JOURNAL OF COMBINATORIAL THEORY, Series A **60**, 168–182 (1992)

A Combinatorial Construction of the Schubert Polynomials

NANTEL BERGERON

Princeton University, Department of Mathematics Fine Hall-Washington Road Princeton, New Jersey 08544-1000

Communicated by the Managing Editors

Received June 8, 1990; revised February 20, 1991

We show a combinatorial rule based on diagrams (finite nonempty sets of lattice points (i, j) in the positive quadrant) for the construction of the Schubert polynomials. In the particular case where the Schubert polynomial is a Schur function we give a bijection between our diagrams and column strict tableaux. A different algorithm had been conjectured (and proved in the case of vexillary permutations) by A. Kohnert (Ph.D. dissertation, Universität auf Bayreuth, 1990). We give, at the end of this paper, a sketch of how one would show the equivalence of the two rules.

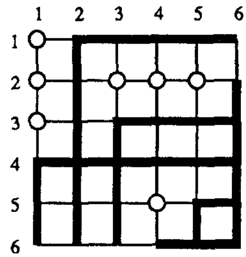
0. INTRODUCTION

Schubert polynomials where first introduced in an algebraic geometric context. They arise in the study of the flag manifolds and the Schubert varieties [2, 3]. Subsequently A. Lascoux and M.-P. Schutzenberger developed, in an impressivve list of papers [5–14], an elegant theory of these polynomials. Our present contribution is to give a combinatorial rule to construct the above-mentioned polynomials. We first recall here (without proof) the basic facts of the theory of Schubert polynomials needed for our investigation. The interested reader may find an excellent overview along with *complete proof* of this theory in a recent book by I. G. Macdonald [16].

Let $w = (w_1, w_2, ..., w_n)$ be a permutation in the symmetric group S_n on n elements. We denote by $l(w)$ the *length* (number of inversions) of w. For $1 \leqslant i \leqslant n-1$ let s_i denote the transposition that interchanges i and $i+1$ and fixes all other elements. We say that a decomposition $w = s_{a_1} s_{a_2} \cdots s_{a_p}$ is *reduced* if $p = l(w)$.

0097-3165/92 $5.00

The *diagram* $D(w)$ of a permutation w is the set of points $(i,j)\in[1\ldots n]^2$ for which $i<w_j^{-1}$ and $j<w_i$. Here $[1\ldots n]$ denotes the set $\{1,2,\ldots,n\}$. Graphically, we obtain $D(w)$ by removing in $[1\ldots n]^2$ all points south or east of a point (i,w_i). For example, if $w=(2,6,3,1,5,4)$, the diagram $D(w)$ consists of all the circled points in the picture below.



If w, a permutation in S_n, is embedded in S_{n+1} by $w'=(w_1,w_2,\ldots,w_n,n+1)$ then $D(w')=D(w)$. That is to say, the diagram of w is stable under such embedding of symmetric groups. Hence we may define for $w\in S_\infty=\bigcup_{n\geqslant 1}S_n$ the diagram $D(w)\subset N^2$. One can show that the diagram of a permutation w is such that

$$|D(w)|=l(w), \tag{0.1}$$

where $|D(w)|$ denotes the cardinality of $D(w)$.

For $w\in S_n$ and $i\in[1\ldots n]$ let $c_i(w)$ denote the number of elements of $D(w)$ in the ith row of $D(w)$. Let $\lambda(w)$ be the sequence of the numbers $c_i(w)$ arranged in decreasing order. The sequences $c(w)=(c_1(w),c_2(w),\ldots,c_n(w))$ and $\lambda(w)$ are called, respectively, the *code* and the *shape* of w.

A permutation w is called *vexillary* if there is no $a,b,c,d\in[1\ldots n]$ such that $a<b<c<d$ and $w_b<w_a<w_d<w_c$. If $c(w)=\lambda(w)$ we say that w is *dominant*. Finally, if $w_1<\cdots<w_r$ and $w_{r+1}<\cdots<w_n$ for some r we say that w is *grassmannian*. Note that both dominant and grassmannian permutations are particular cases of vexillary permutations. We should point out that the probability that $w\in S_n$ is vexillary is at most $\frac{23}{24}^{[n/4]}$ [16]; hence when n is large, the number of vexillary permutations is small compared to the total number of permutations.

