

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 \( d = (n, n-2, \ldots, 1) \)) whose \( j \)-th column, for \( 1 \leq j \leq n-1 \), consists of the numbers \( a(1), \ldots, a(n-j) \) 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: It is easily seen that \( K(v) \leq K(w) \), and hence \( v \leq w \) implies \( K(v) \leq K(w) \).
Conversely, suppose that \( K(v) \subseteq K(u) \) and let \( j = j(v, u) \) be the smallest integer \( \geq 1 \) such that \( v(j) \neq u(j) \). If \( v = u \) we define \( j(v, u) = n \). We proceed by descending induction on \( j(v, u) \).

If \( j(v, u) = n \) we have \( v = u \), so assume \( j(v, u) = j < n \). Then \( v(j) \) is not equal to any \( v(1), \ldots, v(j) \) and hence is equal to \( v(k) \) for some \( k > j \).

For each \( i < j \) the \((n-i)\)th columns of \( K(v) \) and \( K(u) \) are identical, and since \( K(v) \subseteq K(u) \) it follows that \( v(j) < u(j) \), i.e. \( v(j) = v(k) \). Let \( v' = v(j) \), then by (1.10) we have \( v(k) \leq v' \) and hence \( v < v' \) by (1.18). Also \( v'(i) = v(i) \) for \( i < j \), and \( v'(j) = v(j) = u(j) \) so that \( j(v', u) > j \). Hence \( v' \leq v \) by the inductive hypothesis, and therefore \( v < u \).

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

\[ D(w) = \{ (i, j) \mid i \neq j \}. \]

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

\[ (i, j) \in D(w) \iff i < w^{-1}j < j < wi. \]

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

The graph \( G(w) \) of \( w \) is the set of points \((i, w(i)) \) \((1 \leq i \leq n) \), or equivalently \((w^{-1}j, j) \) \((1 \leq j \leq n) \). The complement of \( D(w) \) in \( \Sigma_m \) 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 \( w = (305142) \) and \( n = 6 \), the diagram \( D(w) \) consists of the points circled in the picture below:

```
  1  2  3  4  5  6

  1  2  3  4  5  6
  2  3  4  5  6  1
  3  4  5  6  1  2
  4  5  6  1  2  3
  5  6  1  2  3  4
  6  1  2  3  4  5
```

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

\[ S_m = \bigcup_{n \leq t} S_n \]

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

(1.23) (i) If \( f(w, r) > f(w) \) (i.e., \( w(r) < w(r + 1) \)) then

\[
\triangleq \text{c}_{r}(w) = \text{c}_{r}(w) + e_{r},
\]

where \( e_{r} \) is the sequence with \( 1 \) in the \( r \)th place and \( 0 \) elsewhere.

(ii) If \( (a_{1}, \ldots, a_{k}) \in f(w) \) then

\[
\triangleq \text{c}_{w}(w) = \sum_{i=1}^{k} e_{a_{i}}. 
\]

Proof: (i) follows from (1.3(i)), and (ii) follows from (i) by induction on \( p \).

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

\[
\triangleq c_{i}(w) > c_{i+1}(w) \iff w(i) > w(i + 1).
\]

Proof: Suppose that \( w(i) > w(i + 1) \). Then the \((i + 1)\)th row of \( f(w) \) 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 \((i + 1)\)th row of \( f(w) \) is contained in the \( i \)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 \( u = (w_{1}, w_{2}, \ldots) \) is any sequence and \( r \) is an integer \( \geq 0 \), let

\[
\triangleq \text{c}_{r}u = (w_{1}, w_{2}, \ldots, w_{r}, 0, w_{r+1}, w_{r+2}, \ldots)
\]

so that the operation \( \text{c}_{r} \) introduces a zero after the \( r \)th place. Then we have

(1.25)

\[
\triangleq c(w^{-1}) = \sum_{r \geq 0} c_{r} \cdot c_{r}u(1^{-r})
\]

where \((1^{-r})\) is the sequence consisting of \( c_{1}, \ldots, c_{r} \).

Proof: By induction on the length of \( c(w) \) it is enough to show that if \( w_{1} \) is the permutation whose code is \((c_{1}, c_{2}, \ldots)\) then

\[
\triangleq c(w_{1}^{-1}) = (1^{-r}) \cdot c_{r}c_{r}u(1^{-r}).
\]

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 column, we obtain (1.3).

Permutations

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

\[
\triangleq |\lambda(w)| = \text{Card} D(w) = f(w).
\]

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

\[
\triangleq \lambda(w') \geq \lambda(w^{-1}).
\]

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 \( w \in S_{n} \) for which \( \lambda(w) = \lambda(w^{-1}) \). They may be characterized in various ways:

(1.27) The following conditions on a permutation \( w \in S_{n} \) are equivalent:

(i) the set of rows of \( D(w) \) is totally ordered by inclusion;

(ii) the set of rows of \( f(w) \) is totally ordered by inclusion;

(iii) the set of columns of \( D(w) \) is totally ordered by inclusion;

(iv) the set of columns of \( f(w) \) is totally ordered by inclusion;

(v) there do not exist \( a, b, c, d \) such that \( 1 \leq a < b < c < d \) and \( w(d) < w(c) < w(b) < w(a) \); and

(vi) there exist \( u, v \in S_{2n} \) such that \( (w = u)D(v) = \text{D}(\lambda) \) is the diagram \( \lambda(w) \) in \( S_{n} \).

Proof: Since \( D(w) = (1 \times n)/w\) it is clear that (i) \( \Leftrightarrow \) (ii) and (i) \( \Leftrightarrow \) (iii). Moreover (i) \( \Leftrightarrow \) (ii) for either is false if and only if there is \( D(w) = (a, b, c, d) \) such that \( a < b, c < d \) and \( (a, b, c, d, \ldots) \) belong to \( D(w) \), whilst \((a, d)\) and \((c, b)\) do not. Let \( d = w^{-1}(b) \) and \( w = w^{-1}(d) \), then we have \( a < b < c < d \) and \( w(d) < w(c) < w(b) < w(a) \); thus, (i), (ii) and (iii) are all equivalent.

Next, it is clear that the conjunction of (i) and (ii) is equivalent to (iv). Then 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 \( w \in S_{n} \) is said to be vexillary if it satisfies the equivalent conditions of (1.27). By (1.27) (iii), the first non-vexillary permutation is \( (2421) \) in \( S_{4} \).
For each $w \in S_n$, let
\[ \mathbb{W} = \mathbb{W}_{w,\mathbb{W}} \]
where as before $w_0 = (n, n - 1, \ldots, 1)$ is the longest element of $S_n$. Then
(1.28) (i) \( \mathbb{W}(w) = \mathbb{W}(w) \).
(ii) $I(W) = I(w)$. 
(iii) $I(W) = I(w)^\prime$. 
Proof: (i) follows from (1.6) (or from (ii) below).
(ii) If $i < j$ then
\[ (i, j) \in I(W) \iff w(i) > w(j) \]
\[ \iff w(n + 1 - i) < w(n + 1 - j) \]
\[ \iff (n + 1 - j, n + 1 - i) \notin I(w). \]
(iii) follows from (6).
From (1.27) and (1.28) it follows that
\[ I(W) = \text{permutations that are strict} \iff w \text{ is strict}. \]
\[ I(W) = \text{permutations that are reflexive} \iff w \text{ is reflexive}. \]
\[ I(W) = \text{permutations that are } w^{-1} \text{ is reflexive} \iff \mathbb{W} \text{ is reflexive}. \]

**Dominant permutations**

We consider next two particular types of vexillary permutations.

(1.30) Let $w \in S_m$. 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 $\lambda = (\lambda_1, \ldots, \lambda_m)$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$. We shall show by induction on $i$ that
\[ (i, j) \in D(w) \iff 1 \leq j < \lambda_i. \]
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 $1 \leq k \leq i - 1$. Since $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$ 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 $w \in S_m$. Then its following conditions are equivalent:
(i) $c_i(w) \leq \cdots \leq c_m(w) \leq c_{m+1}(w) = 0$ for $i > r$;
(ii) $w(i) < w(i + 1)$ unless $i = r$.

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

(iii) $\Rightarrow$ (ii). We have
\[ c_i(w) = (w(1) - 1, \ldots, w(r) - r). \]
If $w$ satisfies the equivalent conditions of (1.31), $w$ is called a Grassmannian permutation. By (1.27)(ii), Grassmannian permutations are vexillary, and $w \in S_n$ is Grassmannian if and only if $\mathbb{W} = \mathbb{W}_{w,\mathbb{W}}$ 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 $c_i \geq c_j$ 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
\[ p_i = \{ j : (i, j) \in D(w) \}; \]
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 $p_i \supset c_j \supset p_k$ (where $\supset$ denotes strict containment), hence there exists $t \in p_i, t \notin p_k$. Let $s = w(t)$, then $s < t$ and (since $i \notin p_k$) we have $s \in p_i$ and $s \notin p_k$. Hence we fixed $(i, j)$ such that $i < j$ and $c_i \geq c_j$, the number of $k$ between $i$ and $j$ such that $c_k > c_j$ is at most $\text{Card}(p_i - p_k) = c_i - c_j$, so that (V1) is satisfied.

Next let $w$ be vexillary, $i < k < j$ and $c_i < c_j$, so that $p_i \supset c_j \supset p_k$. Let $s \in p_i$. If $s \notin p_k$ then $w(k) \leq s < w(t)$, so that $w(k)$ lies in $p_i$, but not in $p_k$, which is impossible. Hence $s \in p_k$ and therefore $p_i \supset c_k \supset p_j$. So we have $c_k \geq c_i$, 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 $\{ p_2, p_3, \ldots \}$ is totally ordered by inclusion.
Let $j > 1$ and suppose first that $c_1 \geq c_j$. If $p_0 \not\subseteq p_j$, there exists $t \in p_j$ such that $t \not\in p_0$, so that $w(t) < w(j)$. There are at least $c_1 - c_j + 1$ elements $t \in p_j$ such that $t \not\in p_j$, and since each such $t$ satisfies $t < w(j)$, it is of the form $t = w(k)$ for some $k$ between $1$ and $j$. Since $w(k) < t < w(j)$, it follows that $s \not\subseteq p_0$. Since either $p_0 \not\subseteq p_j$ or $p_j \not\subseteq p_0$, we conclude that $p_0 \not\subseteq p_j$ (strict inclusion) and hence that $c_0 < c_j$. Hence there are at least $c_1 - c_j + 1$ values of $k$ between $1$ and $j$ for which $c_0 < c_j$, contradicting (V1). Hence $p_0 \not\subseteq p_j$.

Finally, let $j > 1$ and $c_0 < c_j$, so that $w(1) < w(j)$. If $p_j \not\subseteq p_0$ there exists $s \in p_0$ such that $s \not\subseteq p_j$; we have $s = w(k)$ for some $k$ between $1$ and $j$, and since $w(k) < w(j)$ we have $c_0 < c_j$, contradicting (V2). Hence $p_0 \not\subseteq p_j$ in this case, and the proof is complete.

**Remark.** It is stated in [LS4, prop. 2 d] 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 $c_j > c_i$ for all $j > i$.

Since (V3) is implied by (V2), it follows from (1.22) that every vexillary code satisfies (V1) and (V2). However, the conjunction of (V1) and (V2) is not sufficient for vexillarity: for example, the permutation $w = (2576134)$ 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 $e_1 \leq \ldots \leq e_m$. The sequence

$$\varphi(w) = (e_1, e_2, \ldots, e_m)$$

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

**Remark.** There is another definition of the flag of a permutation $w_1$ due to M. Wachs [W2]. 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 $d_1' \leq \ldots \leq d_m'$, and let

$$\varphi^*(w) = (d_1', d_2', \ldots, d_m').$$

These two notions are not equivalent. In fact (1.32) (J. Allain). We have $\varphi(w) = \varphi^*(w)$ 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 $\varphi(w) = \varphi^*(w)$ if and only if $d_i - 1 = e_i$ for each $i$. Thus 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 $\varphi(w)$.

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

$$(1.34) \quad \lambda = (\lambda_1^e, \lambda_2^e, \ldots, \lambda_m^e)$$

where $\lambda_1 > \lambda_2 > \ldots > \lambda_m \geq 0$ and each $\lambda_i \geq 1$. For $1 \leq r \leq k$ let

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

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 $\lambda_r$ for some $r$, and $e_i = \max\{j : j \geq i \text{ and } c_j \geq \lambda_r\} = f_r$.

It follows that (whether $w$ is vexillary or not)

$$(1.35) \quad \varphi(w) = (f_1, f_2, \ldots, f_k).$$

Moreover we must have

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

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

(1.37) Suppose $w$ is a vexillary permutation with shape $\lambda$ and flag $\varphi$ given by (1.34) and (1.35). Then the $f_r$ must satisfy the inequalities

$$0 \leq f_r - f_{r-1} \leq m_r + \cdots + m_1 - \lambda_{r-1}.$$
Also

(2) \[ \text{Card } \{ k : s < k \leq f_r \text{ and } c_k = r \} \leq m_r. \]

since exactly \( m_r \) terms of the sequence \( c \) are equal to \( r \).

Finally we have

(3) \[ \text{Card } \{ k : s < k \leq f_r \text{ and } c_k > p_s \} = f_r - s \]

because \( c_k \leq p_s \) for all \( k > f_{r-1} \), and \( c_k \geq p_s \) for all \( k \) such that \( s < k \leq f_{r-1} \), by virtue of (V). From (1), (2), and (3) we deduce that

\[ f_r - s \leq p_{r-1} - p_s + m_r + f_{r-1} - s \]

which proves (1.37).

(1.38) For each sequence \( (f_1, \ldots, f_s) \) satisfying (1.36) and (1.37) there is a unique vexillary permutation \( w \) with shape \( \lambda \) and flag \( \phi = (\phi_1, \ldots, \phi_m) \). 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 \( m_s \) entries in \( c \) equal to \( p_s \) are inserted in the rightmost available spaces in the interval \( [1, f_s] \), and so on: for each \( r \geq 1 \), when all the terms \( c_k > p_s \) 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 \( 1 \leq i \leq f_r \) and \( c_i = p_s \), then by (V) we have \( c_j \geq p_s \)

for all \( j \) such that \( i < j \leq f_r \). Hence the entries equal to \( p_s \) in the sequence \( c \) must be inserted as above.

Conversely, if the sequence \( c \) is constructed as above, we claim that \( c \) satisfies (V) and (V)2, and hence \( w \) is vexillary by (1.32). Suppose first that \( i < j \) and \( c_i \geq c_j \): say \( c_i = p_s, c_j = p_r, r > s \).

Then the number of blanks such that \( i < k < j \) and \( c_k < p_r \) is equal to the number of black spaces in the interval \([f_r, f_s]\) after all the entries \( p_s \), \( r + 1 \leq i \leq s \) have been inserted, hence at most

\[ f_r - f_s = (m_{r+1} + \cdots + m_s) \]

which by (1.37) is \( p_r - p_s \). Hence the sequence \( c \) satisfies (V1). Suppose next that \( i < j \) and \( c_i < c_j \); say \( c_i = p_s, c_j = p_r, r < s \). Then we have \( j \leq f_r \leq f_s \). From the definition of the sequence \( c \), it follows that for each \( \hat{k} \) such that \( i \leq \hat{k} \leq f_r \) we have \( c_{\hat{k}} = p_s \), and hence \( c_k = c_s \)

whenever \( i < k < j \). Consequently the condition (V2) is satisfied, and the proof is complete. ||

\[ \phi(w) \]

\[ (1.39) \quad \phi(w) = (\phi_1, \phi_2, \phi_3) \text{ must satisfy } 0 \leq f_3 - f_2 \leq 3, 0 \leq f_2 - f_1 \leq 2, \]

and there are \((2 + 1)(2 + 1) = 12\) vexillary classes, and the representatives of these classes for which \( w(I) \neq 1 \) (or equivalently \( \phi_1(w) \neq 0 \) are as follows:

\[ \phi(w) \]

\[ (1.38) \quad \]
Proof: We proceed by induction on \( \ell(w) = \ell \). Let \( c = (c_1, c_2, \ldots) \) be the code of \( w \), and let \( w' \) be the permutation with code \( c' = (c_1, c_2, \ldots) \). We may assume that \( c_1 \neq 0 \). Then \( c_1 = p_r \) for some \( r \), and we have
\[
\lambda(w') = (p_r^{n_r}, \ldots, p_1^{n_1}),
\]
\[
\phi(w') = ((f_1 - 1)^{n_1}, \ldots, (f_r - 1)^{n_r}, \ldots, (f_s - 1)^{n_s}).
\]
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 \( \lambda(w') = \lambda(w')' \), so that
\[
\lambda(w''') = ((f_1 - 1)^{n_1}, \ldots, (f_s - 1)^{n_s}).
\]
where \( s = k + 1 - r \). We have \( \ell(w') = \ell(w) - c_1 \), so that the inductive hypothesis applies to \( w' \).
Hence if \( q_1, \ldots, q_k \) are defined by the formula (s), we have
\[
(1) \quad c(w') = (d_1 + 1, \ldots, d_p, d_p + 1, 0, d_{p+1}, d_{p+1}, \ldots).
\]
From (1) and (2) and (1.40) it follows that
\[
\phi(w') = (p_r^{n_r}, \ldots, p_1^{n_1}, p_1^{n_1}, \ldots, p_1^{n_1})
\]
as required.

If \( w \in S_n \), let \( \Pi_w = \Pi_{w_0}w \), where \( w_0 \) is the longest element in \( S_n \). If \( w \) is vexillary, of shape \( \lambda \), then \( \Pi_w \) is vexillary of shape \( \lambda' \), by (1.27) and (1.28). Let
\[
\phi(\Pi_w) = (\lambda^{n_1}, \ldots, \lambda^{n_k})
\]
be the flag of \( \Pi_w \). Then we have
\[
(1.42) \quad \lambda_i = n - f_{i+1} \quad (1 \leq i \leq k).
\]
For once we shall leave the proof to the reader.
Let \( N_{\lambda} \) denote the number of non-vexillary \( w \in S_n \), and let
\[
P_\lambda = N_\lambda/n!
\]
be the probability that an element of \( S_n \) is non-\( \text{vexillary} \). The first few values of \( N_n \) and \( P_n \) are:

\[
\begin{array}{c|c|c}
\hline
n & N_n & P_n \\
\hline
1 & 0 & 0 \\
2 & 0 & 0 \\
3 & 0 & 0 \\
4 & 1 & 0.042 \\
5 & 17 & 0.142 \\
6 & 207 & 0.288 \\
7 & 2279 & 0.452 \\
\hline
\end{array}
\]

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 \( \text{vexillary} \) condition \((1.27)(\text{iii})\) is 23/24 (because \( S_4 \) contains only one \( \text{non-\( \text{vexillary} \)} \) permutation), we see that the probability that \( w \in S_n \) is \( \text{vexillary} \) is at most \((23/24)^{n/4}\), hence decreases exponentially to zero. (A. Lascoux.) Thus the \( \text{vexillary} \) permutations in \( S_n \) become sparser and sparser as \( n \) increases.

Instead of counting \( \text{non-\( \text{vexillary} \)} \) permutations, we may attempt to count \( \text{vexillary} \) permutations. Let us say that a permutation \( w \in S_n \) is \( \text{primitive} \) 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 \( \text{vexillary} \) (resp. \( \text{primitive} \)) permutations \( w \in S_n \).

Since each \( \text{vexillary} \) permutation \( w \in S_n \) gives rise to \( r + 1 \) \( \text{vexillary} \) permutations in \( S_{nr} \), namely \( I_p \cdot w \cdot I_q \) 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

\[
V(t) = \frac{1}{1 - U(t)} U(t) - 1
\]

For each partition \( \lambda \neq \emptyset \), let \( U_{\lambda, k} \) denote the number of \( \text{vexillary} \) permutations of shape \( \lambda \) in \( S_n \), and let

\[
U_{\lambda}(t) = \sum_{k \geq 1} U_{\lambda, k} t^k
\]

\[* N_1 \text{ was computed by A. Garvin. I would guess that } N_4 \text{ is of the order of } 24000.\]

so that

\[
U(t) = \sum_{\lambda \neq \emptyset} U_{\lambda}(t)
\]

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

\[
\lambda = (p_1^{\alpha_1}, p_2^{\alpha_2}, \ldots, p_k^{\alpha_k})
\]

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

\[
f_r \geq m_1 + \cdots + m_r \quad (1 \leq r \leq k),
\]

\[
0 < f_r - f_{r-1} \leq m_r + m_{r-1} - p_r \quad (2 \leq r \leq k).
\]

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

\[
f_r = m_1 + \cdots + m_r \quad \text{for some } r = 1, \ldots, k.
\]

In general, if \( e = (e_1, \ldots) \) is the code of a permutation \( w \), then \( w \in S_n \) if and only if \( n \geq e_1 + \cdots \) for \( 1 \leq i \leq r \), where \( r \) is the length of \( e \). In other words, the least \( n \) for which \( w \in S_n \) is \( n = \max\{e_1 + \cdots, 1 \leq i \leq r \} \). In the case of a \( \text{vexillary} \) 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 \([f_r+1, f_{r+1}] \cap (1, \ldots, k)\), and hence \( w \) will be \( \text{vexillary} \) in \( S_n \) if and only if \( w \) satisfies \((1)\) above and

\[
\begin{align*}
n &= \max\{p_r + f_r : 1 \leq r \leq k\}; \\
\end{align*}
\]

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

\[
\begin{align*}
x_1 + u_1 & \leq x_2 + u_2 \leq \cdots \leq x_k + u_k \\
\end{align*}
\]

and

\[
\begin{align*}
m_r + p_{r-1} - p_r & \geq f_r - f_{r-1} \\
&= (u_r + x_r) - (u_{r-1} + x_{r-1}) \\
&= m_r + u_r - u_{r-1}
\end{align*}
\]
Notes on Schubert Polynomials

so that

\[ p_1 + n_1 2 + n_2 3 \ldots 2 p_n + n_n. \]

It now follows that

\[ U_k(i) = \sum_{u} \text{divided differences} \]

summed over the integer vectors \( u = (u_1, \ldots, u_k) \in \mathbb{N}^k \) having at least one zero component, and

satisfying the inequalities (3), (4) above. We have

\[ U_k(i) = \sum_{u} \text{divided differences} \]

and

\[ U_k(i) = U_k(i) \]

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

Added in proof.

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

\[ V_\lambda = \sum_{(\lambda_1, \ldots, \lambda_k) \in D_\lambda} (f^\lambda)^2 \]

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

\[ V_\lambda \sim c n^{-d} \]

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

The formula (1) gives that \( N_k = 24003 \).

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 \( \partial_{xy} \) 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 \( f = x^p y^q \) we have

\[ \partial_{xy}(x^p y^q) = \frac{x^{p'} y^{q'} - x^p y^q}{x - y} = \sigma(r - s) \sum_{0 < r - s} \]

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 \( fg \), \( \partial_{xy} \) acts according to the rule

\[ \partial_{xy}(fg) = (\partial_{xy} f)g + (s_{xy} f)(\partial_{xy} g). \]

In particular we have

\[ \partial_{xy}(fg) = f \partial_{xy} g \]

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

\[ \partial_{xy}(s_{xy} f) = \partial_{xy} f, \]

\[ \partial_{xy}(s_{xy} f) = \partial_{xy} f, \]

\[ \partial_{xy}(s_{xy} f) = \partial_{xy} f. \]

\[ \partial_{xy}(s_{xy} f) = \partial_{xy} f. \]

\[ \partial_{xy}(s_{xy} f) = \partial_{xy} f. \]
Notes on Schubert Polynomials

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

\[(x-p)^{-1}(x-p)^{-1}(y-z)^{-1}\sum_{w\in S_n} f(w)u^w,
\]

where the symmetric group $S_n$ permutes $x, y$ and $z$, and $f(w)$ is the sign of the permutation $w$.

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

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

for each $n \geq 1$, and

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

For each $i \geq 1$, let

$$\Delta_i = \Delta_{x_i} \ldots \Delta_{x_1}.
$$

Each $\Delta_i$ is a linear operator on $P_n$ (and on $P_i$ for $n > i$) of degree $-1$. From (2.3) we have (compare with (1.1))

\[(2.4) \begin{cases}
q^0 = 0, \\
\Delta_i q^0 = \Delta_i q^0, \\
\Delta_i q^0 = \Delta_i (a_i (q^0) + q^0 + 1).
\end{cases}
\]

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

$$\Delta_a = \Delta_{a_p} \ldots \Delta_{a_1}.
$$

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

(2.5) If $a, b \in R(u)$ then $\Delta_a = \Delta_b$.

Proof: We proceed by induction on $p = f(u)$. Let us write $a \equiv b$ to mean that $\Delta_a = \Delta_b$. The inductive hypothesis then implies that

$$a \equiv b \text{ if either } a_i \equiv b_i \text{ or } a_i \equiv b_i.
$$

By the exchange lemma (1.8) we have

$$a_i = (b_i, a_i, \ldots, a_i) \in R(u)
$$

for some $i = 1, \ldots, p$. If $i \neq p$ then $b \equiv a \equiv a$ by virtue of (x), so that $a = b$. If $i = p$ and $|b_i - a_i| > 1$ then by (2.4) and (1.1)

$$a_i' = (a_i, b_i, a_i, \ldots, a_i) \in R(u)
$$

and $a \equiv a_i' \equiv a_i = b$, so again $a \equiv b$.

Finally, if $i = p$ and $|b_i - a_i| = 1$, we apply the exchange lemma again, this time to $a_i$ and $a_i'$, this shows that

$$a_i = (a_i, b_i, a_i, \ldots, a_i) \in R(u)
$$

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

$$\Delta_a = (a_i, b_i, a_i, \ldots, a_i) \in R(u)
$$

and $a \equiv a_i' \equiv a_i = b$. Hence $a \equiv b$ in all cases.

Remark. For any permutation $u$, let $G(u)$ denote the graph whose vertices are the reduced words for $u$, and in which a reduced word $a$ is joined by an edge to each of the words obtained from $a$ by either interchanging two consecutive terms $i, j$ such that $|i-j| > 1$, or by replacing three consecutive terms $i, j, k$ such that $|i-k| = 1$ by $j, i, j$. Then the proof of (2.6) shows that

(2.6) The graph $G(u)$ is connected.

From (2.5) it follows that we may define

$$\Delta_a = \Delta_u
$$

unambiguously, where $a$ is any reduced word for $u$. By (2.2), the operators $\Delta_u$ for $u \in S_n$ are $\Delta_n$ linear, where

$$L_n = \mathbb{Z}[t_1, t_2, \ldots, t_n]^{S_n} \subset P_n
$$

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

An sequence $a = (a_1, \ldots, a_p)$ will be said to be reduced if $a \subset R(u)$ for some permutation $u$.

(2.6) If $a = (a_1, \ldots, a_p)$ is not reduced, then $\Delta_a = 0$.

Proof: By induction on $p$. If $a' = (a_1, \ldots, a_p)$ is not reduced, then $\Delta_{a'} = 0$ and hence $\Delta_a = \Delta_{a'} \Delta_a = 0$. So we may assume that $a'$ is reduced. Let $u = x_{a_1} x_{a_2} \ldots x_{a_p} = x_{a_1} x_{a_2} \ldots x_{a_1}$. We have

$$d(u) = p - 1 \text{ and } f(u) = p - 1,\text{ hence by (1.3) } f(u) = p - 2, \text{ so that } d(u) = f(u) = f(u) + 1.
$$

Consequently $\Delta_a = \Delta_a \Delta_{a'}$ and therefore $\Delta_a = \Delta_a \Delta_a = \Delta_a \Delta_a = 0$.


(2.7) Let \( u, v \) be permutations. Then
\[
\delta_u \delta_v = \begin{cases} 
\delta_{uv} & \text{if } (uv) = (v) + (u), \\
0 & \text{otherwise}.
\end{cases}
\]

Proof: (2.5), (2.6).

(2.8) Let \( w \) be a permutation, \( i \geq 1 \). Then
\[
\delta_w = \delta_w \quad \text{and} \quad \delta_w = 0, \text{hence the result follows from (2.7).}
\]

As before let \( \omega = (v, n - 1, 2, \ldots, 1) \) be the longest element of \( \mathbb{S}_n \). One element of \( R(\omega w) \) is
\[
(1, 2, \ldots, n - 1, 2, \ldots, n - 2, \ldots, 1).
\]

(2.10) We have
\[
\delta_w = a_w^{-1} \sum_{u \in \mathbb{S}_n} c(u)w
\]
where \( a_w = \prod_{1 \leq i < j \leq n} (x_i - x_j) \), and \( c(u) = \pm 1 \) is the sign of \( u \).

Proof: From the definition it follows that \( \delta_w \) is of the form
\[
\delta_w = \sum_{u \in \mathbb{S}_n} c_u w
\]
with coefficients \( c_u \) rational functions of \( x_1, \ldots, x_n \). By (2.8) we have \( \delta_w = \delta_w \) for \( 1 \leq i \leq n - 1 \), so that \( \delta_w = \delta_w \) for all \( v \in \mathbb{S}_n \), and therefore
\[
\delta_w = \sum_{u \in \mathbb{S}_n} c(u)w.
\]

Comparison of (1) and (2) shows that
\[
c_{uv} = c(u w) \quad (u, v \in \mathbb{S}_n).
\]

Hence all the coefficients \( c_{uv} \) are determined by one of them, say \( c_{uv} \) from the sequence (2.9) for \( w_0 \) it is easily checked that the coefficient of \( w_0 \) in \( \delta_w \) is
\[
c_{w_0} = c(w_0) a_w^{-1}.
\]

Hence from (3) we have
\[
c_{w} = c(w_0) a_{w_0}^{-1} = c(w) a_{w}^{-1}
\]
which proves (2.10).

From (2.10) it follows that, for any \( a = (a_1, \ldots, a_n) \in \mathbb{N}^n \),
\[
\delta_{a_1} \cdots \delta_{a_n} = \delta_{a_1} \cdots \delta_{a_n} = \delta_{a_1} \cdots \delta_{a_n}
\]
where \( a^\ast \) means \( a^\ast_1 \cdots a^\ast_n \), \( \delta = (n - 1, n - 2, \ldots, 0) \) and \( R(\omega) \) is the Schur function indexed by \( \delta = \delta \). Then \( \delta_{a_1} \cdots \delta_{a_n} \) is a \( \lambda_{a_1} \)-linear mapping of \( R(n) \) onto \( R(n) \).

For \( w \in \mathbb{S}_n \), let \( \omega = w_0 w w_0 \). Then
\[
\delta_w = c(w) w_0 \delta_{w_0} w_0.
\]

Proof: From the definition of \( \delta_w \) we have
\[
w_0 w w_0 = - \delta_{w_0}.
\]

from which (2.12) follows easily, since \( w_0 = 1 \).

If \( f \) and \( g \) are polynomials in \( x_1, x_2, \ldots \), the expression of \( \delta_{a_1}(f g) \) as a sum of polynomials \( \delta_{a_i} f / \delta_{a_i} g \) (i.e. the "Liebnitz form") for \( \delta_{a_1} \) 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:
\[
(2.13) \quad \text{If } f = \sum a_i x_i \text{ then }
\]
\[
\lambda_i(fg) = f(w_0) g + \sum (a_i - a_j) \delta_{a_i} g
\]

summed over all pairs \( i < j \) such that \( f(w_0) = f(1) = 1 \), where \( \lambda_i \) is the transpose that interchanges \( i \) and \( j \).

Proof: Let \( (a_1, \ldots, a_n) \) be a reduced word for \( w \). Since \( f \) is linear it follows from (2.3) that
\[
\delta_{a_1} f = \delta_{a_1} \cdots \delta_{a_n} f = \delta_{a_1} \cdots \delta_{a_n} f + \sum_{r \geq 1} \delta_{a_1} \cdots \delta_{a_r} f(\delta_{a_{r+1}} \cdots \delta_{a_n} f)
\]

Now \( \delta_{a_1} \cdots \delta_{a_r} \cdots \delta_{a_n} = 0 \) unless \( (a_1, \ldots, a_r, a_{r+1}) \) is reduced, and then by (1.11) it is equal to
\[
\delta_{a_1} \cdots \delta_{a_r} \cdots \delta_{a_n} w = w \delta_{a_1} \cdots \delta_{a_r} \cdots \delta_{a_n} \delta_{a_{r+1}},
\]

where \( w = a_{r+1} \cdots \delta_{a_n} \cdots \delta_{a_1} \), has length \( p = l - f(w) = 1 \), and \( c = \delta_{a_1} \cdots \delta_{a_r} \cdots \delta_{a_n} f = a_{r+1} f \) where
\[
(i,j) = (s_{a_1}, \ldots, s_{a_r}(\delta_{a_{r+1}} a_{n+1})), \quad a_i = a_j = a_{r+1} \|.
\]

We also introduce the operators \( \delta_i (i \geq 1) \) defined by
\[
\delta_i = \delta_i (x_i f).
\]
In place of (2.4) we have
\[
\begin{align*}
\pi_i^2 &= \pi_i, \\
\pi_i \pi_j &= \pi_j \pi_i, \\
\nu \pi_i \kappa \pi_j &= \pi_i \kappa \pi_j \pi_i \kappa.
\end{align*}
\]
If we define \( \pi_a \) to be \( \pi_{a_1} \ldots \pi_{a_n} \) for any sequence \( a = (a_1, \ldots, a_n) \) of positive integers, then corresponding to (2.5) we have
\[
\nu \pi_a = \pi_a.
\]
(2.15) If \( a, b \in \mathbb{R}(w) \) then \( \pi_a = \pi_b \).

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
\[
\pi_a = \pi_a
\]
unambiguously, where \( a \) is any reduced word for \( w \).

In place of (2.10) we have
\[
(2.15) \quad \text{For any } f \in P_w, \quad \nu_a f = a^{(2)} \sum_{x \in S^*} c(u) (x^2 f) = \delta_u (x f).
\]
In particular, if \( a \in \mathbb{N}^* \),
\[
(2.16) \quad \nu_a \pi^a = \pi_a (x_1, \ldots, x_a)
\]
Proof. We have
\[
\nu f = 0 (x_1 f),
\]
\[
\nu \pi_i \nu f = 0 (x_i \pi_i (x_1 f)) = \delta_i (x_1 \pi_i f)
\]
and generally
\[
\nu x \ldots \nu f = \delta_i (x_i \ldots x_1 f)
\]
for each \( n \geq 1 \). From this and (2.10) it follows easily that \( \nu_{a_n} f = \delta_u (x^2 f, a) \).

Let \( (a_1, \ldots, a_n) \) be a reduced word for \( w \). Then
\[
\delta_u = \delta_{a_1} \delta_{a_2} \delta_{a_3} \ldots
\]
\[
= (x_{a_1} = x_{a_1+1})^{-1} (1 - x_{a_1} (x_{a_1} - x_{a_1+1})^{-1} (1 - x_{a_1}) \ldots
\]
which shows on expansion that \( \delta_u \) is of the form
\[
\delta_u = \sum_{s \in S^*} f_{s^u}
\]
where \( f_{s^u} \) are rational functions of \( x_1, x_2, \ldots, \), and in particular (by (1.3))
\[
f_{s^u} = (-1)^{s} \prod_{(i,j) \in T_{s^u}} (x_i - x_j)^{-1}
\]
and thus is \( \neq 0 \). It follows that the \( \delta_u \) are linearly independent over the field of rational functions \( \mathbb{Q}_{w} = Q(x_1, x_2, \ldots) \).

Now from (2.2) we have
\[
\delta_u (f g) = \delta_u (f) g + (x_u) (\delta_u g)
\]
or equivalently, if \( f : P_w \otimes P_w \rightarrow P_w \) is the multiplication map,
\[
\delta_u \circ f = f \circ (\delta_u \otimes I + I \otimes \delta_u).
\]

From this it follows that
\[
\delta_u \circ f = f \circ (\delta_u \otimes I + I \otimes \delta_u) \circ (\delta_u \otimes I + I \otimes \delta_u).
\]

On expansion this is a sum over subsequences \( \delta \) of \( a = (a_1, \ldots, a_n) \), say
\[
\delta_u \circ f = \sum_{\delta \in S^*} \delta_u (\delta) \circ \delta,
\]
where
\[
\delta_u (\delta) = \delta_u (\delta, a) = \sum_{a \in S^*} \delta_u (a).
\]
and
\[
\delta_u (a, b) = \delta_u (a, b) = \{ \delta_u, \text{ if } a \in S, \text{ otherwise } \}
\]
Since \( \delta_u = 0 \) if \( b \) is not reduced (2.6), the sum is over reduced subsequences \( \delta \) of \( a \), and by (1.17) we can write
\[
\delta_u \circ f = \sum_{\delta \in S^*} \delta_u (\delta) \circ \delta,
\]
where for \( \nu \leq w \)
\[
(3) \quad \delta_u \circ f = \sum_{\delta \in S^*} \delta_u (\delta)
\]
summed over subsequences \( \delta \subset \mathbb{R}(w) \) such that \( k \) is a reduced word for \( \nu \).

So for each pair of permutations \( w, \nu \) such that \( w \geq \nu \) we have a well-defined operator \( \delta_u \nu \) on \( P_w \), defined by (3). Since the \( \delta_u \) are linearly independent, the definitior (3) is independent of the reduced word \( a \in \mathbb{R}(w) \).
For each pair \( u, v \in S_n \) such that \( u \geq v \) there is a linear operator \( \partial_{uv} \) on \( \mathbb{F}_m \) such that

\[
\partial_{uv}(f) = \sum_{i=1}^{m} (\partial_{u,v} f) \cdot \partial_i g.
\]

\( \partial_{uv} \) has degree \( -(u) + (v) \).

Examples.

1. Let \( u = v \), then

\[
\partial_{uu} = u^{-1}(u,v) = u^{-1}a_n \cdots a_2 = 1.
\]

2. Let \( u = 1 \), then

\[
\partial_{u1} = \partial_{0u} = \partial_{0} = \partial_u.
\]

3. Suppose that \( v = s_a \), so that \( v = s_{a_1} \cdots s_{a_r} \) for an unique \( r \geq 1 \). Then \( b = \{a_1, \ldots, a_r, b\} \) and

\[
\partial_{ub} = s^{-1}(u,b) = \partial_a \cdots \partial_{a_r} = \partial_v.
\]

Now \( u = w t \) where \( t \) is the transposition

\[
t = t_j = a_n \cdots a_j \quad (1 < j)
\]

so that \( (i, j) = s_{a_n} \cdots s_{a_2} (a_n, a_j + 1) \) and therefore

\[
\partial_{u,j} = s_{a_n} \cdots s_{a_2} (s_{a_1} - s_{a_j} + 1) (1 - s_{a_2}) \cdots (1 - s_{a_n}) = (a_j - a_i)^{-1}(1 - t_j)
\]

is the divided difference operator \( \partial_{u,j} \).

The product formula for \( \partial_{uv} \) is

\[
(2.28) \quad \partial_{uv}(f) = \sum_{u \geq v} w^{-1}(\partial_{u,v} f) \partial_{uw}.
\]

Proof. We have

\[
\partial_{uv}(f) = \sum_{u \geq v} w \partial_{u,v}(f) \partial_u h
\]

and on the other hand

\[
(2) \quad \partial_{uv}(fgh) = \sum_{u \geq v} w \partial_{u,v}(f) \partial_{u,v}(gh)
\]

Comparison of (1) and (2) gives

\[
\partial_{uv}(f) = \sum_{u \geq v} \partial_{u,v}(f) \cdot \partial_{uv}(g)
\]

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 \( X^r(X) \) for the \( r \)th exterior power, and \( h_r(X) \) in place of \( S^r(X) = (1 - (-1)^r X^r(X)) \) 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 \( \lambda, \mu \) are partitions and \( X \in R \), the skew Schur function \( s_{\lambda/\mu}(X) \) is defined by the formula

\[
s_{\lambda/\mu}(X) = \det(h_{\lambda_i - \mu_j}(X))_{i,j \geq 0}
\]

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 \( \lambda, \mu \) be partitions of length \( \leq n \); then the multi-Schur function \( s_{\lambda/\mu}(X_1, \ldots, X_n) \) is defined by

\[
s_{\lambda/\mu}(X_1, \ldots, X_n) = \det(h_{\lambda_i - \mu_j}(X_i))_{i,j \geq 0}.
\]

We also define

\[
s_{a}(X_1, \ldots, X_n) = \det(h_{a_i - \mu_j}(X_i))_{i,j \geq 0}
\]

for any sequence \( a = (a_1, \ldots, a_n) \) 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

\[
s_{\lambda/\mu}(X_1, \ldots, X_n) = \det(h_{\lambda_i - \mu_j}(X_i))_{i,j \geq 0}
\]

with arguments constant in each column. However, we get nothing essentially new: if we define partitions \( \lambda, \mu \) by

\[
\lambda_i = N - \mu_{i+1}, \quad \mu_i = N - \mu_{i+1} \quad (1 \leq i \leq n)
\]

then

\[
s_{\lambda/\mu}(X_1, \ldots, X_n) = s_{\lambda_1/\mu_1}(X_1) \cdots s_{\lambda_n/\mu_n}(X_n)
\]
where \( N \geq \text{max}(\lambda_1, \mu_1) \) (so that \( \lambda \) and \( \mu \) are the respective complements of \( \lambda \) and \( \mu \) in the rectangle \( (N^2) \)), then we have

\[
e_{ij}(X_1, \ldots, X_n) = e_{ij}(X_n, \ldots, X_1)
\]
as one sees by replacing \((i,j)\) by \((n+1-j, n+1-i)\) in the determinant \((3.1)\).

(3.2) We have

\[
e_{ij}(X_1, \ldots, X_n) = 0
\]
unless \( \lambda \geq \mu \).

Proof: If \( \lambda \geq \mu \) then \( \lambda_i < \mu_i \) for some \( i \leq n \), and hence

\[
h_i = \mu_i - \lambda_i \leq j \leq \lambda_i - \mu_i < 0
\]
whenever \( i \geq r \) and \( j \leq r \). It follows that the matrix \((h_{\lambda_i-\mu_i+1}(X_i))\) has an \((n-r+1) \times r\) block of zeros in the south-west corner, and hence its determinant vanishes.

(3.3) If \( \lambda \geq \mu \) and \( (\lambda, \mu) \) are an \( r \times n \) matrix, then

\[
e_{ij}(X_1, \ldots, X_n) = e_{ij}(X_n, \ldots, X_1)
\]

Proof: We have \( \lambda_i = \mu_i \leq 0 \) for \( r+1 \leq i \leq n \), and hence for each \( s > r \) the \( s^{th} \) row of the matrix \((h_{\lambda_i-\mu_i+1}(X_i))\) 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 \( e_a(X) = 0 \) for all sufficiently large \( n \). We then define the rank \( r(X) \) of \( X \) to be the largest \( r \) such that \( e_a(X) \neq 0 \). If \( X, Y \) both have finite rank, the formula

\[
e_a(X + Y) = \sum e_b(X) e_b(Y)
\]
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 \( a \in Z^k \),

\[
e_a(X_1, \ldots, X_n) = \text{det}(h_{a_{n-1}+1}(X_1 - Y_1), \ldots, h_{a_{n-1}+1}(X_n - Y_n)) .
\]

Proof: We have

\[
h_{a_{n-1}+1}(X_i - Y_j) = \sum_{i \leq j} h_{a_{n-1}+1}(X_i - Y_i),
\]
since \( h_r(-Y_j) = (-1)^{Y_j} e_r(Y_j) = 0 \) if \( r \geq j \). Hence the matrix

\[
(h_{a_{n-1}+1}(X_i - Y_i))_{i,j \leq n}
\]
is the product of the matrix

\[
(h_{a_{n-1}+1}(X_i))_{i < j, j < n}
\]
and the matrix

\[
(h_{a_{n-1}+1}(Y_i))_{i,j < n}
\]
which is unitriangular. Now take determinants.

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

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

\[
X_i = x_1 + \ldots + x_i, \quad Y_i = y_1 + \ldots + y_i
\]
for each \( i \geq 0 \). Then for all \( a \in N^n \) we have

\[
s_\lambda(X_1 - Y_1, \ldots, X_n - Y_n) = \prod_{i=1}^n (x_i - y_i) .
\]

In particular, if \( \lambda \) is a partition of length \( \leq n \),

\[
s_\lambda(X_1 - Y_1, \ldots, X_n - Y_n) = \prod_{i \in \lambda^\circ} (x_i - y_i) .
\]

Proof: From (3.4) we have

\[
s_a(X_1 - Y_1, \ldots, X_n - Y_n) = \text{det}(h_{a_{n-1}+1}(X_1 - Y_1), \ldots, h_{a_{n-1}+1}(X_n - Y_n))_{i,j \leq n} .
\]

If \( j > i \), then

\[
h_{a_{n-1}+1}(X_i - Y_i, X_{j-1} - Y_{j-1}) = \pm s_{a_{n-1}+1}(Y_n, X_{j-1} - X_i)
\]
which is 0 because

\[
rk(Y_n + X_{j-1} - X_i) \leq a_i + (j - 1) - 1 < a_i + i + 1 .
\]
Notes on Schubert Polynomials

Hence the determinant at \( (-) \) is triangular, with diagonal elements
\[
h_{\alpha}(X_1 - Y_n, ..., X_{n-1}) = h_{\alpha}(x_n - Y_n)
\]
\[
= \sum_{\tau \in S_n} (-1)^{\tau} h_{\alpha}(x_n, ..., x_1)
\]
\[
= \sum_{\tau \in S_n} (-1)^{\tau} e_{\tau}(Y_n) e_{n-\tau}
\]
\[
= \prod_{i=1}^{n} (x_n - y_i).
\]

The formula (3.5) now follows.\

In particular, when all the \( y_i \) are zero we have
\[
\delta_{\alpha}(X_1, ..., X_n) = \prod_{i=1}^{n} x_i^{\alpha_i} = x^\alpha
\]
for all \( \alpha \in \mathbb{N}^n \). Also, when all the \( x_i \) are zero (and \( \alpha \) is a partition \( \lambda \)) we have
\[
s_{\lambda}(-Y_1, ..., -Y_n) = (-1)^{\lambda_1} \prod_{i=1}^{n} y_i
\]
\[
= (-1)^{\lambda_1} y^\lambda.
\]

If we replace the \( y_i \)'s by \( x_i \)'s, and \( \lambda \) by \( \lambda' \), this becomes
\[
x^\lambda = (-1)^{\lambda_k} h_{\lambda}(x_1, ..., x_n).
\]

(3.5') Let \( \lambda = (\lambda_1, \lambda_2, ..., \lambda_{n+1}) \) be a partition of length \( \leq n \), and \( X_1, ..., X_n \) elements of a \( \lambda \)-ring \( R \). Suppose that \( i < j \) are such that \( \lambda_i = \lambda_{i+1} = \cdots = \lambda_j \) and
\[
r\lambda(X_j - X_i) \leq j - k \quad \text{for} \quad i \leq k \leq j.
\]

Then
\[
s_{\lambda}(X_1, ..., X_n) = s_{\lambda}(X_1, ..., X_i, X_{i+1}, X_{i+2}, ..., X_n),
\]
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 \( r\lambda(Y) \leq j - i \). For all \( m \geq 0 \) we have
\[
h_m(X_j) = h_m(X_j - Y)
\]
\[
= \sum_{\tau \in S_m} (-1)^{\tau} e_{\tau}(Y) h_{m-\tau}(X_j).
\]

It follows that if we replace the \( i \)th row of the determinant \( s_{\lambda}(X_1, ..., X_{n-1}, X_j, X_{j+1}, ..., X_n) \) by
\[
\sum_{j=\lambda_1}^{\lambda_{n+1}} (-1)^{\lambda_j} e_{\lambda_j}(Y) h_{\lambda_{n+1}}(X_1, ..., X_{j-1}, X_{j+1}, ..., X_n)
\]
we shall obtain
\[
s_{\lambda}(X_1, ..., X_{n-1}, X_j, X_{j+1}, ..., X_n)
\]
with \( j - i \) arguments equal to \( X_j \). The proof is now completed by induction on \( j - i \).

Duality

Let \( (X_n)_{n \in \mathbb{Z}} \) be a sequence in the \( \lambda \)-ring \( R \) such that
\[
r\lambda(X_n - X_{n-1}) \leq 1
\]
for all \( n \in \mathbb{Z} \).

Let \( I \) be any interval in \( \mathbb{Z} \). Then the inverse of the matrix
\[
H = (h_{\lambda}(X_j))_{j,k \in I}
\]
is zero because \( r\lambda(X_{j+1} - X_k) = j - k \) for all \( k \leq j \).

Proof: Let \( K \) denote the matrix \( (h_{\lambda}(X_j))_{j,k \in I} \). The \((i,k)\) element of \( HK \) is then
\[
= \sum_{\tau \in S_m} (-1)^{\tau} h_{\lambda_{n+1}}(X_1, ..., X_{j-1}, X_{j+1}, ..., X_n)
\]
which is zero because \( r\lambda(X_j - X_{k+1}) \leq i - (k+1) < i - k \).

(3.8) (Uniqueness Theorem, 1st version) Let \( \lambda \geq \mu \) be partitions of length \( \leq n \), such that \( q(\lambda') \leq m \).