Let $P=\mathbf{Z}[x_1,x_2,x_3,\ldots]$ denote the ring of polynomials in infinitely many variables with coefficients in $\mathbf{Z}$. For any $i\geqslant 1$ we define the linear operator *divided difference*

$$\partial_i f(x_1,x_2,\ldots)=\frac{f(x_1,\ldots,x_i,x_{i+1},\ldots)-f(x_1,\ldots,x_{i+1},x_i,\ldots)}{x_i-x_{i+1}}.$$

By direct calculation we have that

$$\partial_i^2 = 0,$$
$$\partial_i \partial_j = \partial_j \partial_i \qquad \text{if } |i-j| > 1,$$
$$\partial_i \partial_{i+1} \partial_i = \partial_{i+1} \partial_i \partial_{i+1}.$$

From this one may deduce that the operator

$$\partial_w = \partial_{a_1} \dots \partial_{a_p} \tag{0.2}$$

is well defined for any reduced decomposition $w = s_{a_1} s_{a_2} \cdots s_{a_p}$. Moreover, if $s_{a_1} s_{a_2} \cdots s_{a_q}$ is not reduced then

$$\partial_{a_1} \dots \partial_{a_q} = 0. \tag{0.3}$$

Let $\delta = (n-1, n-2, \dots, 1, 0)$ and let $x^\delta = x_1^{n-1} x_2^{n-2} \cdots x_{n-1}$. For each permutation $w \in S_n$ the *Schubert polynomial* $\mathcal{S}_w$ is defined to be

$$\mathcal{S}_w = \partial_{w^{-1} w_0} x^\delta,$$

where $w_0 = (n, n-1, \dots, 2, 1)$ is the longest (with respect to $l(—)$) element of S_n. From (0.2) and (0.3) one can prove that

$$\partial_i \mathcal{S}_w = \begin{cases} \mathcal{S}_{ws_i} & \text{if } \; w_i > w_{i+1}, \\ 0 & \text{if } \; w_i < w_{i+1}. \end{cases} \tag{0.4}$$

We give below a list of properties for the Schubert polynomials:

$$\mathcal{S}_{w_0} = x^\delta, \qquad \mathcal{S}_1 = 1. \tag{P.1}$$

For each $w \in S_n$, $\mathcal{S}_w$ is homogeneous of degree $l(w)$. Moreover,

$$\mathcal{S}_w = \sum_\alpha c_\alpha x^\alpha, \tag{P.2}$$

where $\alpha \in \mathbf{N}^{n-1}$, $\alpha \subset \delta$ (i.e., $\alpha_i \leqslant n-i$ for each i) and $c_\alpha \in \mathbf{N}$:

$$\mathcal{S}_{s_i} = x_1 + x_2 + \cdots + x_i. \tag{P.3}$$

Let $i: S_n \hookrightarrow S_{n+1}$ be the embedding $w \to (w_1, w_2, \dots, w_n, n+1)$. Then

$$\mathcal{S}_w = \mathcal{S}_{i(w)} \tag{P.4}$$

for all $w \in S_n$. It follows that $\mathcal{S}_w$ is a well-defined polynomial for each permutation $w \in S_\infty = \bigcup_n S_n$. If $u \in S_m$ and $v \in S_n$, we denote by $u \times v$ the permutation $(u(1), \dots, u(m), v(1)+m, \dots, v(n)+m)$ in S_{m+n}. We have then

$$\mathcal{S}_{u \times v} = \mathcal{S}_u \cdot \mathcal{S}_{1_m \times v}, \tag{P.5}$$

where 1_m is the identity element of S_m. If w is dominant of shape λ then

$$\mathscr{S}_w = x^\lambda. \tag{P.6}$$

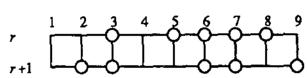
If w is Grassmannian of shape λ, then

$$\mathscr{S}_w = S_\lambda(x_1, x_2, ..., x_r), \tag{P.7}$$

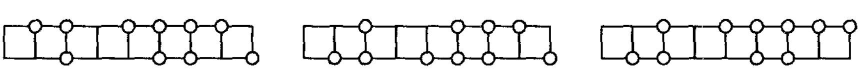
where $S_\lambda(X_r)$ is the Schur function [15] indexed by λ and r is the unique descent of w. Finally, if w is vexillary then $\mathscr{S}_w$ is a *multi-Schur function.* We refer the reader to [16 or 17] for more details.