Then
\[
s_{\lambda}(X_1, ..., X_{n-1}, ..., X_{n-m}) = (-1)^{n-m} s_{\lambda}(X_1, ..., X_n) \lambda_{n-m}.
\]

Proof: Let
\[
\xi_1 = \lambda_1 - i, \quad \eta_1 = \mu_1 - i \quad (1 \leq i \leq n),
\]
\[
\xi_j = \lambda_j - j, \quad \eta_j = \mu_j - j \quad (1 \leq j \leq n),
\]
Then the integers \( \xi_i (1 \leq i \leq n) \) and \( \eta_j (1 \leq j \leq m) \) fill up the interval \([-m, n-1]\), and so do the \( \eta_i \) and the \(-\eta_j - 1\).
The \((i_1, \ldots, i_n)\) minor of the matrix \(H\) is
\[
\det(h_{i_j, i_j}(-X_{i_j})) = s_{i_1} \cdots s_{i_n}(-X_{i_1}, \ldots, -X_{i_n}).
\]
The complementary cofactor of \((H^{*Y}) = (h_{i_j, i_j}(X_{i_j}))_{m \times n+1}\) has row indices \(-q_i - 1\) \((1 \le j \le m)\) and column indices \(-q_j - 1\) \((1 \le j \le m)\). Hence it is
\[
(-1)^{m+n} s_{i_1} \cdots s_{i_n}(-X_{i_1}, \ldots, -X_{i_n}).
\]

Since each minor of \(H\) is equal to the complementary cofactor of \((H^{*Y})\) (because \(m = 1\)) the result follows. \(\square\)

Remark. Observe that
\[
\text{rk}(X_{i_1}, \ldots, X_{i_{m+1}}) \le (\lambda_m - i) - (\lambda_{m+1} - i - 1) = \lambda_i - \lambda_{m+1} + 1.
\]
Hence (3.8) gives a duality theorem for the multi-Schur function \(s_{i_1+\ldots+i_k}(Y_1, \ldots, Y_n)\) provided that
\[
\text{rk}(Y_{i+1} - Y_i) \le \lambda_i - \lambda_{i+1} + 1 \quad (1 \le i \le n - 1).
\]

At first sight the formula (3.8) is disconcerting, because the arguments \(-X_{i+n}\), on the left are not in general the negatives of the arguments \(X_{i+n}\) on the right. However, we can use (3.6) to rewrite (3.8), as follows. As in Chapter 1, let us write the partition \(\lambda\) in the form
\[
\lambda = (\lambda_1, \ldots, \lambda_m, \lambda_{m+1}, \ldots, \lambda_n)
\]
where \(p_1 > p_2 > \ldots > p_k > 0\) and each \(m_i \ge 1\). Then in
\[
s_{i_1} \cdots s_{i_n}(-X_{i_1}, \ldots, -X_{i_{m+n}})
\]
the first \(m_1\) arguments are
\[-X_{i_1}, \ldots, -X_{i_{m_1}}
\]
which by (3.6) may all be replaced by \(-X_{i_1}\), where \(c_1 = p_1 - m_1\). The next \(m_2\) arguments are
\[-X_{i_2}, \ldots, -X_{i_{m_1+m_2}}
\]
which by (3.6) may all be replaced by \(-X_{i_2}\), where \(c_2 = p_2 - m_1 - m_2\). In general, for each \(i = 1, 2, \ldots, k\) the \(r^\text{th}\) group of \(m_i\) arguments may all be replaced by \(-X_{i_r}\), where \(c_i = p_i - (m_1 + \ldots + m_i)\). Now if
\[
X' = (q_1^*, \ldots, q_k^*)
\]
is the conjugate partition, we have \(m_1 + \ldots + m_i = q_i+1, q_k\) and \(c_i = p_i - q_i+1, q_k\) is the content of the square \(s_i = (q_{i+1}, q_k)\) in the diagram of \(\lambda\). The squares \(s_1, \ldots, s_n\) are the "salients" of the border of \(\lambda\), read in sequence from north-east to south-west. Hence the duality theorem (3.8) now takes the form

\[(3.8')\quad (\text{Duality Theorem, 2nd version}).\quad \text{With the above notation, we have}
\]
\[
x_{i_1} \cdots x_{i_n}(-X_{i_1})^m, \ldots, (-X_{i_n})^m = (-1)^{m_1} a_{i_1, \ldots, i_n}(X_{i_1})^m, \ldots, (X_{i_n})^m.
\]

Finally, if we set \(Z_i = -X_i\), \((1 \le i \le k)\) we have

\[(3.8'')\quad x_{i_1} \cdots x_{i_n}(-Z_{i_1})^m, \ldots, (-Z_{i_n})^m = (-1)^{m_1} a_{i_1, \ldots, i_n}(-Z_{i_1})^m, \ldots, (-Z_{i_n})^m
\]

provided
\[
\text{rk}(Z_{i_1}, \ldots, Z_i) \ge m_{i+1} + q_{i+1, i} \quad (1 \le i \le k - 1).
\]

Let now \(x_1, x_2, \ldots\) be independent indeterminates over \(Z\). We may regard \(Z(x_1, x_2, \ldots)\) as a \(Z\)-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
\[
\mathfrak{h}_i(X_i) = 
\]
by (3.8)
\[
\mathfrak{h}_i(X_i) = h_{i+1}(X_{i+1}),
\]
\[
\mathfrak{h}_i(X_i) = h_{i+1}(X_{i+1} + Z_1),
\]
\[
\mathfrak{h}_i(X_i) = h_i(X_{i+1}).
\]

**Proof:** Consider the generalizing functions \(\mathfrak{h}_i(X_i)\) is the coefficient of \(t^i\) in
\[
\Delta\left(\sum_{i \ge 0} \mathfrak{h}_i(X_i)t^i\right) = \Delta\left(\prod_{i = 1}^n (1 - x_id)^i\right),
\]
and
\[
\Delta\left(\prod_{i = 1}^n (1 - x_id)^i\right) = \prod_{i = 1}^n (1 - x_id) = \prod_{i = 1}^n (1 - x_id)^i.
\]

so that
\[
\Delta\left(\sum_{i \ge 0} \mathfrak{h}_i(X_i)t^i\right) = \prod_{i = 1}^n (1 - x_id)^i = \sum_{i \ge 0} h_i(X_{i+1})t^{i+1}
\]
in which the coefficients of \(t^i\) is \(h_{i+1}(X_{i+1})\).

The other two relations are proved similarly.\(\square\)
(3.10) Let 0 ∈ ℤ and let r1, . . . , rn ≥ 0. If i is such that r1 ̸= rj for all j ̸= i then
\begin{align*}
\delta_{r_1}A(X_1, \ldots, X_n) &= A_{r_1}(X_1, \ldots, X_{n}), \\
\delta_{r_1}A(X_1, \ldots, X_n) &= -A_{r_1}(X_1, \ldots, X_{n}), \\
\tau_{r_1}A(X_1, \ldots, X_n) &= A_{r_1}(X_1, \ldots, X_{n}), \\
\tau_{r_1}A(X_1, \ldots, X_n) &= -A_{r_1}(X_1, \ldots, X_{n}).
\end{align*}
where ri has its coordinate equal to i, and all other coordinates zero.

Proof: By definition, we have A = det(h1, ..., hn, X) and δi acts only on the ith row of the determinant, the entries in the other rows being symmetrical to xi, and xi,1 (because of the condition rj ̸= rj if j ̸= 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 (2.8) in the form
\begin{equation}
\tau_{r_1}A(X_1, \ldots, X_n) = (-1)^{|r_1|}\delta_{r_1}A(X_1, \ldots, X_n).
\end{equation}

Let (i, j) be a corner square of the diagram of λ, so that j = λi and i = λj. Let μ be the partitions obtained from λ by removing the square (i, j). By operating on either side of (4) with δi or τi, we obtain the same relation with μ replacing λ. Hence it is enough to show that (4) is true when λ = (m*), in that case both sides are equal to (X1 ... Xm)*, by (3.5), (3.5') and (2.9).

(3.11) Let 0 be the longest element of Sn. Then for any 0 ∈ ℤ we have
\begin{align*}
\delta_0A(X_1 + Z_1, \ldots, X_n + Z_n) &= A_{r_0}(X_1 + Z_1, \ldots, X_n + Z_n), \\
\tau_0A(X_1 + Z_1, \ldots, X_n + Z_n) &= A_{r_0}(X_1 + Z_1, \ldots, X_n + Z_n),
\end{align*}
where Z1 = z1 + ... + zn (1 ≤ i ≤ m) and the Zi are independent of x1, ..., xn.

Proof: The sequence
\begin{equation}
(n = 1, n = 2, n = 3, \ldots, n = 2, 1, 2, 3, \ldots, n = 1)
\end{equation}
is a reduced word for 0, so that
\begin{equation}
\tau_0 = \tau_{r_0} = (τ_{r_1} τ_{r_2} \cdots τ_{r_m}) (τ_{r_1} τ_{r_2} \cdots τ_{r_m})
\end{equation}
and likewise for δ0. By (3.10), x1, ..., xn applied to A(X1 + Z1, ..., Xn + Zn) will produce
\begin{equation}
A(X_1 + Z_1, X_2 + Z_2, \ldots, X_n + Z_n).
\end{equation}

Multi-Schur functions

We have next to operate on this with x1x2 ... xn, which will produce
\begin{equation}
x_1X_2 + X_1X_3 + X_1X_4 + X_2X_3 + X_2X_4 + X_3X_4 + X_5X_6 + X_5X_7 + X_6X_7.
\end{equation}

By repeating this process we shall obtain the formula for ρA. That for ρ0A is proved similarly.

Remark. Let 0 ∈ ℤ and Z1 = ... = Zm = Z = 0 in (3.11). Then by (2.5) we have
\begin{align*}
\delta_0A(x) &= δ_0A(X_1, \ldots, X_n) = A_{r_0}(X_1, \ldots, X_n), \\
\tau_0A(x) &= τ_0A(X_1, \ldots, X_n) = A_{r_0}(X_1, \ldots, X_n).
\end{align*}

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

Sergeev's formula

Let x1, ..., xn, y1, ..., yn be independent variables and let
\begin{equation}
X_1 = x_1 + \cdots + x_n, Y_1 = y_1 + \cdots + y_n
\end{equation}
for all i ≥ 1, with the understanding that xj = 0 if j > m and yj = 0 if j > n.

(3.12) (Sergeev) For all partitions λ we have
\begin{equation}
x_1X_m - Y_m = \sum_{\lambda_1 \leq m} \sigma_{\lambda_1}(x_1, y_1)/D(x)D(y)
\end{equation}
where
\begin{equation}
D(x) = \prod_{1 \leq i \leq m} (x_i - y_i),
\end{equation}
and in view of (2.16) Sergeev's formula may be restated in the form
\begin{equation}
x_1(X_m - Y_m) = \tau_{r_1}D(x_1, y_1).
\end{equation}

From (3.11) and (1) above we have
\begin{equation}
\tau_{r_1}D(x_1, y_1) = x_1X_m - Y_m.
\end{equation}


If \( \lambda = (p_1^n, \ldots, p_k^n) \), (2) can be rewritten in the form

\[
\tau_{\lambda}(x, y) = s_\lambda(x_1^n, \ldots, z_k^n)
\]

where \( Z_i = X_m - Y_m \). Since

\[
\text{rk}(Z_{i+1} - Z_i) = \text{rk}(Y_m - Y_{m+1}) = p_i - p_{i+1},
\]

the duality theorem (3.8') applies, and gives

\[
s_\lambda(x_1^n, \ldots, z_k^n) = (-1)^{\beta_{\lambda}} s_\lambda((-Z_1)^n, \ldots, (-Z_k)^n)
\]

\[
= (-1)^{\beta_{\lambda}} s_\lambda((Y_1 - X_1)^n, \ldots, (Y_k - X_k)^n)
\]

\[
= (-1)^{\beta_{\lambda}} s_k(Y_1 - X_1, Y_2 - X_2, \ldots, Y_k - X_k)
\]

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

\[
\tau_{\lambda}(x, y) = (-1)^{\beta_{\lambda}} s_\lambda(Y_1 - X_1, \ldots, Y_k - X_k)
\]

\[
= (-1)^{\beta_{\lambda}} s_k(Y_1 - X_1, Y_2 - X_2, \ldots, Y_k - X_k)
\]

\[
= s_\lambda(X_m - Y_m)
\]

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

\[
\tau_{\lambda}(x, y) = (-1)^{\beta_{\lambda}} s_\lambda(Y_1 - X_1, \ldots, Y_k - X_k)
\]

\[
= (-1)^{\beta_{\lambda}} s_k(Y_1 - X_1, Y_2 - X_2, \ldots, Y_k - X_k)
\]

\[
= s_\lambda(X_m - Y_m)
\]

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

\[
\tau_{\lambda}(x, y) = (-1)^{\beta_{\lambda}} s_\lambda(Y_1 - X_1, \ldots, Y_k - X_k)
\]

\[
= (-1)^{\beta_{\lambda}} s_k(Y_1 - X_1, Y_2 - X_2, \ldots, Y_k - X_k)
\]

\[
= s_\lambda(X_m - Y_m)
\]
(4.3) (i) \( G_w = x^1 \), \( G_1 = 1 \).

(ii) For each \( w \in \mathcal{S}_n \), \( G_w \) is a non-zero homogeneous polynomial in \( x_1, \ldots, x_n \) of degree \( \ell(w) \), of the form

\[
G_w = \sum a_\alpha x^\alpha
\]

summed over \( \alpha \in \mathbb{N}^{n-1} \) such that \( \alpha \subset \delta \) (i.e., \( \alpha_i \leq n-i \) for each \( i \)) and \( |\alpha| = \ell(w) \). Hence \( G_w \) is \( \sum a_\alpha x^\alpha \) is homogeneous of degree \( \ell(w) \). If now \( \alpha \in \mathbb{N}^{n-1} \) is such that \( \alpha \subset \delta \), then by (4.2) \( \delta \alpha \) is a linear combination of monomials \( x^\alpha \) such that \( \delta \alpha = \alpha \), if \( i \neq r, r+1 \), and

\[
\max(\delta_1 \alpha_1, \delta_n \alpha_n) \leq \max(\alpha_1, \alpha_n) - 1 \leq n - i - 1,
\]

so that \( \beta \subset \delta \). Hence the linear space \( \mathcal{H}_\alpha \) of the monomials \( x^\alpha \), \( \alpha \subset \delta \), is mapped into itself by each \( \delta_\beta \) (1 \( \leq r \) \( n-1 \)) and hence by each \( \delta_\alpha \), \( w \in \mathcal{S}_n \). Hence \( G_w \) is \( \mathcal{H}_\alpha \) for each \( w \in \mathcal{S}_n \).

(iii) \( G_w \) is symmetrical in \( x_i \), \( i \leq j \), and only if \( \delta_j G_w = 0 \), that is to say if and only if \( \delta_j a_{\alpha} = 0 \), which by (4.2) is equivalent to \( u(\alpha) < u(\alpha') \).

(iv) \( G_w \) is symmetrical in \( x_i \), \( i \leq j \), by (iii) above, but does not contain \( x_n \), hence does not contain any of \( x_{n-1} \), \( \ldots, x_n \).

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

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

\[
G_i = x_1 + x_2 + \cdots + x_i.
\]

Proof: By (4.3), \( G_i \) is a homogeneous symmetric polynomial of degree \( \ell(i) = 1 \) in \( x_1, \ldots, x_i \), hence is equal to \( \ell(x_1 + \cdots + x_i) \) for some integer \( c \). But \( \delta_i G_i = G_i = 1 \) by (4.2) and (4.3)(i), hence \( c = 1 \).

(4.5) (Stability) Let \( m \geq n \) and let \( i : \mathcal{S}_n \to \mathcal{S}_m \) be the embedding. Then

\[
G_m = G_{m(n)}
\]

Proof: By (4.3), \( G_m \) is a homogeneous symmetric polynomial of degree \( \ell(m) = 1 \) in \( x_1, \ldots, x_n \), hence is equal to \( \ell(x_1 + \cdots + x_m) \) for some integer \( c \). But \( \delta_i G_m = G_m = 1 \) by (4.2) and (4.3)(i), hence \( c = 1 \).

For all \( w \in \mathcal{S}_n \).

Proof. We may assume that \( m = n+1 \). Let \( w_0 \) be the longest element of \( \mathcal{S}_{n+1} \), then \( w_0 = \) \( w_{n+1} \), \( \ldots, w_1 \), \( w_0 \) is the longest element of \( \mathcal{S}_n \), and hence

\[
G_{w_0} = \delta_{i_{n+1}}(G_{w_{n+1}}) = \delta_{i_{n+1}}(G_{w_1}) = \delta_{i_{n+1}}(G_{w_0}) = \delta_{i_{n+1}}(G_{w_0}) = \delta_{i_{n+1}}(G_{w_0})
\]

(because \( \delta_{i_{n+1}}(G_{w_{n+1}}) = G_{w_1} \), \( G_{w_0} \) is not necessarily homogeneous of degree \( m \)).

From (4.5) it follows that \( G_m \) is a well-defined polynomial for each permutation \( w \in \mathcal{S}_n \).

If \( w \in \mathcal{S}_m \) and \( w \in \mathcal{S}_n \), we denote by \( w \times w \) the permutation

\[
(w_1, \ldots, w(m), 1, \ldots, n) + m
\]

in \( \mathcal{S}_{m+n} \). We have then

\[
G_{m+n} = G_m \cdot G_{n+n}
\]

where \( 1_m \) is the identity element of \( \mathcal{S}_m \).

Proof. We shall make use of the following fact: if \( f \) is a polynomial in \( x_1, x_2, \ldots \), and \( \delta_1 f = 0 \) for all \( i \geq 1 \), then \( f \) is a constant. For \( f \in \mathcal{P}_k = \mathbb{Z}[x_1, \ldots, x_n] \) for some \( n \), and is symmetric in \( x_1, \ldots, x_n \), because \( \delta_1 f = \cdots = \delta_{n-1} f = 0 \).

To prove (4.5) we proceed by induction on \( \ell(w) + \ell(u) \). If \( \ell(w) = \ell(u) \) then \( m = n \), \( w = 1_m \), and both sides of (4.6) are equal to 1. Let

\[
F(u, v) = G_{m+n} = \delta_{i_m} G_{m+n}
\]

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

Suppose first that \( i < m \). Then

\[
\delta_i F(u, v) = \delta_i (G_{m+n}) = \delta_i (G_m) \cdot G_{n+n}
\]

because \( \delta_i (G_{m+n}) = 0 \) by (4.2). Hence we have \( \delta_i F(u, v) = 0 \) if \( \ell(w_i) > \ell(u) \); and if \( \ell(w_i) < \ell(u) \) then

\[
\delta_i F(u, v) = \delta_i (G_{m+n}) = \delta_i (G_m) \cdot G_{n+n}
\]
which is zero by the inductive hypothesis. Likewise, if \( i > m \) we have
\[
\partial_i F(u, v) = \begin{cases} 
F(v(u), v) & \text{if } f(u) < f(v), \\
0 & \text{otherwise},
\end{cases}
\]
and so again \( \partial_i F(u, v) = 0 \) by the inductive hypothesis.

Finally, if \( i = m \) we have \( f(u) > f(v) \), because
\[
(u \times v)(m) = u(m) < m + v(1) = (u \times v)(m + 1),
\]
and therefore \( \partial_m F(u, v) \) kills \( \partial_{w_0} F_{w_0} \) and \( \partial_{w_0} F_{w_0} \); moreover, \( \partial_m F(u, v) = 0 \), because \( \partial_m F \in \mathbb{Z}[x_1, \ldots, x_m] \). Hence \( \partial_m F(u, v) = 0 \), and the proof is complete.

For certain classes of permutations there are explicit formulas for \( \Theta_n \). We consider first the case where \( w \) is dominant, of shape \( \lambda \) (so that the diagram of \( w \) coincides with the diagram of \( \lambda \)).

(4.7) If \( w \) is dominant of shape \( \lambda \), then
\[
\Theta_n = x^\lambda.
\]

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

Suppose \( w \in S_n \), \( w \neq w_0 \) and \( w \) is dominant of shape \( \lambda \). Then \( \lambda \not\supsetneq \delta \) and \( \lambda \not\supsetneq \delta \). Let \( r \geq 0 \) be the largest integer such that \( \lambda' = \lambda - i \) for \( 1 \leq i \leq r \), and let \( c = \lambda_{r+1} + 1 - n - r - 1 \). Then \( w_0 w \) is dominant of length \( f(w) + 1 \), and \( \lambda(w_0 w) = \lambda + c \), where \( c \) is the vector whose \( \alpha \)-th component is 1 and all other components zero. Hence we have
\[
\Theta_n = \partial_m \Theta_{w_0} = \partial_m (x^\lambda) = x^\lambda,
\]
because \( \lambda = \lambda_{\alpha+1} \).

Conversely, every monomial \( x^\lambda \) (where \( \lambda \) is a partition) occurs as a Schubert polynomial, namely as \( \Theta_n \), 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 \( \lambda \), then \( \Theta_n \) is the Schur function \( s_{\lambda}(w) \), 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)(ii), \( G_1 = 1 \). Then \( r \geq 1 \) and the code of \( w \) is
\[
(w(1) - 1, w(2) - 2, \ldots, w(r) - r)
\]
so that \( \lambda = (w(r) - r, \ldots, w(2) - 2, w(1) - 1) \). Let \( w = w_0 \) be the longest element of \( S_n \). Then \( w_0 = (w(r), \ldots, w(1), w(r + 1), w(r + 2), \ldots) \)
is dominant of shape \( \lambda + \delta \), where \( \delta = (r - 1, r - 2, \ldots, 0) \), and \( f(w_0) = f(w) + f(\delta) \). Hence
\[
\Theta_n = \partial_m \Theta_{w_0} = \partial_m (x^{\lambda+\delta}) = x_1(X_r)
\]
by (4.2), (4.7) and (2.11).

Conversely, every Schur function \( s_{\lambda}(w) \) (where \( \lambda \) is a partition of length \( \leq r \)) occurs as a Schubert polynomial, namely as \( \Theta_n \) where \( w \) is the permutation with code \( c(w) = \lambda \).

More generally, let \( w \) be vexillary with shape \( \lambda = (\lambda_1, \ldots, \lambda_m) \) (where \( m = f(w) \)) and flag \( \phi = (\phi_1, \ldots, \phi_m) \) (Chapter I). Then \( \Theta_n \) is a multi-Schur function (Chapter III), namely
\[
\Theta_n = \phi_{\lambda}(X_{\phi_1}, \ldots, X_{\phi_m})
\]
where \( X_i = x_1 + \cdots + x_i \), for each \( i \geq 1 \).

Proof: The idea is to convert \( w \) systematically into a dominant permutation. Recall ((1.23), (1.24)) that if \( c(w) = (\phi_1, \ldots, \phi_m) \), and \( c_i \leq c_{i+1} \) for some \( i \geq 1 \), then \( f(w_0) = f(w) + 1 + \phi_i \).

As in Chapter I let
\[
\lambda(w) = (\phi_1^{\#}, \ldots, \phi_m^{\#})
\]
where \( \phi_1 < \cdots < \phi_m > 0 \) (and each \( \phi_i^{\#} \geq 1 \)), and let
\[
\phi(w) = (\phi_1^{\#}, \ldots, \phi_m^{\#})
\]
where \( f_1 \leq \cdots \leq f_k \).

Consider first the terms equal to \( p_k \) in the sequence \( c(w) \). They occupy the positions \( f_k - m_k + 1, \ldots, f_k \). We shall use \( \phi \) to move them all to the left until they occupy the first \( m_k \) positions, by multiplying \( w \) on the right by
\[
w_1 = (f_{k+1}, \ldots, f_k)(f_{k+1}, \ldots, f_k) \cdots (f_{k+1}, \ldots, f_k).
\]
Let \( w_1 = w_{01} \). In the code of \( w_1 \), the first \( m_k \) entries will be equal to \( p_k + f_k - m_k \); the shape of \( w_2 \) is
\[
\lambda^{(1)}(w) = (p_{k1} + f_{k1} - m_{k1}, p_{k2}^{\#}, \ldots, p_{km_k}^{\#})
\]
Notes on Schubert Polynomials

and it follows from the description (1.38) of vexillary codes that the terms equal to $p_3$ 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 $w_2$ on the right by

$$w_2 = (x_{f_2-m_2+1} \cdots x_{m_1+m_2}) \cdot (x_{f_2-m_2+1} \cdots x_{m_1+m_2})^{-1} \cdot (x_{f_2} \cdots x_{m_1}).$$

Let $w_1 = w_2 w_2$; the code of $w_2$ starts off with $m_2$, entries to $p_1$, then $m_2$ entries equal to $p_2 + f_2 - m_2$, $m_3$ entries equal to $p_3 + f_3 - m_2$, then $m_3$ entries equal to $p_3 + f_3 - m_2$. The shape of $w_2$ is

$$\lambda(w_2) = ((p_1 + f_1 - m_2)^{m_1}, (p_2 + f_2 - m_2)^{m_2}, \ldots, (p_k + f_k - m_2)^{m_k}).$$

and the terms equal to $p_2$ in the sequence $c(w_2)$ will occupy the positions $f_4 - m_3 + 1, \ldots, f_4$.

We continue in this way; at the $n$th stage we define $w_n = w_{n-1} w_n$, where

$$w_n = (x_{f_{n+1} - m_{n+1} - \cdots - m_1}) \cdot (x_{f_{n+1} - m_{n+1} - \cdots - m_1})^{-1} \cdot (x_{f_1} \cdots x_{m_1}).$$

and $w_n$ has shape

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

where $a_1 = f_1 - (m_1 + \cdots + m_k) \geq 0$ by (1.36). Notice also that

$$(p_{n+1} + a_{n+1}) - (p_1 + a_1) = (m_1 + \cdots + m_k) - (f_{n+1} - f_1) \geq 0$$

by (1.37).

Finally we reach $w_k = w_1 \cdots w_k$, which is dominant with shape (and code)

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

We have

$$\ell(w_1) = |w| = \sum m_i,$$

$$\ell(w_k) = \lambda(w_k) = \sum m_i (p_i + a_i),$$

and

$$\ell(w) = a_i m_i \quad (1 \leq i \leq k)$$

so that

$$\ell(w_k) = \ell(w) + \sum_{i=1}^{k} \ell(w_i)$$

and therefore, since $w = w_k w_1 \cdots w_k^{-1}$,

$$G_w = \partial_{\mu} \partial_{\lambda} G_{w_k}$$

by (4.2). Now by (4.6) and (3.15') we have

$$G_{\mu} = s_{\mu} = s_{\lambda} = s_{\lambda} \cdot G_{\lambda}$$

where $\mu = m_1 + \cdots + m_k = \lambda(\lambda)$. Hence by repeated use of (2.10) we obtain

$$G_{\lambda(w_k)} = \partial_{\mu} G_{\lambda(w_k)} = s_{\lambda(w_k)} = s_{\lambda(w_k)} \cdot G_{\lambda(w_k)}$$

by virtue of (2.8). If we now operate with $\partial_{\mu}$, we shall obtain in the same way

$$G_{\lambda(w_k)} = \partial_{\mu} G_{\lambda(w_k)} = s_{\lambda(w_k)} \cdot G_{\lambda(w_k)}$$

and so finally

$$G_w = s_{\lambda(w_k)} \cdot (X_{f_{k+1}} \cdots (X_{f_k})^{m_k})$$

**Remarks.**

1. As in Chapter I, let

$$X = (q_1^{m_1}, \ldots, q_k^{m_k})$$

be the conjugate partition, so that

$$m_1 + \cdots + m_k = q_{k+1}$$

and therefore

$$p_i + a_i = p_i + a_i = q_{k+1} + i \quad (1 \leq i \leq k)$$

by (1.41), where $(p_1^{a_1}, \ldots, p_k^{a_k})$ is the flag of $w^{-1}$. Thus

$$\mu = \lambda(w_k) = (p_1^{a_1}, \ldots, p_k^{a_k})$$

by (4.10).

2. The result (4.9) admits a converse. If $\lambda = (p_1^{a_1}, \ldots, p_k^{a_k})$ 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 (3.9'), namely

$$0 \leq f_1 + a_i \leq f_2 + a_2 + \cdots + a_{k+1} \leq f_3 + a_3 + \cdots + a_{k+2} + \cdots + a_k \quad (1 \leq i \leq k - 1),$$

is the Schubert polynomial of a vexillary permutation, namely the permutation with shape $\lambda$ 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 $\phi$ coincide with those of (1.37). The conditions (1.36), namely

$$f_i \geq m_1 + \cdots + m_i \quad (1 \leq i \leq k)$$
Let $H_n$ denote the additive subgroup of $P_n = \mathbb{Z}[x_1, \ldots, x_n]$ spanned by the monomials $x^\alpha, \alpha \subseteq \Delta_n = (n-1, n-2, \ldots, 0)$.

(4.11) The Schubert polynomials $\Theta_w, w \in S_n$, form a $\mathbb{Z}$-basis of $H_n$.

Proof: By (4.3) each $\Theta_w$ lies in $H_n$. If

$$
\sum_{\alpha \subseteq \Delta_n} a_\alpha \Theta_\alpha = 0 \quad (a_\alpha \in \mathbb{Z})
$$

is a linear dependence relation, then by homogeneity we have

(1)

$$
\sum_{\alpha \subseteq \Delta_n} a_{\alpha} \Theta_\alpha = 0
$$

for each $p \geq 0$, and by operating on (1) with $\Theta_\alpha$ we see that $a_\alpha = 0$. Hence the $\Theta_\alpha$ are linearly independent and hence form a $\mathbb{Q}$-basis of $H_n \otimes \mathbb{Q}$. It follows that each monomial $x^\alpha, \alpha \subseteq \Delta_n$, can be expressed in the form

(2)

$$
x^\alpha = \sum_{\alpha \subseteq \Delta_n} \lambda_{\alpha} \Theta_\alpha
$$

with rational coefficients $\lambda_{\alpha}$; by operating on (2) with $\Theta_\alpha$ we have $\lambda_{\alpha} = \delta_\alpha x^\alpha$, and hence the $\lambda_{\alpha}$ are integers.]

From (4.11) it follows that

(4.12) The $\Theta_w, w \in S_n$, form a $\mathbb{Z}$-basis of $P_n = \mathbb{Z}[x_1, x_2, \ldots]$. 

Proof: Let $x^\alpha$ be a monomial in $P_n$. Then $\alpha \subseteq \Delta_n$ for sufficiently large $n$, hence $x^\alpha$ is a linear combination of the $\Theta_\alpha$.

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 $\Theta_w, w \in S(n)$, form a $\mathbb{Z}$-basis of $P_n$.

Proof: By (4.3)(iii) we have

$$
\Theta_w \in P_n \quad \iff \quad \delta_w \Theta_\alpha = 0 \text{ for all } m > n
$$

$$
\iff \quad w \in S(n).
$$

Let $P'_n \subseteq P_n$ be the $\mathbb{Z}$-span of the $\Theta_w, w \in S(n)$. If $P'_n \neq P_n$, choose $f \notin P_n - P'_n$; by (4.12) we can write $f$ as a linear combination of Schubert polynomials, say

(1)

$$
f = \sum_w a_w \Theta_w
$$

where there is at least one term with $a_w \neq 0$ and $w \notin S(n)$. Hence for some $m > n$ we have $\delta_w \Theta_\alpha = \Theta_{w \alpha}$, and since $\delta_w f = 0$ 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 $\eta: P_n \to \mathbb{Z}$ be the homomorphism defined by $\eta(w) = 0 \quad (1 \leq i \leq n)$. 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)

$$
f = \sum_{w \in S(n)} \eta(\Theta_w) \Theta_w
$$

Proof: By (4.12) and linearity, it is only necessary to verify this formula when $f$ is a Schubert polynomial $\Theta_w, w \in S(n)$, and then it follows from (4.12) and (4.3)(ii) that $\eta(\Theta_\alpha \Theta_\beta)$ is equal to 1 when $w = w$ and is zero otherwise.]

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

$$
\Theta_w = \sum (a_i - a_j) \Theta_{w_i w_j},
$$

where $t_i$ is the transposition that interchanges $i$ and $j$, and the sum is over all pairs $i < j$ such that $f(w_{i j}) = f(w) + 1$.

Proof: The polynomial $\Theta_w$ is homogeneous of degree $f(w) + 1$, and hence by (4.14) we have

$$
\Theta_w = \sum \delta_w (f(\Theta_w)) \Theta_w
$$

summed over $w$ of length $f(w) + 1$. Now by (2.13)

$$
\delta_w (f(\Theta_w)) = f(\Theta_w) + \sum (a_i - a_j) \delta_{w_i w_j} \Theta_w
$$

summed over $i < j$ such that $f(w_{i j}) = f(w) + 1 = f(w)$. It follows that $\delta_w (f(\Theta_w)) = a_i - a_j$ if $w = w_{ij}$, and in zero otherwise.]

In particular,

(4.15')

$$
\Theta_w = \sum x(\alpha) \Theta_w
$$

summed over transpositions $w = t_{i j}$ such that $f(w_{i j}) = f(w) + 1$, where $x(\alpha) = \{1$ or $-1$ according as $i < j$ or $i > j\}$

(4.15") (Monk's formula) $\Theta_w \Theta_w = \sum \Theta_w$ summed over transpositions $w = t_{i j}$ such that $i < j$ and $f(w_{i j}) = f(w) + 1$.]

\[\begin{align*}
\text{(4.15')} & \\
\text{(4.15") (Monk's formula)} & \\
\end{align*}\]
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 \( v \) is the last descent of \( w \), so that \( w(r) > w(r + 1) \) and \( w(r + 1) < w(r + 2) \). Choose the largest \( s > r \) such that \( w(r) > w(s) \) and let \( v = w_r \). Then from (4.15) applied to \( v \) we have

\[ x_s \Theta_v = \Theta_v - \sum_{j} \Theta_{w'} \]

summed over all permutations \( w' = w_r \) where \( q < r \) and \( f(w') = f(w) + 1 = f(w) \). Hence \( w'(q) = w(q) > w(q) = w(q) \), 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 \( f(u) = f(w') = p \) then \( w' \) precedes \( u \) if and only if for some \( i \geq 1 \) we have

\[ w'(j) = w(j) \] for \( j < i \), 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 \( \Theta_w \) can be expressed in the form

\[ \Theta_w = x_s \Theta_v + \sum_{j} \Theta_{w'} \]

where \( r \) is the last descent of \( w \), \( f(v) = f(w) + 1 \) 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 \( w \), \( \Theta_w \) is a polynomial in \( z_1, z_2, \ldots \) with positive integral coefficients.

For we may assume, as inductive hypothesis, that (4.17) is true for all permutations \( v \) such that either \( f(v) = f(w) \), or \( f(v) = f(w) \) 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 \( \Delta_p^{(p)} \) by (4.7).)

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

\[ \Theta_w(z_1, \ldots, z_n) = \sum_{a} x_{a} \Theta_{(a)}(z_{m+1}, \ldots, z_{n}) \]

summed over \( a \in S_{n-m} \) and \( v \in S_{n-m}^{m} \).
Notes on Schubert Polynomials

[4.20] For each permutation \( w \) we have

\[
\sigma_w = \sum_{(1 \leq i < j \leq n)} x_i^j
\]

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

**Example.** If \( w = (132) \), then \( D(w) \) consists of the diagram:

and \( \theta_w = x_1 x_2 + x_1 x_3 + x_2 x_3 + x_2 x_1 + x_3 x_2 + x_3 x_1 + x_1 x_3 + x_2 x_1 + x_3 x_2 \).

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 \( w \) arbitrary \([11]\), but open in general.

The shift operator

Let \( f \in \mathbb{R} \) and let \( m \geq 0 \). Then

\[
r^m f = r_m f = \partial_1 \ldots \partial_m (x_1 \ldots x_m f)
\]

(4.21)
is independent of \( m \), because \( \sigma_m f = f \) 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 \( r : \mathbb{R} \to \mathbb{R}_{n+1} \) is called the shift operator. For example, we have

\[
r^1 = \partial_1 (x_1) = x_1 + x_2
\]

and for \( i \geq 2 \),

\[
r^i = \partial_1 \ldots \partial_i (x_1 \ldots x_i + x_2 x_3 + \ldots + x_i x_{i+1} + x_{i+1} x_{i+2} + \ldots + x_n)
\]

(4.22)
so that by (4.4)

\[
\tau \theta_w = \tau (x_1 + \ldots + x_n) = x_1 + \ldots + x_{n+1} = \theta_{w+1}.
\]

More generally, (4.22) For all permutations \( w \),

\[
\tau \theta_w = \theta_{u_w} = \theta_{v_w}
\]

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

**Proof.** For each \( r \geq 1 \) let \( w^R \) be the longest element of \( S_n \) and let \( \delta_r = (r-1, r-2, \ldots , 1) \). Then if \( w \in S_n \) we have

\[
\tau \theta_w = \partial_1 \ldots \partial_n (x_1 \ldots x_n w^{-1}(x_1)) = \partial_1 \ldots \partial_n (x_1 \ldots x_n w^{-1}(x_1))
\]

Now \( x_1 \ldots x_n \) is the cycle \( 1 - 2 \ldots - n + 1 = 1 \), and hence

\[
x_1 \ldots x_n w^{-1}(u)^R = (1 \times u)^{-1} u^R
\]

so that

\[
\tau (x_1 \ldots x_n w^{-1}(u)^R) = \tau (x_1 \ldots x_n) + \tau (u^{-1} u^R)
\]

and therefore by (2.7) we have

\[
\tau \theta_w = \partial w(1 \times u)^{-1} u^R = \theta_{u(w)} = \theta_{v(w)}
\]

(4.23) Let \( n \in \mathbb{N}^n \) and \( 0 \leq p_1 \leq \cdots \leq p_n \). Then

\[
\tau_{p_1}(X_{p_1}, \ldots , X_{p_n}) = s_{p_1}(X_{p_1+1}, \ldots , X_{p_n+1}).
\]

**Proof.** Since \( \tau = \tau_1 \tau_2 \ldots \tau_n \), this follows from (3.10) \( \parallel \)

(4.24) We have

\[
\partial^r = 0
\]

for \( 1 \leq i \leq r \).

**Proof.** By (4.12) it is enough to show that \( \partial_i \tau = \theta_w = 0 \) for all permutations \( w \), and this follows from (4.22) and (4.2) \( \parallel \)

For each \( r \geq 1 \) let \( \rho : \mathbb{R} \to \mathbb{R} \) be the homomorphism defined by

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

(4.25) Let \( u^R \) be the longest element of \( S_n \). Then

\[
\tau_{u^R}(f) = \rho \tau^R(f)
\]

for all \( f \in \mathbb{R} \).

**Proof.** By linearity we may assume that \( f = x_0 \) where \( 0 \in \mathbb{N}^n \). Since \( x^0 = s_{p_1}(X_{p_1+1}, \ldots , X_{p_n}) \) by (3.4), we have

\[
\tau^R(f) = s_{p_1}(X_{p_1+1}, \ldots , X_{p_n})
\]
by (4.23), and hence
\[ f_k^*(x_0) = x_k f(x_1, \ldots, x_n) \]
which is equal to \( x_k = f^*(x^*) \) by (2.16).]

Transitions

A transition is an equation of the form
\[ T(w, r) \]
where \( r \geq 1 \), and \( w \) and \( r \) are permutations and \( \Phi \) is a set of relations. It exists only for certain values of \( r \), depending on \( w \). An example is (4.10), in which \( r \) is the last descent of \( w \).

By (4.15) we have
\[ \phi = \sum_{r \in \Phi} \phi \]
summed over transpositions \( t = (u, v), \) such that \( f(u, \perp) = f(u, \perp) \), where \( f(u, \perp) \) is the sign of \( u \). So for \( T(w, r) \) to hold there must be exactly one \( r > r \) such that
\[ f(u, \perp) = f(u, \perp) + 1 \]
and
\[ w = u \tau \]

Consider the graphs \( G(u) \) and \( G(v) \) of \( w \) and \( v \). They differ only in rows \( r \) and \( j \):

\[ G(w) \quad G(v) \]

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 \( s(j) > s(r) \) and \( A \cap G(u) = \emptyset \), and this will be the case if and only if \((A \cup B \cup C) \cap G(u) = \emptyset \). Since \((A \cup B \cup C) \cap G(u) = (A \cup B) \cap G(u) \), it follows that

\[ A \cap G(u) = \emptyset. \]

Thus (4.26) is the relation \( T(w, r) \) if and only if
\[ (A \cup B \cup C) \cap G(u) = \emptyset. \]
Notes on Schubert Polynomials

We may notice directly one corollary of (1.26). Let

\[ \Theta_\nu(1) = \Theta_\nu(1, 1, \ldots) \]

be the number of monomials in \( \Theta_\nu \), each counted with its multiplicity. (By (4.17), \( \Theta_\nu \) is a positive sum of monomials.) If \( T(u, r) \) is a transition, we have

\[ \Theta_\nu(1) = \Theta_\nu(1) + \sum_{c \in C} \Theta_c(1) \]

and also, by (4.20)

\[ \Theta_{\nu^+}(1) = \Theta_{\nu^+}(1) + \sum_{c \in C} \Theta_c(1). \]

From these two relations it follows, by induction on \( \nu(u) \) and on the integer \( \Theta_\nu(1) \), that

(4.30)

or in other words that \( \Theta_\nu \) and \( \Theta_{\nu^+} \) each contain the same number of monomials. So if Kohnert's algorithm (4.20) is true, we should have

\[ \text{Card } K(\Theta(u)) = \text{Card } K(\Theta(u^{-1})). \]

Doubtless the combinatorialists will seek a "bijective" proof of this fact.

Let \( T(u, r) \) be a transition and let \( u \in \Phi(u, r) \). Consider again the graphs of \( u \) and \( v \):

```
      i      j
     |      |
    M | 0 | N
     | 0 | 0
     r  j
```

Then we have

(4.31)

\[ c_\nu(u) = n + p, \quad c_\nu(v) = n + p + 1, \]

\[ c_\nu(s) = n + p + 1, \quad c_\nu(r) = n, \]

and \( c_\nu(v) = c_\nu(s) \) if \( k \neq i, r \). In particular, \( c_\nu(v) > c_\nu(s) \) for all \( v \in \Phi(u, r) \).

Schubert polynomials (I)

Proof. \( c_\nu(u) \) is the number of positive integers \( k > i \) such that \( u(k) < u(i) \). 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 \( \Phi \). Then from (4.31) we have

\[ c_\nu(u) = c_\nu(v) \quad \text{and} \quad c_\nu(s) = c_\nu(v) \quad \text{Hence in this case } \lambda(u) = \lambda(v) \quad \text{and therecfe } \lambda(v) = \lambda(u). \]

If on the other hand \( m > 0 \), there are two possibilities:

- either

\[ c_\nu(u) > c_\nu(v) \geq c_\nu(s) > c_\nu(v), \]

- or

\[ c_\nu(v) > c_\nu(u) \geq c_\nu(s) > c_\nu(v). \]

In both cases it follows that \( \lambda(v) \) is of the form \( R^m \lambda(u) \), where \( R \) is a raising operator and \( m \geq 1 \).

Hence \( \lambda(v) > \lambda(u) \) for the dominance partial ordering on partitions), and we have proved

(4.22) \( \lambda(v) > \lambda(u) \) for all \( u \in \Phi(u, r) \), with equality if and only if \( v \) is the leader of \( \Phi \).

Recall (1.26) that for any permutation \( u \) we have

\[ \lambda(u) > \lambda(u^{-1}). \]

Hence for \( v \in \Phi(u, r) \) we have

(4.33)

\[ \lambda(u) > (v) \geq \lambda(u^{-1}) \geq (v) \]

by (4.29) and (4.22). Moreover, at least one of the inequalities \( (\cdot) \) is strict unless \( v \) is the leader of \( \Phi(u, r) \) and \( u^{-1} \) is the leader of \( \Phi(u^{-1}, r) \) (in the notation of (4.29)). In the notation of the diagram preceding (4.27) this means that

\[ (A' \cup B' \cup C') \cap G(u) = \emptyset \]

and hence, as in the proof of (1.26), that \( \text{Card } \Phi \leq 1 \).

(4.34) \( \lambda(u) > (v) \) is a transition with \( v \) a core, then \( \Phi(u, r) \) is either empty or consists of one rectilinear permutation.

Proof. Suppose that \( \lambda(u) \) is not empty, and let \( u \in \Phi \). By (1.27) we have \( \lambda(u) = \lambda(u^{-1}) \), and hence all the inequalities in (4.33) are equalities. Thus \( v \) is rectilinear, and by the remarks above it is the only member of \( \Phi \).
Notes on Schubert Polynomials

(4.35) Let \( T(w, r) \) be a transition with \( r > d_r(w) \). Then

\[ d_r(v) \geq d_r(w) \]

for all \( v \in \Phi(w, r) \).

Proof: As before, let \( v = w_i \) with \( i < r \), and let \( d_r(w) = d \). We have to show that

\[ c_i(v) \leq \cdots \leq c_r(v) \]

where \( c_i(v) = c_i(w) \) for \( 1 \leq k \leq d \), and

\[ c_{r_1}(v) = c_{r_1}(w) \leq c_{r_2}(w) < c_{r_3}(w) \]

by (4.31), so that \( c_{r_1}(v) < c_{r_1}(w) \).

We distinguish three cases:

(a) \( i = d \), so that \( d \leq i \leq 1 \) and therefore \( c_i(v) = c_i(w) \) for \( 1 \leq k \leq d \).

(b) \( i = d + 1 \). In this case we have \( c_i(v) = c_i(w) \) for \( 1 \leq k \leq d - 1 \), and

\[ c_{r_1}(v) = c_{r_1}(w) \leq c_{r_2}(w) < c_{r_3}(w) \]

by (4.28), so that \( c_{r_1}(v) < c_{r_1}(w) \).

(c) \( i > d \). Since \( d < i \leq r \) and \( c_i(w) \leq c_i(w) \), hence \( u(i) \geq u(i) \). The diagram on p. 58 shows that \( u(i+1) > u(i) \), or equivalently \( u(i+1) > u(i) \), so that \( c(i) \leq c_i(w) \).

Hence:

\[ c_{r_1}(v) = c_{r_1}(w) \leq c_{r_2}(w) < c_{r_3}(w) \]

and therefore

\[ c_{r_1}(v) < c_{r_1}(w) \leq c_{r_2}(w) < c_{r_3}(w) \]

Since the sequences \( c_1(v), \ldots, c_{r_2}(v) \) and \( c_1(w), \ldots, c_{r_2}(w) \) differ only in the \( r \)-th place, we have

\[ c_1(v) \leq \cdots \leq c_{r_2}(v) \]

as required.

The maximal transition for \( w \) in \( T(w, d_r(w)) \). Let us temporarily write \( w \rightarrow v \) to mean that \( v \in \Phi(w, d_r(w)) \).

(4.36) Suppose the:

\[ w = w_0 \rightarrow w_1 \rightarrow \ldots \rightarrow w_n \]

is a chain of maximal transitions in which none of the \( w_i \) is Grassmannian. Then

\[ p < (d_i(w) - d_i(w))(f(u)) \]

for some \( 1 \leq i \leq n \).

Proof: For any permutation \( v \), let \( c(v) = d_i(w) - d_i(v) \). Also let \( f(v) \) denote the last nonzero term in the sequence \( c(v) \), i.e. \( f(v) = c_{r_2}(v) \). Recall that \( v \) is Grassmannian if and only if it has only one descent, that is to say if and only if \( c(v) = 0 \).

Schubert polynomials (I)

From (4.35) we have

\[ d_i(w_k) = d_i(w_{k+1}) \]

for \( 1 \leq k \leq p \), and from (4.31) we have

\[ c_i(w_k) < c_i(w_{k+1}) \]

where \( r = d_i(w_k) \), hence \( d_i(w_k) \leq d_i(w_{k+1}) \) and therefore

\[ c_i(w_k) < c_i(w_{k+1}) \]

Moreover, if \( c_i(w_k) = c_i(w_{k+1}) \) we must have \( d_i(w_k) = d_i(w_{k+1}) \) and hence by (1)

\[ f(v_k) < f(v_{k+1}) \]

It follows that the \( p + 1 \) points \( (w_k, p) \) are distinct. Since they all satisfy \( 1 \leq x_k \leq c_i(w) \) and \( 1 \leq y_k \leq f(w) \), we have \( p + 1 \leq c_i(w) \), 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 \( \Phi(d_i(w), f(w)) \) in (4.36).

The rooted tree \( T_w \) of a permutation \( w \) is defined as follows:

(i) if \( w \) is Grassmannian, then \( T_w = \{w\} \);

(ii) if \( w \) is not Grassmannian, take the maximal transition for \( w \):

\[ \Theta_w = \Theta_w + \sum_{i \in I} \Theta_i \]

where \( \Theta \) is empty, replace \( \Theta \) by \( 1 \times \Theta \) as explained above. To obtain \( T_w \), join \( w \) by an edge to each \( v \in \Phi \), and attach to each \( v \) its tree \( T_v \).

By (4.36), \( T_w \) is a finite tree, and by construction all its endpoints are vexillary permutations of length \( f(w) \). It follows from (4.38) that \( w \rightarrow x \rightarrow \ldots \rightarrow w_n \). Then \( T_w \) depends (up to isomorphism) only on the diagonal equivalence class (Chapter I) of the permutation \( w \).

Recall that \( \rho_m : \Theta_m \rightarrow \Theta_m \) is the homomorphism defined by \( \rho_m(x_k) = x_k \) if \( 1 \leq i \leq m \), and \( \rho_m(x_k) = 0 \) if \( i > m \).

(4.37) Let \( \Sigma \) be the set of endpoints of \( T_w \). Then if \( m \leq d_i(w) \) we have

\[ \rho_m(\Theta_w) = \sum_{\Theta_\alpha (\Sigma)} \Theta_\alpha (\Sigma) \]


Notes on Schubert Polynomials

Proof: If \( w \) is vexillary we have \( \rho_m(\Theta_w) = s_{\lambda(w)}(X_m) \) by (4.1), since \( d_i(w) = d_i(\lambda) \geq m \). If \( w \) is not vexillary, it follows from the maximal transition \((*)\) above that

\[
\rho_m(\Theta_w) = \sum_{\kappa \leq \lambda} \rho_m(\Theta_{\kappa})
\]

since \( r = d_i(\lambda) > d_i(\kappa) \geq m \). The result now follows by induction on \( \text{Card}(T_w) \).

Multiplication of Schur functions

Let \( \mu, \nu \) be partitions and let \( u \in S_n, u' \in S_n \) be Grassmannian permutations of shapes \( \mu, \nu \) respectively. Let \( w = u \circ u' \in S_{n+1} \), so that by (4.6) and (4.8)

\[
\Theta_w = \Theta_u \circ \Theta_{u'}
\]

where \( r = d_i(u) \) and \( s = n + d_i(u') \). Hence if \( m \leq r \) we have

\[
\rho_m(\Theta_w) = s_{\lambda(u')}s_{\lambda(u)}
\]

and so by (4.37)

\[
s_{\lambda(u')}s_{\lambda(u)}(X_m) = \sum_{V \in \mathcal{Y}_W} s_{\lambda(V)}(X_m)
\]

where \( V \) is the set of endpoints of the tree \( T_w \). Here the integer \( m \) can be arbitrarily large, because we can replace \( w \) by \( 1 = w \) for any positive integer \( k \). Consequently we have

(4.38)

\[
s_{\lambda(u')}s_{\lambda(u)} = \sum_{V \in \mathcal{Y}_W} s_{\lambda(V)}
\]

where \( V \) is the set of endpoints of the tree \( T_{u \circ u'} \), and \( u \) (resp. \( u' \) ) is Grassmannian of shape \( \mu \) (resp. \( \nu \)).

The same argument evidently applies to the product of any number of Schur functions. If \( \mu^{(1)}, \ldots, \mu^{(k)} \) are partitions, let \( u_i \in S_n \) be a Grassmannian permutation of shape \( \mu^{(i)} \), for each \( i = 1, \ldots, k \) (so that \( n_i \geq \ell(\mu^{(i)}) + \ell(\mu^{(j)}) \)) and let \( w = u_1 \circ \cdots \circ u_k \). Then

(4.39)

\[
\sum_{V \in \mathcal{Y}_W} s_{\lambda(V)}
\]

where \( V \) is the set of endpoints of the tree \( T_{u_1 \circ \cdots \circ u_k} \).

In particular, suppose that each \( \mu^{(i)} \) is one-partition, say \( \mu^{(i)} = \mu_i \), so that the left-hand side of (4.39) becomes \( \lambda_1 \lambda_2 \cdots \lambda_k \). Correspondingly, each \( u_i \) is a cycle of length \( n_i + 1 \), namely \( u_i = (n_i + 1, 1, 2, \ldots, n_i) \). Now [M, Ch.I, §6] the coefficient of a Schur function \( s_\lambda \) in \( \lambda_{n_1} \) is the Kostka number \( K_{\lambda, \mu} \). Hence we have

(4.40)

\[
K_{\lambda_{n_1}} \text{ as the number of endpoints of shape } \lambda \text{ in the tree of } w \equiv u_1 \times u_2 \times \cdots
\]
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. Kohno. 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 \((r+1,j) \notin D_{r,r+1}\). If \(i = k\) or if \((r,j+1) \notin D_{r,r+1}\), let \(D_1\) be the diagram obtained from \(D\) by replacing the element \((r+1,j)\) by \((r,j)\).

Now suppose instead that \((r,j) \notin D_{r,r+1}\). We say that the point \((r,j)\) is \(r\)-fixed with respect to \(D(w)\) if the number of elements of \(D\) in the column \(j\) and in the rows \(r\) 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 \((r+1,j+i) \notin D_{r,r+1}\), let \(D_1\) be the diagram obtained from \(D\) by replacing the element \((r,j)\) by \((r+1,j)\).

In both cases we say that the diagram \(D\) is obtained from \(D\) by a "B-move" (with respect to \(D(w)\)). For example let \(D\) be such that \(D_{2,3}\) is the following:

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
3 & 4 & \_ & \_ \\
\_ & \_ & \_ & \_ \\
\_ & \_ & \_ & \_ \\
\end{array}
\]
Notes on Schubert Polynomials

For this case \( j(r, D) = (2,5,8,9) \). We can perform on this diagram a B-move in columns 2, 5 or 9 and obtain, respectively, the following diagrams:

\[ \begin{array}{cccc}
1 & 2 & 3 & 4 \\
5 & 6 & 7 & 8 \\
9 & 10 & 11 & 12 \\
\end{array} \]

The element in column 8 is not allowed to move since \( r + 1 \geq 9 \). Let \( D(u) \) denote the set of all diagrams (including \( D(u) \)) obtainable from \( D(v) \) by any sequence of B-moves.

Next for \( D \in D(u) \) let \( p^D \) denote the monomial \( x_1^{p_{11}} x_2^{p_{22}} \cdots \) where \( a_i \) is the number of elements of \( D \) in the \( i \) row. For any permutation \( w \) we shall have the following theorem.

\[ \theta_w = \sum_{D \in D(u)} p^D. \]

To prove this we will proceed by reverse induction on \( f(u) \). If \( w = w_k \) (the longest element of \( S_{r+1} \)) then (B.1) holds since \( D(w_k) \) contains only the element 1. \( D(w_k) \) and \( p^{D(w_k)} = x^1 \). On the other hand from (4.3), \( \theta_w = x^1 \). Now if \( w \neq w_k \) then let \( r = \min \{ i : w(i) < w(i+1) \} \). From (4.2) we have

\[ \theta_w = \bar{\theta}_w \theta_{w_{r+1}}. \]

Let \( v = w_k \). By the induction hypothesis equation (B.1) holds for \( \bar{\theta}_w \). The induction step will be to "apply" the operator \( \bar{\theta}_w \) to the diagrams in \( D(v) \) to this end we need more tools.

For the moment let us fix \( D \in D(v) \). Let \( a = a_i(D) \) and \( b = a_{i+1}(D) \) respectively be the number of elements of \( D \) in the \( r \) and \( r + 1 \) rows. We have

\[ \bar{\theta}_w D^D = \bar{\theta}_w \cdots \bar{\theta}_w D_{i+1} \cdots \bar{\theta}_w D \begin{cases} \sum_{i=0}^{r-1} \cdots \sum_{j=0}^{r-1} \cdots \sum_{k=0}^{r-1} & \text{if } a > b, \\
0 & \text{if } a = b, \\
\sum_{i=0}^{r-1} \cdots \sum_{j=0}^{r-1} \cdots \sum_{k=0}^{r-1} & \text{if } a < b. \end{cases} \]

This suggests we define the operator \( \bar{\theta}_w \) directly on the diagram \( D \). For this we need only to concentrate our attention on the rows \( r \) and \( r + 1 \) of \( D \). Let \( j(r, D) = (j_1, j_2, \ldots, j_k) \). Notice that in all columns \( j \leq w(r) \) of \( B_{r+1} \) there are exactly two elements and in column \( w(r) = j_k \) of \( B_{r+1} \) there is exactly one element in position \( (r, j_k) \). We shall now reduce the sequence of indices \( j(r, D) \) according to the following rule. Let \( B_{r+1} = (j_1, j_2, \ldots, j_k) \) and remove from \( B_{r+1} \) all pairs \( j_a, j_{a+1} \) for which \( (r, j_k) \in D \) and \( (r + 1, j_{a+1}) \in D \). 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 \( (r, D) = (f_1, f_2, \ldots, f_k) \) the final sequence. From construction, the sequence \( j(r, D) \) is such that if \( (r, f_k) \in D \) then \( (r, f_{k+1}) \notin D \). Let \( \mathfrak{u}(r, D) \) be the minimal \( k \) such that \( (r, f_k) \in D \). If \( (r+1, f_{k+1}) \in D \)

then set \( \mathfrak{u}(r, D) = r + 1 \). We are now in a position to define the operation of \( \bar{\theta}_w \) on the diagram \( D \). To this end let us first assume that \( a > b \). This means that we have \( a > b \) elements in row \( r \) then in row \( r + 1 \). Hence \( q = \mathfrak{u}(r, D) + 1 \geq a + b - 1 \) for \( q \) the length of \( \mathfrak{u}(r, D) \). The equality holds if and only if \( \mathfrak{u}(r, D) = 1 \). In the case \( a > b \) the operator \( \bar{\theta}_w \) on the diagram \( D \) is defined by the map

\[ (B.4a) \quad \bar{\theta}_w D = \{ D_0, D_1, \ldots, D_{a+b-1} \} \]

where \( D_0 \) is identical to \( D \) except that we remove the element in position \( (r, w(r)) \) and for \( k = 1,2,\ldots,a+b-1 \) we successively set \( D_k \) to be identical to \( D_{a+b-1} \) except that the elements \( (r, f_k, D_{a+b-1}) \) is replaced by \( \mathfrak{u}(r, D_{a+b-1}) \). Now if \( a < b \) we have \( \mathfrak{u}(r, D) = a \leq a + b + 1 \) with equality iff \( \mathfrak{u}(r, D) = q + 1 \). So \( \mathfrak{u}(r, D) = r + 1 \geq b + a \). In this case the operator \( \bar{\theta}_w \) on the diagram \( D \) is defined by the map

\[ (B.4b) \quad \bar{\theta}_w D = \{ D_0, D_1, \ldots, D_{a+b-1} \} \]

where \( D_0 \) is identical to \( D \) except that we remove the element in position \( (r, w_r) \) and the element \( (r + 1, f_k, D_{a+b-1}) \) is replaced by \( (r, f_k, D_{a+b-1}) \). Now if \( a = b \) we successively set \( D_0 \) to be identical to \( D_{a+b-1} \) except that the element \( (r + 1, f_{k+1}, D_{a+b-1}) \) is replaced by \( (r, f_{k+1}, D_{a+b-1}) \). Finally if \( a = b \) then

\[ (B.4c) \quad \bar{\theta}_w D = \{ \} \]

With this definition of \( \bar{\theta}_w \) we have that

\[ (B.5) \quad \bar{\theta}_w D^D = \sum_{a, b} \theta_w D^D \]

with the positive sign in case \( (B.4a) \) and the negative sign in case \( (B.4b) \). For \( (B.4c) \) the result is zero.

We shall now show that

\[ (B.6) \quad \bar{\theta}_w \text{ maps } D(u) \text{ into } D(u) \]

Proof: The reader will notice that in \( D(u) \) the rectangle defined by the rows \( 1, 2, \ldots, r + 1 \) and the columns \( 1, 2, \ldots, w(r) - 1 \) is filled with elements and none of these elements can be B-move. Hence these elements are fixed in any diagram \( D \in D(u) \). The same applies to all elements in column \( w(r) \). If both are in the smallest rows and there are no elements in the rows strictly greater than \( r \). Now let \( D \) be a diagram in \( D(u) \) and assume that \( \bar{\theta}_w D = \{ D_0, D_1, \ldots, D_{a+b} \} \) is non-empty. The remark
above implies that the element in position \((r, w(r))\) does not affect the sequence of B-moves from \(D(w)\) to \(D\). Hence we can apply the same sequence of B-moves to \(D(w) - \{(r, w(r))\}\) and obtain \(D_b\). Moreover \(D(w) - \{(r, w(r))\}\) 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 \(D_b\) is obtainable from \(D(w)\) by a sequence of B-moves, that is \(D_b \notin \Omega(w)\). Now from the construction of \(\partial, D, D_b (b > 0)\) is obtained from \(D_{b-1}\) by exactly one B-move. Hence \(\partial, D \in \Omega(w)\).

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 \(v = (6, 3, 9, 5, 1, 2, 11, 8, 4, 7, 10)\). 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 elements in position \((r, w(r))\) is colored black.

![Diagram](image)

Now let \(D\) be the following diagram of \(\Omega(w)\):

![Diagram](image)

Here, \(a_0(D) = 7, a_{+1}(D) = 4\) and \(j(r, D) = (3, 5, 7, 8, 10)\). The reduced sequence \(f(r, D)\) is \((8, 10)\) and \(\text{up}(r, D) = 1\). Hence \(\partial, D \in \{\partial_0, D_1, D_2\}\) where

\[\sum_{\partial \in \partial_0(D)} \partial x^p = 0.\]

Proof: There are two classes of diagrams in \(\Omega_1(w)\). The first class contains the diagrams \(D\) for which \(a_0(D) = a_{+1}(D)\). In this case it is trivial that \(\partial x^p = 0\). The other class is formed by the diagrams \(D\) such that \(a_0(D) \neq a_{+1}(D)\) and \(\text{up}(r, D) > 1\). In this case we shall construct...
an involution, \( D = D' \), such that \( \partial^2 z^D + \partial z^{D'} = 0 \). Let \( f(r, D) = (f_1, f_2, \ldots, f_i) \), a = \( u(r, D) \) and \( b = a_{x+1}(D) \). We first define the involution for the case \( a > b \). Since \( u(r, D) > 1 \) we must have \( a = u(r, D) + 1 \geq a - b \). So let \( D' \) be identical to \( D \) except that the elements in positions \( r, f_{u(r, D)}(D), \ldots, (r, f_{u(r, D)+a-1}(D)) \) \( D \)-move to the positions \( r + 1, f_{u(r, D)}(D) + 1, \ldots, (r + 1, f_{u(r, D)+a-1}(D) + 1) \). It is clear that \( D' \in \Omega(v) \). But \( f(r, D') = f(r, D) \) and \( u(r, D') > u(r, D) > 1 \), hence \( D' \notin \Omega(v) \). Moreover we have \( a_{x+1}(D) = b \) and \( a_{x+1}(D) = a \), hence \( \partial^2 z^D + \partial z^{D'} = 0 \). 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

\[
\Theta_a = \partial \Theta_b = \sum_{D \in \Omega(x+1)} \partial z^D = \sum_{D \in \Omega(x-1)} \partial z^D + \sum_{D \in \Omega(x)} \partial z^D = \sum_{D \in \Omega(x)} \partial z^D + \partial z^{D'} = 0 \quad \text{ (B.2)}
\]

Kohnert’s construction

Let \( D \) be any diagram. Choose \( (i, j) \notin D \) such that \( (i, j') \notin D \) for all \( j' > j \). Let us suppose that there is a point \( (i', j) \notin D \) and \( i' < i \). Then let \( h < i \) be the largest integer such that \( (h, j) \notin D \) and let \( D_i \) denote the diagram obtained from \( D \) by replacing \( (i, j) \) by \( (h, j) \). We say that \( D_i \) 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

\[
\Theta_a = \sum_{D \in K(D(w))} z^D. \quad \text{(B.9)}
\]

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

\text{Appendix : Combinatorial construction}

on \( i - h \). If \( i = h = 1 \) then the K-move is simply one B-move. Now if \( i = h = 1 \), then we first perform the sequence of B-moves in row \( h = 1 \) 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 any row of \( D \) let \( B_i(D) \) denote the set of all diagrams (including \( D \)) obtained from \( D \) by any sequence of B-moves in the rows \( i, i+1 \) only. It is clear that if \( i \) 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 \( \Omega_i \) denote the set of all diagrams obtainable from \( B_i(D(w)) \) by any sequence of K-moves for which no elements cross the border between the rows \( i, i+1 \). A simple inductive algorithm may be used here to show that for any \( D \in \Omega_i \) we have \( \Omega_i(D) \subset \Omega_{i+1} \). Next let \( \Omega_k \) 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 \( k \)) that \( D \in \Omega_k \) we have \( \Omega_k(D) \subset \Omega_k \). But some of the cases are really hard to formalize! Now this completed would show that \( \Omega(w) = K(D(w)) \).
Orthogonality

Recall that

\[ P_\alpha = \mathbb{Z}[x_1, \ldots, x_n], \]
\[ A_\alpha = \mathbb{Z}[x_1, \ldots, x_n]^{\mathbb{A}} \]

where \( x_1, \ldots, x_n \) are independent indeterminates.

(5.1) \( P_\alpha \) is a free \( A_\alpha \)-module of rank \( n! \) with basis

\[ B_\alpha = \{ x^i : 0 \leq i - 1, 1 \leq i \leq n \}. \]

Proof by induction on \( n \). The result is trivially true when \( n = 1 \), so assume that \( n > 1 \) and that \( P_{\alpha-1} \) is a free \( A_{\alpha-1} \)-module with basis \( B_{\alpha-1} \). Since \( P_\alpha = P_{\alpha-1}[x_n] \), it follows that \( P_\alpha \) is a free \( A_{\alpha-1}[x_n] \)-module with basis \( B_{\alpha-1} \). Now

\[ A_{\alpha-1}[x_n] = A_{\alpha}[x_n], \]

because the identities

\[ e_\alpha(x_1, \ldots, x_n) = \sum_{i=0}^{n} (-1)^i e_i(x_1, \ldots, x_n) \]

show that \( A_{\alpha-1} \subset A_{\alpha}[x_n] \), and on the other hand it is clear that \( A_{\alpha} \subset A_{\alpha-1}[x_n] \). Hence \( P_\alpha \) is a free \( A_{\alpha}[x_n] \)-module with basis \( B_{\alpha-1} \).

To complete the proof it remains to show that \( A_{\alpha}[x_n] \) is a free \( A_{\alpha} \)-module with basis

\[ 1, x_1, \ldots, x_n^{\alpha-1} \]. Since \( \prod_{i=1}^{\alpha} (x_\alpha - x_i) = 0 \), we have

\[ x_\alpha^k = x_1 x_\alpha^{k-1} - \alpha x_\alpha^{k-2} + \ldots + (-1)^{\alpha-1} x_\alpha \]

from which it follows that the \( x_\alpha^i \) (\( 1 \leq i \leq n \)) generate \( A_{\alpha}[x_n] \) as a \( A_{\alpha} \)-module. On the other hand, if we have a relation of linear dependence

\[ \sum_{i=1}^{\alpha} \lambda_i x_{\alpha}^{i-1} = 0 \]
with coefficients \( f_i \in \Lambda_n \), then we have also
\[
\sum_{i=1}^{n} f_i x_i^{n-i} = 0
\]
for \( j = 1, 2, \ldots, n \), and since
\[
det(\varepsilon^T) = \prod_{i<j} (x_i - x_j) \neq 0,
\]
it follows that \( f_1 = \cdots = f_n = 0 \).

As before, let \( \delta = (n - 1, n - 2, \ldots, 1, 0) \). By reversing the order of \( x_1, \ldots, x_n \) in (5.1) it follows that
\[
(5.4) \quad \text{The monomials } x^\alpha, \alpha \subseteq \{1, \ldots, n\}, 0 \leq n - \delta \leq n \text{ for } 1 \leq \delta \leq n \text{ form a } \Lambda_n \text{-basis of } P_\alpha.
\]

We define a scalar product on \( P_\alpha \), with values in \( \Lambda_n \), by the rule
\[
(f, g) = \delta_{\alpha\beta}(fg) \quad (f, g \in P_\lambda)
\]
where \( w \) is the longest element of \( \Lambda_n \). Since \( \delta_{\alpha\beta} \) is \( \Lambda_n \)-linear, so is the scalar product.

(5.5) Let \( w \in \Sigma_n \) and \( f, g \in P_\alpha \). Then
(i) \(< \partial_1 f, g > = f, \partial_1 g > = f, \partial_{\alpha}(wfg) > = (wfg, (wfg))
\]
because \( \partial_1 f \) is symmetrical in \( x_1, x_{\alpha+1} \). The last expression is symmetrical in \( f \) and \( g \), hence
<br>
(ii) \(< \partial_1 f, g > = f, \partial_{\alpha}(wfg) > = (wfg, (wfg))
\]
and since \( \delta_{\alpha\beta} = -\delta_{\beta\alpha} \) this is equal to
\[
-\delta_{\alpha\beta}(\partial_1 f(x_1)) = \delta_{\alpha\beta}(f(x_1)) = -f, \partial_{\alpha} g > = 0
\]
(5.4) Let \( u, v \in \Sigma_n \) be such that \( (u) + (v) = (\varepsilon) \). Then
\[
< \Theta_n, \Theta_n > = \begin{cases} 1 & u = wv, \\ 0 & \text{otherwise.} \end{cases}
\]
Orthogonality

Proof: We have
\[
< \Theta_n, \Theta_n > = < \delta_{\alpha\beta}, \delta_{\alpha\beta} > = (\varepsilon, \varepsilon) = 0
\]
by (3.3). Also \( f(wu) - f(u) = f(v) \), hence
\[
\delta_{\alpha\beta}(wv) = \begin{cases} 1 & v = wu, \\ 0 & \text{otherwise.} \end{cases}
\]
It follows that
\[
< \Theta_n, \Theta_n > = \begin{cases} 0 & v \neq wu, \\ < e, e > = 1 & v = wu. \end{cases}
\]
(5.5) Let \( u, v \in \Sigma_n \). Then
\[
< \Theta_n, \Theta_n > = < \delta_{\alpha\beta}, \delta_{\alpha\beta} > = (\varepsilon, \varepsilon) = 0
\]
by (5.5) and (4.12). By (4.12) the scalar product is therefore zero unless \( (u) = (v) \), hence
\[
< \Theta_n, \Theta_n > = 0
\]
and it is equal to \( (v) < \Theta_n, \Theta_n > = (v) \). Now \( \alpha_{<v,w>} \) is a linear combination of monomials \( x^\alpha \) such that \( \alpha \subseteq \delta \) and \( |\alpha| = |(u) - (v)| \). Hence \( \Theta_n < \Theta_n > \) is a sum of monomials \( x^\alpha \) and
<br>
(5.6) The Schubert polynomials \( \Theta_n, w \in \Sigma_n \), form a \( \Lambda_n \)-basis of \( P_\alpha \).

Proof: Let \( u, v \in \Sigma_n \), and let
\[
\text{Let } u, v \in \Sigma_n \text{ and let }
\]
(1)
\[
\Theta_n = \sum_{\alpha} \delta_{\alpha\beta} x^\alpha.
\]
Notes on Schubert Polynomials

(2)

\((\psi|\phi)_{\mathcal{S}_n} = \sum_{x\in\mathcal{S}_n} \lambda_x \psi^2,\)

with coefficients \(a_{x \in \mathcal{S}_n} \lambda_x \). Let \(a_{x \in \mathcal{S}_n} = < \psi, \phi > .\) Then from (5.5) we have

\[ \sum_{x \in \mathcal{S}_n} a_{x \in \mathcal{S}_n} \lambda_x = \lambda_n, \]

or in matrix terms

\[ \mathbf{AC}\mathbf{B} = \mathbf{1} \]

where \( \mathbf{A} = (a_{x \in \mathcal{S}_n}), \mathbf{B} = (b_{x \in \mathcal{S}_n}) \) and \( \mathbf{C} = (c_{x \in \mathcal{S}_n}) \) are square matrices of size \( n \), with coefficients in \( \mathcal{S}_n \).

From (2) it follows that each of \( \mathbf{A}, \mathbf{B}, \mathbf{C} \) has determinant \( \pm 1 \), hence the equations (2) can be solved for \( \psi^2; \rho \subseteq \lambda \), as \( \lambda \)-linear combinations of the Schubert polynomials \( \mathcal{S}_n, \psi \in \mathcal{S}_n \). Since by (5.1') the \( \psi^2 \) from a \( \lambda \)-basis of \( \mathcal{S}_n \), we also do the \( \mathcal{S}_n \) in (5.7).

We have

\[ \langle f, g \rangle = \sum_{x \in \mathcal{S}_n} \langle \psi(x) \mathcal{S}_n \rangle \mathcal{S}_n \]

for all \( f, g \in \mathcal{S}_n \).

Proof: Let \( \langle f, g \rangle \) denote the right-hand side of (5.7). We claim first that

\[ \langle f, g \rangle \in \mathcal{S}_n. \]

(1)

For this it is enough to show that \( \partial_i \psi = 0 \) for \( 1 \leq i \leq n - 1 \). Let

\[ \mathcal{A}_n = \{ w \in \mathcal{S}_n : \ell(w) > \ell(w) \}, \]

then \( \mathcal{A}_n \) is the disjoint union of \( \mathcal{A} \) and \( \mathcal{A}_n \), and \( \mathcal{A} \mathcal{A}_n = \mathcal{A}_n \). Hence

\[ \phi(f, g) = \sum_{w \in \mathcal{S}_n} \langle \psi(w), \partial_i \psi \rangle \mathcal{S}_n \]

for all \( \phi, \psi \in \mathcal{S}_n \) we have

\[ \partial_i \psi \cdot \partial_i \psi = \partial_i \psi \cdot \partial_i \psi = 0, \]

it follows that \( \partial_i \phi(f, g) \equiv 0 \) for all \( i \) as required.

Next, since each operator \( \partial_i \psi \) is \( \lambda_n \)-linear, it follows that \( \phi(f, g) \) is \( \lambda_n \)-linear in each argument.

By (5.6) it is therefore enough to verify (5.7) when \( f = u \psi \), and \( g = \psi \mathcal{S}_n, \psi \in \mathcal{S}_n \). We have then

\[ \phi(u \psi, \psi) = \sum_{w \in \mathcal{S}_n} \langle u, \psi \rangle \mathcal{S}_n \psi - u \psi \mathcal{S}_n \psi = \psi \mathcal{S}_n \psi, \]

which by (4.2) is equal to

\[ \sum_{w \in \mathcal{S}_n} \langle u, \psi \rangle \mathcal{S}_n \psi \]

summed over \( u \in \mathcal{S}_n \) such that

\[ \ell(u) = \ell(u) - \ell(u^{-1}) = \ell(u) - \ell(u), \]

and

\[ \ell(u) = \ell(u) - \ell(u^{-1}) = \ell(u) - \ell(u). \]

Hence the polynomial \( (2) \) is (i) symmetric in \( x_1, \ldots, x_n \) (by (1) above), (ii) independent of \( x_n \), (iii) homogeneous of degree \( \ell(u) - \ell(u) \). Hence it vanishes unless \( \ell(u) = \ell(u) \) and \( u = u^{-1} = u \), in which case it is equal to \( \psi(u) = \psi(u) \). Hence

\[ \phi(u \psi, \psi) = \psi \mathcal{S}_n \psi = < u \psi \mathcal{S}_n \psi > \]

by (5.5). This completes the proof of (5.7).

Now let \( x = (x_1, \ldots, x_n) \) and \( y = (y_1, \ldots, y_n) \) be two sequences of independent variables, and let

\[ \Delta = \Delta(x, y) = \prod_{i \leq \infty} (x_i - y_i) \]

(the "semiresidual"). We have

\[ \Delta \psi(x, y) = \sum_{i \leq \infty} \phi(x_{i+1} - x_i) \]

is non-zero if and only if \( \psi(i) \neq j \) whenever \( i \leq j \leq n \), that is to say if and only if \( x \neq y \), and

\[ \Delta(x, y) = \prod_{x_i - x_i} (x_i - x_i) = \prod_{x_i - x_i} (y_i - y_i) \]

is zero if and only if \( x = y \).

The polynomial \( \Delta(x, y) \) is a linear combination of the monomials \( x^\alpha \), \( \alpha \in \mathcal{S}_n \), with coefficients in \( \Delta(y_1, \ldots, y_n) = \mathcal{S}_n(y) \), hence by (4.11) can be written uniquely in the form

\[ \Delta(x, y) = \sum_{\alpha \in \mathcal{S}_n} a_{\alpha} \mathcal{S}_n(\alpha) \]
with \( T_{u}(y) \in P_{u}(y) \). By (5.5) we have

\[
T_{u}(y) = \sum_{v \in C_{u}} \Theta_{u}(v) \Theta_{v^{\ast}}(-y)
\]

where the suffix \( z \) means that the scalar product is taken in the \( z \) variables. Hence

\[
T_{u}(y) = \Theta_{u}(y)/\Theta_{v^{\ast}}(-y)
\]

by (2.10), where \( z \in S_{u} \) acts by permuting the \( z \).

Now the expression (1) must be independent of \( z_{1}, \ldots, z_{n} \). Hence we may set \( z_{i} = y_{i} \), \( 1 \leq i \leq n \). But then (3.9) shows that the only non-zero term in the sum (1) is that corresponding to \( u = u_{y} \), and we obtain

\[
T_{u}(y) = \Theta_{v^{\ast}}(-y).
\]

Hence we have proved

(5.10) ("Cauchy formula")

\[
\Delta(x, y) = \sum_{v \in C_{u}} \Theta_{u}(x) \Theta_{v^{\ast}}(-y)
\]

Remark. Let \( u = r + s \) where \( r, \ s \geq 1 \), and regard \( S_{r} \times S_{s} \) as a subgroup of \( S_{u} \), with \( S_{r} \) permuting \( 1, \ldots, r \) and \( S_{s} \) permuting \( r + 1, \ldots, r + s \). Let \( w_{r_{1}}^{s_{1}}, w_{r_{2}}^{s_{2}} \) be the longest elements of \( S_{r}, S_{s} \) respectively, and let \( w = w_{r_{1}}^{s_{1}} w_{r_{2}}^{s_{2}} \). If \( w \in S_{u} \), we have \( \Theta_{u} \Theta_{w} = \Theta_{w} \Theta_{u} \) if \( \ell(w) = \ell(u) - \ell(w) \), that is to say if \( w \) is Grassmannian (with its only descent at \( r \)), and \( \Theta_{u} \Theta_{w} = 0 \) otherwise. Hence by applying \( \Theta_{u} \) to the \( x \)-variables in (5.10) we obtain

\[
\Delta_{r}(x, y) = \sum_{v \in C_{u}} \Theta_{u}(x) \Theta_{v^{\ast}}(-y)
\]

where \( G_{r} \subset S_{r} \) is the set of Grassmannian permutations \( v \) with descent at \( r \) (i.e. \( s(i) < s(i+1) \) if \( i \neq r \)). On the other hand, it is easily verified that

\[
\Delta_{u}(x, y) = \prod_{i=1}^{r} (x_{i} - y_{i})
\]

and that \( y' = v w_{y} \) is the permutation

\[
(v(r+1), v(r+2), v(r+3), \ldots, v(r+1))
\]

hence is also Grassmannian, with descent at \( s \).
Notes on Schubert Polynomials

Let \((e_x)_{x \in X}\) be the basis dual to \((x^*_a)_{a \in A}\). If

\[
\begin{align*}
\Theta_0 &= \sum a_x x^*_a, \\
\Theta^* &= \sum b_x e_x,
\end{align*}
\]

then by taking scalar products we have

\[
\sum_n a_n b_n = \langle \Theta_0, \Theta^* \rangle,
\]

and therefore also

\[
\sum_n a_n b_n = \langle \Theta_0, \Theta^* \rangle,
\]

so that

\[
\sum_{x \in X} \Theta_0(x) \Theta^*(y) = \sum_{x \in X} \sum_{a \in A} a_x b_x x^*_a y^*_a = \sum_{x \in X} x^*_x y^*_x.
\]

From (5.13) it follows that \(y_x\) is the coefficient of \(x^*_x\) in \(\prod_{x \in X} (x_i - y_i)\), and hence we find

\[
(5.14) \quad e_x = (-1)^{|x|} \prod_{x \in X} e_x(x_{x_1}, \ldots, x_n)
\]

where \(n = |x|\).

Let

\[
C(x, y) = \langle \Theta_0, \Theta^*(y) \Theta^*(x) \Theta_0 \rangle = \prod_{x \in X} (y_i - x_i).
\]

If \(f(x) \in H_a\) (4.11), let \(f(y)\) denote the polynomial in \(y_1, \ldots, y_n\) obtained by replacing each \(x_i\) by \(y_i\). Then we have

\[
(5.15) \quad < f(x), C(x, y) >_{x, y} = f(y),
\]

where as before the suffixes \(x, y\) mean 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

\[
C(x, y) = \sum_{x \in X} \langle \Theta_0, \Theta^*(y) \Theta^*(x) \Theta_0 \rangle.
\]

Hence by (5.5)

\[
< C(x, y), \Theta^*(x) \Theta_0 >_{x, y} = \langle \Theta_0, \Theta^*(x)^* y^*_x \rangle = \Theta^*(x)^* y^*_x.
\]

Hence (5.15) is true for all Schubert polynomials \(\Theta_0, \Theta^*_0 \in S_n\). Since the scalar product is \(\Theta_0, \Theta^*_0\)-linear it follows from (5.6) that (5.15) is true for all \(f \in H_a\).

Let \(\Theta_0\) be the homomorphism that replaces each \(y_i\) by \(x_i\). Then (5.15) can be restated in the form

\[
(5.15') \quad \Theta^*_0 < f(x), C(x, y) >_{x, y} = f(y)
\]

for all \(f \in H_a\).

Now let \(z = [z_1, \ldots, z_n]\) be a third set of variables and consider

\[
(1) \quad < C(x, y), \Theta^*_0 \Theta_0 >_{x, y} = \Theta^*_0 \Theta_0
\]

for \(u, v \in S_n\), where \(\Theta_0\) and \(\Theta^*_0\) act on the \(x\) variables. By (5.3) this is equal to

\[
(2) \quad c(v) < C(x, z), \Theta^*_0 \Theta_0 >_{x, z} = \Theta^*_0 \Theta_0
\]

and by (5.15') we have

\[
(3) \quad < C(x, z), \Theta^*_0 >_{x, z} = \Theta^*_0
\]

\[
(4) \quad < C(z, y), \Theta^*_0 >_{z, y} = \Theta^*_0
\]

Since \(\Theta^*_0\) and \(\Theta_0\) commute, it follows from (1) (4) that

\[
\Theta^*_0 \Theta_0 C(x, y) = \Theta^*_0 \Theta_0 C(x, y)
\]

\[< C(x, z), \Theta^*_0 \Theta_0 >_{x, z} = \Theta^*_0 \Theta_0
\]

Hence we have

(5.16)

\[
< C(x, z), \Theta^*_0 >_{x, z} = \Theta^*_0 \Theta_0
\]

for all \(u, v \in S_n\), where \(\Delta = \Delta(x, y)\) and \(\theta = \Theta_0\).

Let \(E_x\) denote the algebra of operators \(\phi\) of the form

\[
\phi = \sum_{w \in X} \Theta_0 w
\]

with coefficients \(\Theta_0 \in Q(a_1, \ldots, a_n)\). For such a \(\phi\) we have

\[
(5.17) \quad \phi = \langle \Theta_0, \Theta^*_0 \rangle \theta (\phi w^{-1} \Theta_0)
\]

for all \(w \in S_n\), where \(\phi\) and \(w^{-1} \Theta_0\) act on the \(x\) variables in \(\Delta\).
Notes on Schubert Polynomials

For \( \theta(w^{-1} \omega \Delta) = \sum \omega \phi(w^{-1} \omega \Delta) \), and by (3.8) \( \theta(w^{-1} \omega \Delta) = \Delta(w^{-1} \omega \Delta) \), it is zero if \( w \neq \omega \), and in equal to \( \omega \omega \Delta \) if \( w = \omega \).

Let \( u \in \mathfrak{S}_n \), and let \( (a_1, \ldots, a_k) \) be a reduced word for \( u \), so that \( \Delta_u = \Delta_{a_1} \cdots \Delta_{a_k} \). Since \( \Delta_u = (x_n - x_{n-1})^{-1}(1 - x_n) \) for each \( s \), it follows that we may write

\[
\Delta_u = \omega(\omega \Delta) \sum \omega \phi(w^{-1} \omega \Delta) \]

(5.10)

where \( \varsigma \subseteq u \) means that \( \varsigma \) is of the form \( a_1, \ldots, a_k \), where \( (a_1, \ldots, a_k) \) is a subword of \( (a_1, \ldots, a_k) \).

The coefficients \( a_{\omega \Delta} \) are polynomials, for it follows from (5.18) and (5.17) that

\[
a_{\omega \Delta} = \theta(\omega^{-1} \omega \Delta) \]

(5.19)

For all \( f \in \mathfrak{S}_n \), we have

\[
\theta(\omega \Delta(f)) = \begin{cases} \omega f & \text{if } u = \omega, \\ 0 & \text{otherwise,} \end{cases}
\]

(5.20)

**Proof.** From (5.18) we have

\[
\theta(\omega \Delta(f)) = \omega^{-1} \sum \omega \phi(w^{-1} \omega \Delta) \omega(f) \Delta \omega \]

By (5.9) this is zero if \( u \neq \omega \), and if \( u = \omega \) then by (5.10)

\[
\theta(\omega \Delta(f)) = \omega^{-1} \sum \phi(w^{-1} \omega \Delta) \omega(f) \Delta \omega
\]

(5.21)

and thus we can express any \( \phi \in \mathfrak{S}_n \) as a linear combination of the operators \( \Delta_u \). Explicitly, we have

\[
\phi = \sum \omega \phi(w^{-1} \omega \Delta) \Delta_u.
\]

(5.22)

Orthogonality

**Proof.** By linearity we may assume that \( \phi = \Delta \), with \( f \in \mathfrak{S}_n \). Then

\[
\theta(\delta_0 \omega \Delta(f)) = \theta(\delta_0 \omega \Delta(\omega \Delta(f)))
\]

Now by (4.2) \( \delta_0 \omega \Delta(f) \) is either zero or equal to \( \omega \omega \Delta(\omega \Delta(f)) \). By (5.20) \( \theta(\delta_0 \omega \Delta(\omega \Delta(f))) \) is zero if \( \omega \neq \omega \), and in equal to 1 if \( \omega = \omega \). Hence the right-hand side of (5.22) is equal to \( \theta \), as required.

In particular, it follows from (5.21) that

\[
\beta_u = \theta(\omega \Delta(\omega \Delta(f))
\]

(5.23)

Hence is a polynomial.

The coefficients \( a_{\omega \Delta} \) in (5.18) and (5.23) satisfy the following relations:

(5.24) (i) \( a_{\omega \Delta} = \phi(\omega \Delta(\omega \Delta(f))) \)

(ii) \( a_{\omega \Delta} = \phi(\omega \Delta(\omega \Delta(f))) \)

(iii) \( a_{\omega \Delta} = \phi(\omega \Delta(\omega \Delta(f))) \)

for all \( u, v \in \mathfrak{S}_n \), where \( \omega = \omega \omega \Delta \).

**Proof.** (i) By (5.23) and (2.12) we have

\[
\beta_{\omega \Delta} = \phi(\omega \Delta(\omega \Delta(f))) \]

(5.25)

By (5.10) again.

The matrix of coefficients \( (a_{\omega \Delta}) \) in (5.18) is triangular with respect to the ordering \( \preceq \), and one sees easily that the diagonal entries \( a_{\omega \Delta} \) 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

\[
\phi = \sum \omega \phi(w^{-1} \omega \Delta) \Delta_u
\]

(5.26)

and likewise

\[
\theta(\delta_0 \omega \Delta(\omega \Delta(f))) = \phi(\omega \Delta(\omega \Delta(f))) \Delta_u
\]

(5.27)

again by (5.9). Hence (i) follows from (5.16).

(ii) Since \( \delta_0 = \phi(\omega \Delta(\omega \Delta(f))) \) we have

\[
\sum \omega \phi(w^{-1} \omega \Delta) \omega(\omega \Delta(\omega \Delta(f))) = \phi(\omega \Delta(\omega \Delta(f))) \]

(5.28)
Let $E$ be the subalgebra of operators $\phi \in E_1$ such that $\phi(P) \subseteq P$. Then $E_{\omega}$ is a free $P_\omega$-module with basis $(\partial_\omega)_{\omega \in \Sigma}$.

Proof: If $\phi = \sum_{\omega \in \Sigma} \omega \partial_\omega \in E_{\omega}$, then by (5.22)
\[
\phi = \partial_\omega \phi(\omega, \Delta) \in P.
\]

On the other hand, the $\partial_\omega$ are a $Q_\omega$-basis of $E_\omega$, and hence are linearly independent over $P_\omega$.

Let $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$ be two sequences of independent indeterminates, and recall (5.8) that
\[
\Delta(x, y) = \prod_{i \neq j} (x_i - y_j).
\]

For each $w \in S_n$, we define the double Schubert polynomial $\Theta_w(x, y)$ to be
\[
\Theta_w(x, y) = \partial_{\omega} \Delta(x, y)
\]

where $\partial_{\omega} \Delta$ acts on the $x$ variables.

Since $\Delta(x, 0) = x^n$ we have
\[
\Theta_w(x, 0) = \Theta_w(x).
\]

The (single) Schubert polynomial indexed by $w$.

From the Cauchy formula (5.10) we have
\[
\Theta_w(x, y) = \sum_{\omega \in \Sigma} \partial_{\omega} \Theta_{\omega}(x) \Theta_{\omega}(-y)
\]

and by (4.2)
\[
\partial_{\omega} \Theta_{\omega}(x) = \Theta_{\omega}(x)
\]

if $\ell(\omega w) = \ell(\omega w) - \ell(\omega^{-1} w)$, i.e. if $\ell(\omega w) = \ell(w) - \ell(\omega)$, and
\[
\partial_{\omega} \Theta_{\omega}(x) = 0
\]

otherwise. Hence
\[
\Theta_w(x, y) = \sum_{\omega \in \Sigma} \Theta_{\omega}(x) \Theta_{\omega}(-y)
\]
A summed over all \( u, v \in \mathbb{S}_n \) such that \( w = v^{-1} u \) and \( f(u) = f(v) + f(w) \).

From (6.3) it follows that \( \Theta_w(x, y) \) is a homogeneous polynomial of degree \( f(u) \) in \( x_1, \ldots, x_n \).

We have

\[
\Theta_w(x, y) = \delta(x, y),
\]

\[
\Theta_0(x, y) = 1,
\]

\[
\Theta_{w MN}(x, y) = \Theta_{w MN}(y, x) = (w x, y) \Theta_{w MN}(y, x) \text{ for all } w \in \mathbb{S}_n.
\]

\[
\Theta_w(x, y) = 0 \text{ for all } w \in \mathbb{S}_n \text{ except } w = 1.
\]

Proof: (i) is immediate from the definition (6.1).

(ii) and (iii) follow from (6.3).

(iv) follows from (5.20), since \( \Theta_w(x, z) = \Theta_{w MN}(z, w) \Theta_w(x, y) \) for all \( w \in \mathbb{S}_n \).

(6.5) (Stability) If \( m > n \) and \( i \) is the embedding of \( \mathbb{S}_n \) in \( \mathbb{S}_m \), then

\[
\Theta_{w MN}(x, y) = \Theta_i(x, y)
\]

for all \( w \in \mathbb{S}_n \).

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 \( \Theta_w(x, y) \) are well defined for all permutations \( w \in \mathbb{S}_n \).

For any commutative ring \( K \), let \( K[\mathbb{S}_n] \) denote the \( K \)-module of all functions on \( \mathbb{S}_n \) with values in \( K \). We define a multiplication in \( K[\mathbb{S}_n] \) as follows: for \( f, g \in K[\mathbb{S}_n] \),

\[
(fg)(w) = \sum_{v \in \mathbb{S}_n} f(v)g(v)
\]

summed over all \( u, v \in \mathbb{S}_n \) such that \( w = v^{-1} u \) and \( f(u) = f(v) + f(w) \). For this multiplication, \( K[\mathbb{S}_n] \) is an associative (but not commutative) ring, with identity element \( 1 \), 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 \( f, g \in K[\mathbb{S}_n] \).

(6.6) Let \( f, g \in K[\mathbb{S}_n] \).

\[
(fg) \text{ is a unit in } K[\mathbb{S}_n]
\]

Double Schubert polynomials

(i) If \( fg = f \) and \( f(1) \) is not a zero divisor in \( K \), then \( g = 0 \).

(ii) If \( fg = 1 \) then \( f = g \).

(iii) \( f \) is a unit (i.e., invertible) in \( K[\mathbb{S}_n] \) 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 \( f(u) \) that \( g(u) = 0 \) for all \( u \neq 1 \). So let \( r > 0 \) and assume that \( g(u) = 0 \) for all \( u \in \mathbb{S}_n \) such that \( 1 \leq f(u) \leq r - 1 \). Let \( w \) be a permutation of length \( r \). We have

\[
f(u) = (f g)(w) = f(u)g(1) + f(1)g(w) + \sum_{u \neq 1} f(u)g(v)
\]

where the sum on the right is over \( u, v \in \mathbb{S}_n \) such that \( u \neq 1, v \neq 1, u = w \) and \( f(u) \neq f(v) \neq f(w) \), so that \( 1 \leq f(u) \leq r - 1 \) and therefore \( g(v) = 0 \). Hence (i) reduces to \( f(1)g(1) = 0 \) and therefore \( g(1) = 0 \) as required.

(ii) We have \( f(1)g(1) = 0 \) so that \( f(1) \) is a unit in \( K \). Also \( f(1)g(1) = 1 \) if \( f(1)g(1) = 1 \), whereas \( g(1) = \frac{1}{f(1)} \) by (i) above.

(iii) Suppose \( f \) is a unit in \( K[\mathbb{S}_n] \), with inverse \( g \). Since \( fg = 1 \) we have \( f(1)g(1) = 1 \), whereas \( f(1) \) is an unit in \( K \).

Conversely, if \( f(1) \) is an unit in \( K \) we construct an inverse \( g \) of \( f(1) \) as follows. We define \( g(1) = f(1)^{-1} \) and proceed to define \( g(u) \) by induction on \( f(u) \). Assume that \( g(u) \) has been defined for all \( u \) such that \( f(u) < f(w) \) and set

\[
g(u) = -f(1)^{-1} \sum_{w \neq 1} f(w)g(w)
\]

summed over \( u, v \) such that \( uv = w, v \neq 1 \) and \( f(u) + f(v) = f(w) \). This definition gives \( (fg)(w) - 6 \) as required.

Now let \( \mathcal{O}(x) \) (resp. \( \mathcal{O}(x, y) \)) be the function on \( \mathbb{S}_n \), whose value at a permutation \( w \) is \( \Theta_w(x, y) \) (resp. \( \Theta_w(x, y) \)). (The coefficient ring \( K \) is now the ring \( \mathcal{Z}[x, y] \) of polynomials in the \( x_i \) and \( y_i \).) Since \( \Theta_v(x) = \Theta_v(y, z) = 1 \), it follows from (6.6)(iii) that \( \Theta_v(x) \) and \( \Theta_v(x, y) \) are units in \( K[\mathbb{S}_n] \).

(6.7) (i) \( \mathcal{O}(x, y) = \Theta(x, y) \).

(ii) \( \mathcal{O}(x, z) = 1 \).

(iii) \( \mathcal{O}(x, y) = \Theta(-y, -z) \).

(iv) \( \mathcal{O}(x)^{-1} = \Theta(0, x) \).

(v) \( \mathcal{O}(x)^{-1} = \Theta(-x, -1) \).

(vi) \( \mathcal{O}(y, z) = \Theta(y, z)^{-1} \).

(\( \mathcal{O}(x) = \Theta(x)^{-1} \)).
Proof: (i)-(iii) follow directly from (6.2) and (6.4).

From (6.3) and (6.4) we have

\[ \Theta_n(x,y) = \sum_{u,v} \Theta_n(u,v) p_{u,v}(x,y) \]

summed over \( u,v \in \mathbb{N} \) such that \( u = v \) and \( \ell(u) + \ell(v) = \ell(w) \). In other words,

(1)

\[ \Theta(w,z) = \Theta(0,0)p_{0,0}(x,y). \]

In particular, when \( y = x \) we obtain \( \Theta(0,0) = \Theta(x,x) = \frac{1}{0!} \) by (ii) above, and hence \( \Theta(0,0) = \Theta(x)^{+1} \). This establishes (iv); part (v) now follows from (iv) and (iii), and (vi) from (iv) and (1) above. \( \square \)

From (6.7) (vi) we have

\[ \Theta(x) = \Theta(p) \Theta(n, x) \]

or explicitly

\[ \Theta_n(x) = \sum_{u,v} \Theta_n(u,v) p_{u,v}(x,y) \]

summed over \( u,v \) such that \( u = v \) and \( \ell(u) + \ell(v) = \ell(w) \), so that \( u = uv^{-1} \) and \( \Theta_n = \partial_u \Theta_n \) by (4.2). Hence

\[ \Theta_n(x) = \sum_x \Theta_n(x,y) \partial_u \Theta_n(y) \]

(whence the operators \( \partial_u \) act on the \( y \) variables). The sum here may be taken over all permutations \( v \), since \( \partial_u \Theta_n = 0 \) unless \( \ell(uw^{-1}) = \ell(v) - \ell(w) \). By linearity and (1.13) it follows that

(6.8) (Interpolation Formula) For all \( f \in P_n \), \( \mathbb{Z}[x_1, \ldots, x_n] \) we have

\[ f(x) = \sum_{u,v} \Theta_n(x,y) \partial_u f(y) \]

summed over permutations \( w \in \mathbb{S}^{\langle n \rangle} \).

(The reason for the restriction to \( \mathbb{S}^{\langle n \rangle} \) in the summation is that if \( w \notin \mathbb{S}^{\langle n \rangle} \) we shall have \( u(m) > uw(m + 1) \) for some \( m > n \), and hence \( \partial_u = \partial_u \Theta_n \) where \( v = uw(m) \), but \( \partial_u f = 0 \) for all \( f \in P_n \), since \( m > n \), and therefore \( \partial_u f = 0 \).

Remarks. 1. By setting \( y_1 = 0 \) in (6.8) we regain (1.14).

2. When \( n = 1 \), the sum is over \( n! \), which consists of the permutations \( w_p = \delta_{w_{p-1}} \cdots \delta_{w_0} \) with \( p \geq 0 \); \( w_p \) is dominant, of shape \( (p) \), so that (see (6.15) below) \( \Theta_1(x,y) = (x-y_1) \cdots (x-y_p) \). Hence the case \( n = 1 \) of (6.8) is Newton's interpolation formula

\[ f(x) = \sum_{j=0}^{\ell(w)} (x - y_1) \cdots (x - y_j) \partial_{y_j} f(y_1, \ldots, y_{\ell(w)}) \]

where \( f = \partial_u \Theta_{n-1} \cdots \partial_u f, \) or explicitly

\[ f(y_0, \ldots, y_{\ell(w)}) = \sum_{j=0}^{\ell(w)} f(y_1, \ldots, y_{\ell(w)}) \Theta_{n-1}^{(j)}(y_0, \ldots, y_{\ell(w)}). \]

For any integer \( r \), let \( \Theta_n(x,r) \) denote the polynomial obtained from \( \Theta_n(x,y) \) by setting \( y = y_1 = \cdots = y_r \). Since

\[ \Theta_n(x,r) = \delta(x,r) = \prod_{i=1}^{r} (x_i - r)^{+1} \]

where \( x - r \) means \( (x_1 - r, x_2 - r, \ldots) \), it follows from the definitions (6.1) and (4.1) that

\[ \Theta_n(x,r) = \Theta_n(x - r) \]

for all permutations \( \sigma \). Hence, by (6.7)(vi),

\[ \Theta(x - r) = \Theta(x)^{-(r)} \Theta(r) \]

and in particular, for all integers \( r \),

\[ \Theta(x - r) = \Theta(x)^{-(r)} \Theta(r) \]

from which it follows that

(6.9)

\[ \Theta(r) = \Theta(1)^r \]

for all \( r \in \mathbb{Z} \).

Since \( \Theta_n(x) \) is a sum of monomials with positive integral coefficients (4.17), \( \Theta_n(1) \) is the number of monomials in \( \Theta_n(x) \) (each monomial counted the number of times it occurs). By homogeneity, we have

(6.10)

\[ \Theta_n(r) = r^{\ell(w)} \Theta_n(1). \]

From (6.7)(v) and (6.9) we obtain

\[ \Theta(1) = \Theta(1)^{-1} = \Theta(1) \]

so that we have another proof of the fact (4.20) that \( \Theta_n(1) = \Theta_{n-1}(1) \).

Now consider the function \( F \in \Theta(1) \) whose value at \( w \in \mathbb{S}^{\langle n \rangle} \) is

\[ F(w) = \begin{cases} \Theta_n(w), & \text{if } w \neq 1, \\ 0, & \text{if } w = 1. \end{cases} \]

Notes on Schubert Polynomials
For each positive integer $p$ we have

$$F^\prime = \left( \frac{\Theta(1) - 1}{p} \right)^p$$

$$= \sum_{r=0}^{p-1} \left( \frac{(-1)^r}{p} \right) \Theta(1)^r$$

(1)

by (6.9). The value of (1) at a permutation $w$ of length $p$ is by (6.10) equal to

$$\sum_{r=0}^{p-1} \left( \frac{(-1)^r}{p} \right) \Theta(w)^r$$

which is equal to $p\Theta(w)$ (consider the coefficient of $\phi^r$ in $\left( \frac{1}{\phi} \right)^p$). On the other hand, $F^\prime(w)$ is by definition equal to

$$\sum_{w_1, \ldots, w_p} F(w_1) \ldots F(w_p)$$

summed over all sequences $(w_1, \ldots, w_p)$ of permutations such that $w_1 \ldots w_p = w; f(w_1) + \cdots + f(w_p) = f(w) = p$, and $w_i \neq 1$ for $1 \leq i \leq p$. It follows that each $w_i$ has length 1, hence $w_i = a_i$, say, and that $(a_1, \ldots, a_p)$ is a reduced word for $w$. Since

$$\Theta_a = a_1 + \cdots + a_p$$

by (4.4), we have $F(w) = \Theta_a$, and hence the sum (2) is equal to $\sum_{a_1, \ldots, a_p} \Theta_a$, summed over all $(a_1, \ldots, a_p) \in \mathcal{R}(w)$.

We have therefore proved that

(6.11) The number of monomials in $\Theta_a$ is

$$\Theta_a = \frac{1}{p!} \sum_{a_1, \ldots, a_p} \Theta_a$$

summed over all $(a_1, \ldots, a_p) \in \mathcal{R}(w)$, where $p = f(w)$.

Remarks. 1. The reduced words for $I_m \times w(m \geq 1)$ are $(m + a_1, \ldots, m + a_p)$ where $(a_1, \ldots, a_p) \in \mathcal{R}(w)$. Hence from (6.11) and homogeneity we have

$$\Theta_{a_1+\cdots+a_p} = \frac{1}{m} \sum_{a_1+\cdots+a_p} \Theta_{a_1+\cdots+a_p}$$

summed over $\mathcal{R}(w)$ as before. Letting $m \to \infty$, we deduce that

$$\text{Card } \mathcal{R}(w) = \lim_{m \to \infty} \Theta_{a_1+\cdots+a_p} \left( \frac{1}{m} \right).$$

(6.12)

2. If $w$ is dominant of length $p$, then $\Theta_a$ is a monomial by (4.7), and hence in this case

$$\sum_{a_1+\cdots+a_p} = \frac{1}{p!} \left( \frac{1}{p} \right)^p$$

2. Suppose that $w$ is vexillary of length $p$. Then by (6.9) we have

$$\Theta_a = x_1(X_1, \ldots, X_n)$$

where $\lambda$ is the shape of $w$ and $\phi = (\phi_1, \ldots, \phi_n)$ the flag of $w$. Hence

$$\Theta_{a_1+\cdots+a_p} = x_1(X_1, \ldots, X_n)$$

for all $m \geq 1$. If we now set each $x_i = \frac{1}{\lambda_i}$ and then let $m \to \infty$, we shall obtain in the limit the Schur function $s_\lambda$ for the series $e' (M)$, Chap. 1, §2, Ex. 6), which is equal to $\lambda^n$, where $H(\lambda)$ is the product of the hook-lengths of $\lambda$. Hence it follows from (6.12) that if $w$ is vexillary of length $p$, then

$$\text{Card } \mathcal{R}(w) = \frac{1}{p!} \left( \frac{1}{p} \right)^p$$

where $\lambda$ 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 $S_\lambda$ indexed by $\lambda$.

4. It seems likely that there is a $q$-analogue of (6.11). Some experimental evidence suggests the following conjecture.

$$(6.11') \quad \Theta_a(1, q, q^2, \ldots) = \sum q^{a_1} (1 - q^{a_2}) \ldots (1 - q^{a_p})$$

summed over all reduced words for $w$, where

$$\phi(w) = \sum (a_i - a_{i+1})$$

When $w$ is vexillary the double Schubert polynomial $\Theta_a(e, x)$ 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 $\lambda$, then

$$\Theta_{a_1+\cdots+a_p}(x, y) = \prod_{i+j \in \lambda} (x_i - y_j)$$

$$= x_1(X_1 - Y_1, \ldots, X_n - Y_n)$$

Notes on Schubert Polynomials

Double Schubert polynomials

2. If $w$ is dominant of length $p$, then $\Theta_a$ is a monomial by (4.7), and hence in this case

$$\sum_{a_1+\cdots+a_p} = \frac{1}{p!} \left( \frac{1}{p} \right)^p$$

2. Suppose that $w$ is vexillary of length $p$. Then by (6.9) we have

$$\Theta_a = x_1(X_1, \ldots, X_n)$$

where $\lambda$ is the shape of $w$ and $\phi = (\phi_1, \ldots, \phi_n)$ the flag of $w$. Hence

$$\Theta_{a_1+\cdots+a_p} = x_1(X_1, \ldots, X_n)$$

for all $m \geq 1$. If we now set each $x_i = \frac{1}{\lambda_i}$ and then let $m \to \infty$, we shall obtain in the limit the Schur function $s_\lambda$ for the series $e' (M)$, Chap. 1, §2, Ex. 6), which is equal to $\lambda^n$, where $H(\lambda)$ is the product of the hook-lengths of $\lambda$. Hence it follows from (6.12) that if $w$ is vexillary of length $p$, then

$$\text{Card } \mathcal{R}(w) = \frac{1}{p!} \left( \frac{1}{p} \right)^p$$

where $\lambda$ 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 $S_\lambda$ indexed by $\lambda$.

4. It seems likely that there is a $q$-analogue of (6.11). Some experimental evidence suggests the following conjecture.

$$(6.11') \quad \Theta_a(1, q, q^2, \ldots) = \sum q^{a_1} (1 - q^{a_2}) \ldots (1 - q^{a_p})$$

summed over all reduced words $a = (a_1, \ldots, a_p)$ for $w$, where

$$\phi(w) = \sum (a_i - a_{i+1})$$

When $w$ is vexillary the double Schubert polynomial $\Theta_a(e, x)$ 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 $\lambda$, then

$$\Theta_{a_1+\cdots+a_p}(x, y) = \prod_{i+j \in \lambda} (x_i - y_j)$$

$$= x_1(X_1 - Y_1, \ldots, X_n - Y_n)$$
where \( m = \ell(A) \) and \( X_t = x_1 + \cdots + x_t, Y_t = y_1 + \cdots + y_t \) for all \( t \geq 1 \).

**Proof:** As in (4.6) we proceed by descending induction on \( \ell(u) \), \( u \in \mathcal{S}_e \). The result is true for \( w = w_0 \), since \( w_0 \) is dominant of shape \( \delta \) and

\[
\Theta_w(x,y) = \Delta(x,y) = \prod_{i,j \in \delta} (x_i - y_j).
\]

Suppose \( w \neq w_0 \) is dominant of shape \( \lambda \). Then \( \lambda \in \delta \) (and \( \lambda \neq \delta \)) let \( r \geq 0 \) be the largest integer such that \( \lambda' = \lambda - r \cdot \varepsilon_{\delta} \) for \( 1 \leq i \leq r \) and let \( \lambda = \lambda_{r+1} + 1 \leq n - r - 1 \). Then \( w_{\lambda_r} \) is dominant,

\[
\varepsilon(u_{\lambda_r}) = \ell(u) + 1, \quad \text{and} \quad \lambda(u_{\lambda_r}) = \lambda + \varepsilon_{\delta}, \quad \text{and therefore}
\]

\[
\Theta_u(x,y) = \Theta_{w_{\lambda_r}}(x,y)
\]

by the inductive hypothesis; since \( \lambda_r = \lambda_{r+1} \), it follows that

\[
\Theta_u(x,y) = \prod_{i,j \in \delta} (x_i - y_j)
\]

which is equal to \( x_1(X_{1r} - Y_{1r}), \ldots, x_m - Y_{mr} \) by (2.5) \|

(6.15) If \( w \) is Grassmannian of shape \( \lambda \) then

\[
\Theta_w(x,y) = x_1(X_{1r} - Y_{1r}), \ldots, x_m - Y_{mr}.
\]

**Proof:** This follows from (6.14) just as (4.8) follows from (4.7) \|

Finally, let \( w \) be vexillary with shape

\[
\lambda(u) = (g_1^m, \ldots, g_n^m)
\]

and flag

\[
\delta(u) = (f_1^m, \ldots, f_n^m)
\]

as in Chapter IV. Then \( w^{-1} \) is also vexillary, with shape

\[
\lambda(w^{-1}) = (g_1^m, \ldots, g_n^m)
\]

the conjugate of \( \lambda(u) \), and flag

\[
\delta(w^{-1}) = (f_1^m, \ldots, f_n^m)
\]

where by (1.41)

\[
\phi_1 + \phi_1 = f_{s+1+i} + f_{s+1+i}
\]

for all \( s \geq 1 \) and all permutations \( w \).

**Proof:** By (6.5) and (4.21) we have

\[
\sum_{w} \ell(w) \Theta_w(x,y) = |\mathcal{S}_e| \Theta_{w_0}(x,y)
\]

summed over \( w, v \) such that \( w^{-1} = v \) and \( \ell(u) + \ell(v) = \ell(u) \). By (6.3) again, the right-hand side is equal to \( \Theta_{w_0}(x,y) \).
In particular, suppose that \( w \) is vexillary. With the notation of (6.16), the flag of \( J_x \times_u (w^{-1}) \) is obtained from that of \( w \) (resp. \( w^{-1} \)) by replacing each \( f_j \) by \( f_j + r \) (resp. each \( g_j \) by \( g_j + r \)). Hence by (6.16) we have

\[
\Theta_{1, w}(x, y) = \Theta_S(x_1, (f_1 + r - 3), \ldots, (f_n + r - 3)) - \Theta_S(x_1, y_1, \ldots, y_n)
\]

and hence

\[
(6.18) \quad \Theta_{1, w}(x, y) = \Theta_S(x_1, \ldots, x_n - y_1).
\]

For all \( r \geq 1 \), where \( \rho^{(r)} \) (resp. \( \rho^{(r)} \)) is the homomorphism \( \rho \) of (4.25) acting on the \( x \) (resp. \( y \)) variables.

(6.19) Let \( \sigma_v \) (resp. \( \sigma_v \)) denote \( \sigma_\nu^{(r)} \), acting on the \( x \) (resp. \( y \)) variables. Then if \( w \) is vexillary of shape \( \lambda \), we have

\[
\sigma_v \sigma_w \Theta_{1, w}(x, y) = \Theta_S(x_1, \ldots, y_n).
\]

Proof: By (4.24) we have \( \sigma_v = \rho^{(r)} \sigma_v \) and \( \sigma_v = \rho^{(r)} \sigma_v \). Hence (6.19) follows from (6.17) and (6.18).

In particular, suppose that \( w \) is dominant of shape \( \lambda \), so that \( y \) (6.14)

\[
\Theta_{1, w}(x, y) = \prod_{j \in \lambda} (x_j - y_j) = f_s(x_1, y_1) \text{ in } y.
\]

In this case (6.19) gives

\[
\sigma_v \sigma_w \Theta_{1, w}(x, y) = \Theta_S(x_1, \ldots, y_n)
\]

for all \( r \geq 1 \), which is Sergenov's formula (3.12').

Chapter VII

Schubert Polynomials (2)

Recall the decomposition (4.17) of a Schubert polynomial \( \Theta_{w} \):

\[
\Theta_{w}(x, z_1, \ldots) = \sum_{j} \Theta_{j} \Theta_{w}(z_1, \ldots, z_j)\Theta_{j}(z_{j+1}, z_{j+2}, \ldots)
\]

Our first aim in this Chapter will be to give a method for calculating the coefficients \( \Theta_{j} \). We shall then apply our results to the calculation of \( \text{Card}(R(w)) \), the number of reduced decompositions \( w = s_n \cdots s_1 \) (where \( p = \ell(w) \)) of a permutation \( w \).

For this purpose, we introduce the operators \( \delta' \) and \( \gamma' \) defined by

\[
\delta' \Theta_{w} = \begin{cases} \Theta_{w} & \text{if } \ell(s_i w) < \ell(w), \\ 0 & \text{otherwise.} \end{cases}
\]

(7.1)

Remarks. 1. If \( w \) is the (linear) involution defined by \( \omega(\Theta_{w}) = \Theta_{w} \), for each permutation \( w \), it follows from (4.2) that \( \delta' = \omega \delta \). Hence we may define \( \delta' = \omega \delta \Theta_{w} \) for any permutation \( w \), and we have \( \delta' \Theta_{w} = \Theta_{w} - \Theta_{w} \) whenever \( (a_1, \ldots, a_n) \) is a reduced word for \( w \).

2. If \( w \in S_n \) we have \( \delta' \Theta_{w} = 0 \) for all \( i > n \), because \( \delta' \Theta_{w} = w \delta \Theta_{w} \), which is zero because \( w^{-1}(i) < w^{-1}(i+1) \).

(7.2) \( \delta' \) commutes with \( \delta \) for all \( i, j \geq 1 \).

Proof. We have

\[
\delta' \delta \Theta_{w} = \begin{cases} \delta' \Theta_{w} & \text{if } \ell(s_i w_j) = \ell(w) - 2, \\ 0 & \text{otherwise.} \end{cases}
\]

Likewise

\[
\delta \delta' \Theta_{w} = \begin{cases} \delta \Theta_{w} & \text{if } \ell(s_i w_j) = \ell(w) - 2, \\ 0 & \text{otherwise.} \end{cases}
\]

Hence \( \delta' \delta - \delta \delta' \) vanishes on each Schubert polynomial \( \Theta_{w} \), and therefore vanishes identically.
(7.3) Let $u_0 = u^{(r)}$ be the longest element of $S_n$. Then for $r = 1, 2, \ldots, n-1$ we have

$$(1 + t_0^{u_0}) \cdots (1 + t^{u_0}) \theta_{u_0} = (1 + t_0^{u_0}) \cdots (1 + t^{u_0}) \theta_{u_0}$$

as polynomials in $t, x_1, x_2, \ldots$.

Proof. The coefficient of $t^p$ for $1 \leq p \leq r$ on the left-hand side is

$$(1) \sum_{s_1, \ldots, s_r} \theta_{s_1} \cdots \theta_{s_r} \theta_{u_0}$$

summed over all reduced sequences $(s_1, \ldots, s_r)$ satisfying

$$n - r \leq s_1 \leq \ldots \leq s_r \leq n - 1.$$

Let $b_k = n - a_{k+1}$ for all $1 \leq k \leq r$, so that

$$(2) \quad 1 \leq b_1 < \ldots < b_r \leq r.$$

Let $w = s_a \cdots s_r$, so that $w^{(a)}u = n_1 \cdots n_k$. Then

$$\theta_{s_1} \cdots \theta_{s_r} \theta_{u_0} = \theta_{w^{(a)}u} = \theta_{w^{(a)}} \theta_{u_0}$$

$$= \theta_{s_1} \cdots \theta_{s_r} \theta_{u_0},$$

Hence (1) is equal to

$$\sum_{s_1, \ldots, s_r} \theta_{s_1} \cdots \theta_{s_r} \theta_{u_0}$$

summed over all reduced sequences $(s_1, \ldots, s_r)$ satisfying (1), which is the coefficient of $t^p$ on the right-hand side of (7.3)].

Next, we have

$$(7.4) \quad \theta_{t=r+1}(t, x_1, \ldots, x_{n-1}) = (1 + t_0^{u_0}) \cdots (1 + t^{u_0})(x_1, \ldots, x_{n-1}).$$

Proof. By (4.22) we have to show that

$$(1 + t_0^{u_0}) \cdots (1 + t^{u_0})(x_1, \ldots, x_{n-1}) = s_0(t, x_1, \ldots, x_{n-1})$$

where $x_i = x_1 + \cdots + x_i$ for each $i \geq 1$, and $s_0 = u_0$. For this it is enough to show that

$$(1) \quad (1 + t_0^{u_0})(x_1, \ldots, x_i, t + X_{i+1}) = s_0(x_1, \ldots, x_i, t + X_{i+1})$$

for $i = 1, 2, \ldots, n-1$.

Therefore

$$(1 + t_0^{u_0})(x_1, \ldots, x_{n-1}) = s_0(t, x_1, \ldots, x_{n-1})$$

and on the right-hand side they are $h_4(t - X_1)$, where $h_4$ runs form $n - 2i + 1$ to $2n - 3i - 1$ in each case.

Now we have

$$h_4(t - X_1) + \theta_{t=r}(X_{n+1}) = h_4(t - X_1) - \theta_{t=r}(t + X_1) + \theta_{t=r}(t + X_{n+1}) - t \theta_{t=r}(t + X_{n+1})$$

and

$$= h_4(t - X_1) + \theta_{t=r}(t + X_{n+1})$$

which agrees with (7.1) for $r = 1, 2, \ldots, n-1$.

Hence (7.1) holds, and we have

$$(7.5) \quad \theta_{t=r}(t^{(r)}x^{(r)}_1 \cdots x^{(r)}_n) = \theta_{t=r}(x_1) \theta_{u_0}(x_2, x_n).$$

Proof. Let $s = n - r + 1$ and

$$a = x_1^{s-1}x_2^{s-2} \cdots x_n, \quad b = x_1^{s-2}x_2^{s-3} \cdots x_n, \quad c = x_2 \cdots x_n$$

so that $abc = x_1^{s-1}x_2^{s-2} \cdots x_n$. Hence

$$\theta_{t=r}(a(t^{(r)}x^{(r)}_1 \cdots x^{(r)}_n)) = \theta_{t=r}(a(t^{(r)}x^{(r)}_1 \cdots x^{(r)}_n)) \quad \text{by (a.22)}$$

$$= \theta_{t=r}(t^{(r)}x^{(r)}_1 \cdots x^{(r)}_n) \quad \text{by (7.4)}$$

$$= x_1^{s-1}t - x_1^{s-1} \theta_{u_0}(x_1, x_2) \quad \text{by (4.21)}$$

$$(7.6) \quad \theta_{t=r}(x_1^{s-1}x_2^{s-2} \cdots x_n) = \theta_{t=r}(x_1^{s-1}x_2^{s-2} \cdots x_n)$$

$$= x_1^{s-1}t - x_1^{s-1} \theta_{u_0}(x_1, x_2) \quad \text{by (7.4)}$$

Let $u$ be any permutation. If $w(1) = r$, then $s_1 \cdots s_2w = 1 \times w_1$.
I. Define another (7.6)
Remark. Since (7.6)
Proof: We have

\[ \Phi_w(x_1, x_2, \ldots) = \Phi_{w_1}(x_1) \Phi_{w_2}(x_2, x_3, \ldots) \]

(7.6) summed over all \( w = s_1 \cdots s_n \), where

\[ c_i(w) + 1 = p_i + m \leq s_1 < \cdots < s_p \]

and \( f(w) = \ell(w) - \ell \). The code of \( w \) satisfies \( c_i(w) = c_i(w) \) for \( 1 \leq i \leq m - 1 \), and hence

\[ (w)_m = s_{m-n+1} \cdots s_{m-1} s_m \]

It follows that

\[ \sum w^i \Phi_w(x_1, x_2, \ldots) = x_1^{m-1} \Phi_{w_1}(x_1) \Phi_{w_2}(x_2, x_3, \ldots) \]

and therefore, by the inductive hypothesis,

\[ \Phi_n(x) = \sum_{i=0}^{\lambda(n)} x^{m-i} \cdots s_{m-1} s_m \Phi_{w_i}(x) \Phi_{w_{i+1}}(x_2, x_3, \ldots) \]

Finally, for any permutation \( w \), let \( \nu \) be the unique element of \( S_n \) such that \( \nu(1) > \cdots > \nu(m) \), and let \( \mu = \mu(w, m) \). We have \( \ell(w) = \ell(\nu) + \ell(w) \) and \( (w)_m = \nu(w) \), so that by (7.7)

\[ \Phi_w(x) = x^{m-\nu(w_1)} \cdots x_{m-n+1} \Phi_{\nu(w)}(x_2, x_3, \ldots) \]

Hence

\[ \Phi_n(x) = \Phi_{\nu(w)}(x) \]

(7.8) and let \( w_n \) be the permutation whose code is \( (c_1, c_2, \ldots) \), where \( (c_1, c_2, \ldots) \) is the code of \( w \).

With this notation established, we have

\[ \Theta_{w_n}(x) = \sum_{i=0}^{\lambda(n)} x^{m-i} \cdots s_{m-1} s_m \Phi_{w_i}(x) \Phi_{w_{i+1}}(x_2, x_3, \ldots) \]

Proof: We proceed by induction on \( m \), the case \( m = 1 \) is (7.6). From (7.6) we have

\[ \Theta_{w_n}(x) = \Phi_{w_n, s_1(\nu)}(x) \Phi_{w_{n+1}}(x_2, x_3, \ldots) \]

(\( x_1 = 1 \), \( x_m = n \))...
where $S^m(n)$ consists of the permutations whose codes have length $\leq m$, and $\eta(\alpha, \beta)$ is the constant term of the polynomial $\alpha \beta$. Applying this to (7.8), we obtain our final result:

\[ \Theta_u(x) = \sum \Theta_u(x_1, \ldots, x_m)(e^{z_u} \Phi_j(x_1, \ldots, x_m)) \Theta_{\nu_\lambda}(x_{\lambda(1)}, x_{\lambda(2)}, \ldots) \]

summed over all $u \in S^m(n)$ such that $\ell(u) = \ell_\lambda + \ell(\nu)$. [1]

For each such $u$, the constant term $\eta(\beta_u, e^{z_u} \Phi_j(x_1, \ldots, x_m))$ is a polynomial in the (non-commuting) operators $\Phi_j$ with integer coefficients. Hence (7.9) gives a decomposition of the Schubert polynomial $\Theta_u(x)$ of the form

\[ \Theta_u(x) = \sum_{\nu} \binom{z}{\nu} \Theta_{\nu}(x), \]

where $z = (x_1, \ldots, x_m)$ and $\nu = (x_{\nu(1)}, x_{\nu(2)}, \ldots)$. If $u \in S^m(n)$, we have $\Theta_u(x) = \Theta_{nu}$, then $u \in S^n(n)$ and $\nu \in S^n$ in the sum. From (4.18) we know that the coefficients $\eta(\nu)$ in (7.10) are $> 0$.

In particular, if we apply (7.7) to a permutation of the form $u^{(n)} \times u$, we shall obtain

\[ \Theta_{u^{(n)} \times u}(x) = e^{z_u} \Phi_j(x_1, \ldots, x_m) \Theta_{u^{(n)} \times u}(x_{\lambda(1)}, x_{\lambda(2)}, \ldots). \]

On the other hand, by (4.6) we have

\[ \Theta_{u^{(n)} \times u} = \Theta_{u^{(n)}} \Theta_{u} \]

and comparison of (1) and (2) gives

\[ \Theta_{u}(x) = \sum \Theta_{\nu}(x), \]

where $(\nu) = u^{(n)} \times u$. By (4.3), $\Theta_{u}(x)$ is symmetrical in $x_1, \ldots, x_m$. Hence so is the operator $\Phi_j(x_1, \ldots, x_m)$, and we may therefore write $\Phi_j$ in the form

\[ \Phi_j(x_1, \ldots, x_m) = \sum_{(\lambda, \nu)} a_{\lambda, \nu}(h, v) x_{\lambda(1)} x_{\lambda(2)} \ldots \]

summed over partitions $\lambda$ of length $\leq m$ and permutations $\nu$, with integral coefficients $a_{\lambda, \nu}(h, v)$. From (7.13) and (7.14) we have

\[ \Theta_{u}(x) = \sum_{(\lambda, \nu)} a_{\lambda, \nu}(h, v) x_{\lambda(1)} x_{\lambda(2)} \ldots \]

summed over $\lambda$ of length $\leq m$ and $\nu$ such that $\ell(\nu) = \ell(u)$. The Schur functions occurring here are precisely the Schubert polynomials $\Theta_u$, where $u$ is Grassmannian with descent at $m$. Hence, by (4.18),

\[ \Phi_j(x_1, \ldots, x_m) = \sum_{(\lambda, \nu)} a_{\lambda, \nu}(h, v) x_{\lambda(1)} x_{\lambda(2)} \ldots \]

Schubert polynomials (2)

(7.15) The coefficients $a_{\lambda, \nu}(h, v)$ in (7.12) are $\geq 0$.

Since $\Phi_j(x_1, \ldots, x_m) = \Phi_j(x_{\lambda(1)}, x_{\lambda(2)}, \ldots)$ and $s_1(x_1, \ldots, x_m) = s_1(x_1, \ldots, x_m)$ if $\ell(\lambda) \leq m$, it follows from (7.13) that

\[ a_{\lambda, \nu}(h, v) = a_{\lambda, \nu}(h, v) = \alpha(\lambda, v) \]

for all partitions $\lambda$ such that $\ell(\lambda) \leq m$.

We may also calculate the operator $\Phi_j(x_1, \ldots, x_m)$ as follows. For each integer $p \geq 1$ and each subset $D$ of $\{1, 2, \ldots, p-1\}$ let

\[ Q_D(x_1, \ldots, x_m) = \sum_{a_m \in D} x_{a_1} \ldots x_{a_m} \]

summed over all sequences $(a_1, \ldots, a_m)$ such that $1 \leq a_1 \leq \cdots \leq a_m \leq m$ and $a_m < a_{m+1}$ whenever $i \in D$. Then $Q_D(x_1, \ldots, x_m)$ is a homogeneous polynomial of degree $p$, and is zero if $m \leq \text{Card}(D)$.

Now let $u = (a_1, a_2, \ldots)$ be a reduced word, so that $\sum a_m \leq a_{m+1} = p$. The descent set of $u$ is

\[ D(u) = \{ i : a_i > a_{i+1} \}. \]

We now define, for each permutation $u$,

\[ F_u(x_1, \ldots, x_m) = \sum_{a_m \in D} Q_D(x_{a_1}, \ldots, x_{a_m}), \]

a homogeneous polynomial of degree $\ell(u)$.

With these definitions we have

\[ \Phi_j(x_1, \ldots, x_m) = \sum_{u \in S^m(n)} F_u(x_1, \ldots, x_m) a_{\lambda, \nu}(h, v). \]

Proof. Let $u = (a_1, \ldots, a_p)$ be a reduced word. Since

\[ \Phi_j(x_1, \ldots, x_m) = (1 + x_1) \cdots (1 + x_p) \cdots \]

it is clear from the definitions that the coefficient of $z_u$ is $\sum a_{\lambda, \nu}(h, v) \Phi_j(x_{\lambda(1)}, x_{\lambda(2)}, \ldots)$ is just $Q_{D(u)}(x_1, \ldots, x_m)$. Hence

\[ \Phi_j(x_1, \ldots, x_m) = \sum_{u \in S^m(n)} Q_D(x_{a_1}, \ldots, x_{a_m}) a_{\lambda, \nu}(h, v), \]

\[ = \sum_{u \in S^m(n)} F_u(x_1, \ldots, x_m) a_{\lambda, \nu}(h, v). \]
Notes on Schubert Polynomials

Comparison of (7.17) and (7.13) now shows that $P_a(x_1, \ldots, x_n)$ is a symmetric polynomial in $x_1, \ldots, x_n$, and that 

\[ F_a(x_1, \ldots, x_n) = \sum \mu(a, \lambda) x_1^{\lambda_1} \cdots x_n^{\lambda_n} \]

where $\mu(a, \lambda) = \mu(a, \lambda) = \mu(a, \lambda)$ for all $m \geq \ell(\lambda)$. 

Since the number of $x_1 \cdots x_p$ in $Q_{D(p)}(x_1, \ldots, x_n)$ is $1$ if $m \geq p$, it follows that the coefficient of $x_1 \cdots x_p$ in $F_a(x_1, \ldots, x_n)$ is equal to Card($R(u)$) whenever $m \geq \ell(u)$. On the other hand, the coefficient of $x_1 \cdots x_p$ in a Schur function $s_t$, where $|t| = p$, is equal to $f_t$, the number of standard tableaux of shape $\lambda$, or equivalently the degree of the irreducible representation $\chi^t$ of $S_t$ indexed by the partition $\lambda$ ([M], Ch. 1, §7). It follows therefore from (7.19) that 

\[ \text{Card}(R(u)) = \sum_{\lambda \vdash m} \alpha(\lambda, u) s_t \]

Remark. Since the coefficients $\alpha(\lambda, u)$ are $\geq 0$ 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 $S_n$. 

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 $R(u)$, with character $\sum \alpha(\lambda, u)s_t$. 

We shall conclude with some properties of the symmetric functions $F_a$ and the coefficients $\alpha(\lambda, u)$.

(7.21) \text{Let } u \in S_n, v \in S_n. \text{ Then}

\[ F_a(u) = F_a(v)F_a(u) \]

Proof: By (7.18), we have for any $N$,

\[ F_a(x_1, \ldots, x_N) = \sum \mu(a, \lambda) x_1^{\lambda_1} \cdots x_N^{\lambda_N} \]

where $\mu(a, \lambda) = \mu(a, \lambda) = \mu(a, \lambda)$ for all $m \geq \ell(\lambda)$. 

(7.22) \text{Let } u \in S_n, \text{ and let } \omega = \text{ the longest element of } S_n. \text{ Then}

\[ F_{\omega u} = F_{\omega} \circ F_u \]

where $\omega$ is the involution that interchanges $s_1$ and $s_2$. In other words

\[ \alpha(\lambda, u) = \alpha(\lambda, \omega) = \alpha(\lambda, \omega) \]

for all partitions $\lambda$.

For the proof of (7.22) we require a lemma. If $t$ is a standard tableau of shape $\lambda$, 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

\[ s_\lambda = \sum_{\pi \in \lambda} Q_{D(\pi)} \]

where the sum is over the standard tableaux of shape $\lambda$, and $\pi \in \lambda$.

Proof: In the notation of [M, Ch. 1, §5], $s_\lambda$ is the sum of monomials $x^t$ where $T$ runs through the (column-strict) tableaux of shape $\lambda$. Each such tableau $T$ determines a standard tableau $t$, as follows. If a square in the $j^{th}$ column of the diagram of $\lambda$ is occupied by the number $i$, replace $i$ by the pair $(i, j)$ and since $T$ is column-strict the pairs $(i, j)$ obtained are all distinct. If we now order them lexicographically (so that $(i, j)$ precedes $(i', j')$ if and only if either $i < i'$, or $i = i'$ and $j < j'$) and replace them as $1, 2, \ldots, p$, we have a standard tableau $t$, say $t = t$. It follows easily that $\sum x^t = Q_{D(\pi)}$, which proves the lemma.

If $D$ is any subset of $\{1, 2, \ldots, p - 1\}$, let $D'$ denote the complementary subset, and let $D''$ be $\{p - i : i \in D\}$. From the definition of $Q_{D''}$ we have

\[ Q_{D''}(x_1, x_2, \ldots, x_n) = Q_{D'(p)}(x_1, x_2, \ldots, x_n) \]

If $a = (a_1, \ldots, a_p) \in R(u)$, let $\hat{a} = (n - a_1, \ldots, n - a_p)$ and $a' = (n - a_1, \ldots, n - a_1)$. Then we have

\[ \hat{a} \in R(u), \quad a' \in R(u) \]

(3) \text{Moreover, if } t \text{ is a standard tableau we have}

\[ D(t) = D(t') \]
where $t'$ is the transpose of $t$, obtained by reflecting $t$ in the main diagonal. For $i \in\mathbb{N}$ if not 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 $i \notin\mathbb{N}'$.

Since $F_w$ is symmetric, it follows from (1) and (3) that

$$F_w(e_1, \ldots, e_n) = F_w(e_{n-1}, \ldots, e_1) = F_w(e_1, \ldots, e_n)$$

and hence by (7.16) that $F_w = F_{w*}$. From (7.23) and (4) above we have

$$\omega_{i_1} = a_i = \sum_{i \in \ell(\lambda)} q_{\lambda,i}$$

for all partitions $\lambda$ of $p$, where $\ell(\lambda)$ is the set of standard tableaux of shape $\lambda$, and hence it follows from (2) and (3) and the definition of $F_w$ that $\omega_{F_w} = \omega_F$. Hence

$$\omega_{F_{w*}} = F_{w*} = F_w,$$

which completes the proof of (7.23).

(7.24) Suppose $\omega(\mu, w) \neq 0$. Then the monomial $s^\mu$ occurs in $F_w$, and hence there is a reduced word $(a_1, \ldots, a_n)$ for $w$ such that

$$(1) \quad a_1 \cdots < a_{n+1} < \cdots < a_{n+k},$$

By (1.14) the code of $w$ is

$$c(w) = \sum_{i=0}^{n+k} t_{a_i} \cdots t_{a_{i+1}}(e_n).$$

If $w(1) = a_{n+1} \cdots a_{n+k}$, the sum of the first $\mu_1$ terms of this series is

$$w(1)(c_{a_{n+1}}, \ldots, c_{a_{n+k}}) + \cdots + a_{n+1} \cdots a_{n+k}(e_n),$$

and since $a_1 < \cdots < a_{n+1}$, this is equal to

$$w(1)(c_{a_1}, \ldots, c_{a_{n+1}}) + \cdots + c_{a_{n+1}} = V_1,$$

where $V_1$ is a $(0,1)$ vector (i.e., a vector with each component 0 or 1) of weight $\mu_1$. Likewise the sum of the next block of $\mu_2$ terms of the series (2) is a $(0,1)$ vector $V_2$ of weight $\mu_2$, and so on. Hence

$$c(w) = V_1 + \cdots + V_n,$$

and therefore $a_1 + 1, \ldots, a_n + 1$ are the terms of the sequence $w$ that have a smaller element somewhere to the right, in increasing order of magnitude. Hence $\mu$ has no smaller elements to the right of it, and therefore lies to the right of $c_1 + 1$, so that $t(e_{\mu_1} w) = 1$. The same argument shows that $t(e_{\mu_2} w) = 1$ and so on. Hence if $w_1 = e_{\mu_1}, \ldots, e_{\mu_1}$ we have $t(w_1) = t(w) = \mu_1$, and $A(w) = (\mu_2, \mu_3, \ldots)$. It follows by induction on $t(\mu)$ that the word $(a_1, \ldots, a_n)$ determined by the matrix $V$ is reduced, and hence $\omega(\mu, w) = 1$ when $\mu = \lambda(w')$. By (7.23) it follows that $\omega(\mu, w) = 1$ when $\mu = \lambda(w')$. (ii) Suppose now that $\mu = \lambda(w')$. Then there is only one $(0,1)$ matrix $V$ with row sums $\mu$, and column sums $\mu$. Its first row $V_1$ is $1_{\sum_{j} c_j}$ summation over $j$ such that $c_j \neq \emptyset$, i.e., such that there exists $k > j$ such that $w(k) = w(j)$. From (3) it follows that

$$wV_1 = \sum_{i=1}^{\mu} e_{a_i+1}$$

and therefore $a_1 + 1, \ldots, a_n + 1$ are the terms of the sequence $w$ that have a smaller element somewhere to the right, in increasing order of magnitude. Hence $\mu$ has no smaller elements to the right of it, and therefore lies to the right of $c_1 + 1$, so that $t(e_{\mu_1} w) = 1$. The same argument shows that $t(e_{\mu_2} w) = 1$ and so on. Hence if $w_1 = e_{\mu_1}, \ldots, e_{\mu_1}$ we have $t(w_1) = t(w) = \mu_1$, and $A(w) = (\mu_2, \mu_3, \ldots)$. It follows by induction on $t(\mu)$ that the word $(a_1, \ldots, a_n)$ determined by the matrix $V$ is reduced, and hence $\omega(\mu, w) = 1$ when $\mu = \lambda(w')$. By (7.23) it follows that $\omega(\mu, w) = 1$ when $\mu = \lambda(w')$. (iii) This follows immediately from (i) and (ii), and the characterization (1.27) of vexillary permutations."
Notes on Schubert Polynomials

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 $U = (U_i)_{i \in \mathbb{N}}$ of subspaces of $V$ such that

$$0 = U_0 \subsetneq U_1 \subsetneq \cdots \subsetneq U_n = V$$

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 $V = (V_i)_{i \in \mathbb{N}}$ 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 automorphisms 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

$$g_{ij} = \sum_{\ell=1}^{\infty} a_{i\ell} e_{\ell j}, \quad (1 \leq j \leq n)$$

then $g$ is identified with the matrix $(a_{ij})$.

The group $G$ acts on $F$: if $U = (U_i)$ and $g \in G$, then $gU$ is the flag $(gU_i)$. Let $B$ be the subgroup of $G$ that fixes the standard flag $V$. Then $g \in B$ if and only if $g_{ij}$ is a linear combination of $e_1, \ldots, e_i$, 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 $U = (U_i)$ is a sequence $(u_1, \ldots, u_n)$ in $V$ such that $u_i \in U_i \setminus U_{i-1}$ for $1 \leq i \leq n$, or equivalently such that $u_1, \ldots, u_n$ is a basis of $U_i$ for each $i$. Given such a basis of $U$, there is a unique $g \in G$ such that $g_{ij} = u_{\ell j}$ for each $i$, and we have $U = gV$. Hence $G$ acts transitively on the flag manifold $F$, and the mapping $gV \mapsto gB$ is a bijection of $F$ onto the coset space $G/B$.

For a flag $U = (U_i)$, let

$$E_i = E_i(U) = \{ j : 1 \leq j \leq n \text{ and } U_i \cap V_j \neq U_i \cap V_{j-1} \}$$
we see that there exists $k \in B$ such that $u_{w(i)} = 1$ for all $j$, or equivalently
\[ u_w = \delta_{k(i)} = 1 \text{ for } j.
\]

Hence $U = \delta_0V$ as required.

For the converse it is enough to show that (i) $\delta(wv) = w$ and (ii) $\delta(\delta U) = \delta(U)$ for all $k \in B$ and $U \in F$. As to (i), $wV \cap V_j$ is spanned by the basis vectors $e_{w(k)}$ such that $k \leq i$ and $w(k) \leq j$, and therefore $wV_i \cap V_j \neq wV_{i+1} \cap V_j$ if and only if $j = w(k)$ for some $k \leq i$. Thus the set $E_i(wV)$ consists of $w(1), \ldots, w(i)$, which establishes (i). Finally as to (ii), we have $wV_i \cap V_j = wV_{i+1} \cap V_j$ if $k \in B$, so that $E_i(\delta U) = E_i(U)$ and hence $\phi(\delta V) = \phi(U)$ as required.

From (A.2) we have immediately

(A.3) [Bruhat decomposition] $G$ is the disjoint union of the double cosets $BuB, w \in S_n$.

For each $w \in S_n$, let
\[ C_w = (BuB)/B \subset G/B = F.
\]

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 $U \in F$. Then $U \in C_w$ if and only if $U$ has a basis $(v_1, \ldots, v_n)$ such that $v_i \in V_{w(i)} - V_{w(i) - 1}$ for each $i$. We may normalize the $v_i$ by taking
\[ u_w = \sigma_{w(i)} + \text{lower terms}.
\]

We can then subtract from $u_w$ suitable multiples of the $v_k$ for which $k < i$ and $w(k) < w(i)$, so as to make the coefficient of $e_{w(k)}$ in $u_w$ zero for each such $k$. Then $u_w$ is replaced by a vector of the form
\[ u_w = \sum_j a_j e_j,
\]

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 $U \in F$. Then $U \in C_w$ if and only if $U$ has a basis $(v_1, \ldots, v_n)$ of the form
\[ u_w = \sigma_{w(i)} + \sum_j a_j e_j,
\]

where the sum is over all $j$ in the $i$th row of the diagram of $w$, and the coefficients $a_j$ are arbitrary elements of the field $K$. Moreover, the $a_j$ are uniquely determined by the flag $U$, and the mapping $C_w \rightarrow \mathbb{K}^{2n+1}$ so defined is a bijection.
\textbf{Proof:} Clearly each "matrix" $u = (u_i)$ of shape $(n,\ldots, n)$ determines a basis $(v_1, \ldots, v_n)$ of $V$ as above, and hence a flag $U \subseteq C_w$. If $v^* = (v_i^*)$ determines $(v_1^*, \ldots, v_n^*)$ and the same flag $U$, then each $v_i^*$ must be expressible as
\[ v_i^* = u_i + \sum_{j \neq i} u_j v_j, \]
and from the form of $v_i^*$ and the $u_j$ it follows that $v_i^* = u_i$ for each $i$, and hence $v^* = u$.

Since $\text{Card} D(u) = f(u)$ it follows from (A.1) that the Schubert cell $C_u$ is isomorphic to affine space of dimension $f(u)$.

Let $U \in F$ and let $(v_1, \ldots, v_n)$ be any basis of $U$. Since $v_1, \ldots, v_n$ is a basis for $U$, for each $i = 1, \ldots, n$, the flag $U$ determines each of the exterior products $v_1 \wedge \cdots \wedge v_i \in V^i$ up to a nonzero scalar multiple, and hence $U$ determines the vector
\[ v_i = (v_1 \wedge v_i) \wedge \cdots \wedge (v_1 \wedge \cdots \wedge v_n) \in V^i \]
up to a nonzero scalar multiple, where $E = V \oplus V^2 \oplus \cdots \oplus V^i$. 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
\[ \pi: F \to P(E) \]
(the Plücker embedding) for which $\pi(U)$ 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 $\pi$ 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 $C_u$ is a locally closed subvariety of $F$, isomorphic to affine space of dimension $f(u)$.

For each $w \in S_n$, let
\[ X_w = C_w \]
be the closure of $C_w$ in $F$. The $X_w$ are the Schubert varieties in $F$, and a flag $U$ lies in $X_w$ if and only if $U$ has a basis $(v_1, \ldots, v_n)$ such that $v_i \in V_{4(i)}$ for each $i$. Each $X_w$ is in fact a union of Schubert cells $C_{s_o}$: if $(a_1, \ldots, a_k)$ is a reduced word for $w$, then $C_w \subseteq L$ if and only if $w$ is of the form $s_{a_1} \cdots s_{a_k}$ where $(a_1, \ldots, a_k)$ is a subsequence of $(1, \ldots, n)$, that is to say if and only if $w = \sigma_1 \cdots \sigma_{n-1}$ in the Bruhat order. In particular, $X_{1} = C_1$ is the single point $V \in F$. 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 $F$ in $f(u) = f(n-1)$.

Let $H^*(F; \mathbb{Z})$ be the cohomology ring (with integral coefficients) of the projective variety $F$. Each closed subvariety $X$ of $F$ determines an element $[X] \in H^*(F; \mathbb{Z})$, and cup-product in $H^*(F; \mathbb{Z})$ corresponds, roughly speaking, to intersections of subvarieties. In particular, for each $w \in S_n$, we have a cohomology class $[X_w] \in H^*(F; \mathbb{Z})$, and it is a consequence of the cell decomposition (A.3) of $F$ that the $[X_w]$ form a basis of $H^*(F; \mathbb{Z})$. In particular, $[X_{w_0}]$ is the identity element.

The connection between the classes $[X_w]$ and the Schubert polynomials $\Theta_w(u) \in S_n$ is given by
\[ \alpha: \mathbb{Z}[x_1, \ldots, x_n] \to H^*(F; \mathbb{Z}) \]
such that
\[ \alpha(\Theta_w) = [X_{w_0}] \]
for each $w \in S_n$.

\textbf{Proof:} Let us temporarily write
\[ \sigma_w = [X_{w_0}] \]
for $w \in S_n$. Monk [60] proved that for all $w \in S_n$ and $r = 1, \ldots, n-1$
\[ \sigma_w \cdot \sigma_r = \sum \sigma_{w'}. \]
where the sum on the right hand side is over all transpositions $t = (i, j)$ such that $i < j \leq n$ and $f(w) = f(u) + 1$, as in (4.15).

Define $\xi_1, \ldots, \xi_n \in H^*(F; \mathbb{Z})$ by
\[ \xi_1 = \sigma_1 \]
\[ \xi_i = \sigma_i \sigma_{i-1} \quad (2 \leq i \leq n-1) \]
\[ \xi_n = \sigma_{n-1} \]
From (1) we deduce the counterpart of (4.15): if $r$ is the last descent of $w$ (so that $r \leq n-1$), then we have
\[ \sigma_w = \sigma_r \xi_r + \sum \sigma_{w'}. \]
where $\sigma_w$ as in (4.15). Now iteration of (4.16) will ultimately express $\Theta_w$ as a sum of monomials, i.e. as a polynomial in $x_1, \ldots, x_{n-1}$, and iteration of (2) will express $\sigma_r$ as the same polynomial in $x_1, \ldots, x_{n-1}$. Hence if we define $a_i : P_n = H^*(F; \mathbb{Z})$ by $a_i = \xi_i \ (1 \leq i \leq n)$, we have $\sigma_w = a(\Theta_w)$ for all $w \in S_n$, and the proof of (A.5) is complete. \[ \square]
Notes on Schubert Polynomials

In fact the kernel of the homomorphism $\phi$ is generated by the elementary symmetric functions $e_1, \ldots, e_n$ of the $r_i$.

We shall draw one consequence of (A.6) that we have not succeeded in deriving directly from the definition (4.1) of the Schubert polynomials. Since the $e_w, w \in S_n$, form a $\mathbb{Z}$-basis of $H^*(F; \mathbb{Z})$, any product $e_w e_v(u, v \in S_n)$ is uniquely a linear combination of the $e_w$, and it follows from intersection theory on $F$ that the coefficient of $e_w$ in $e_w e_v$ is a non-negative integer. From this we deduce

(A.6) Let $u, v$ be permutations, and write $\Theta_u \Theta_v$ as an integral linear combination of the $\Theta_w$, say

$$\Theta_u \Theta_v = \sum_w c_{uv}^{w} \Theta_w.$$  

Then the coefficients $c_{uv}^{w}$ are non-negative.

We have only to choose $u$ 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 $\phi$ of (A.5).

Remark. The coefficients $c_{uv}^{w}$ in (A.6) are zero unless

(a) $\phi(u) = \phi(v)$,
(b) $u \leq v$ and $v \leq w$.

For $\Theta_u \Theta_v$ is homogeneous of degree $\phi(u) + \phi(v)$, which gives condition (a). Also we have

$$c_{uv}^{w} = \xi_n ^{-1} \Theta_w \Theta_v$$

by (2.17), and the only possible nonzero term in this sum is that corresponding to $n_1 = n$. 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 theorem (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 $\Theta_w$ is well-defined for all permutations $w \in S_n$. 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) is appropriately modified, is valid for any root system, and it is a more general form which 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 $(e_w)$ $(e_v)$ are indicated in [LS6].

Chapter VI. Double Schubert polynomials were introduced in [L2]. For the interpolation formulas (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 [LS7]. Propositions (7.21)-(7.24) are due to Stanley [S].


Notes on Schubert Polynomials