1. Combinatorial Construction

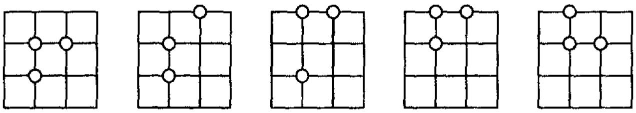
Here a "*diagram*" is any finite nonempty set of lattice points (i, j) in the positive quadrant $(i \geqslant 1, j \geqslant 1)$. For example, the diagram $D(w)$ of a permutation w is a diagram in the above sense. A *labeled diagram* is a diagram D with a label $e_{i,j} \geqslant 1$ for each $(i, j) \in D$. For example, we can label $D(w)$ by setting $e_{i,j} = i$. Let D be any labeled diagram. We denote by $D_{(r,r+1)}$ the labeled diagram D restricted to the row r and $r+1$. Let $j(r, D) = (j_1, j_2, ..., j_k)$ be the columns of D in which there is exactly one element of $D_{(r,r+1)}$ per column. Choose a column $j_i \in j(r, D)$. Assume first that $(r+1, j_i) \in D_{(r,r+1)}$. If $i = k$ or $(r, j_{i+1}) \in D_{(r,r+1)}$, let D_1 be the diagram obtained from D by replacing the element $(r+1, j_i)$ by (r, j_i) keeping the same labeling as in D. Now suppose instead that $(r, j_i) \in D_{(r,r+1)}$. If $i = 1$ or $(r+1, j_{i-1}) \in D_{(r,r+1)}$ and if $e_{r,j_i} > r$, let D_1 be the diagram obtained from D by replacing the element (r, j_i) by $(r+1, j_i)$ and labeled by the following rule. Let $e = e_{r,j_i}$ in D, we put in D_1 the same labels as in D except if $e_{r+1,j} > e$ with $j \leqslant j_i$ in D then $e_{r+1,j}$ is re-labeled by e in D_1. In both cases we say that the diagram D_1 is obtained from D by a "*B-move.*" For example, let D be such that $D_{(r,r+1)}$ is



with labels $e_{i,j}$. For this case, $j(r, D) = (2, 5, 8, 9)$. If $e_{r,5} > r$, we can perform on this diagram a B-move in column 2, 5, or 9 and obtain, respectively, the following diagrams:



In the second diagram, the labels $e_{r+1,2}$ and $e_{r+1,3}$ are the only labels subject to some re-labeling. The element in column 8 is not allowed to move since $(r+1, 5) \notin D_{(r,r+1)}$. Let $\Omega(w)$ denote the set of all diagrams (including $D(w)$) obtainable from $D(w)$, with labels $e_{i,j} = i$, by any sequence of B-moves. In $\Omega(w)$, we forget about the labels. For example, $\Omega(1, 4, 3, 2)$ is



Here the labels in the last diagram were $e_{1,2} = 2$, $e_{2,2} = 2$, and $e_{2,3} = 2$.

Next for $D \in \Omega(w)$ let x^D denote the monomial $x_1^{a_1} x_2^{a_1} x_3^{a_3} \cdots$, where a_i is the number of elements of D in the ith row. For any permutation w we have

THEOREM 1.1.

$$\mathscr{S}_w = \sum_{D \in \Omega(w)} x^D. \tag{1.1}$$

To prove this theorem we will proceed by reverse induction on $l(w)$. If $w = w_0$ (the longest element of S_n) then (1.1) holds, since $\Omega(w_0)$ contains only the element $D(w_0)$ and $x^{D(w_0)} = x^\delta$. On the other hand, from (P.1), $\mathscr{S}_{w_0} = x^\delta$. Now if $w \neq w_0$ then let $r = \min\{i \colon w_i < w_{i+1}\}$. From (0.4) we have

$$\mathscr{S}_w = \partial_r \mathscr{S}_{ws_r}. \tag{1.2}$$

Let $v = ws_r$. By the induction hypothesis equation (1.1) holds for $\mathscr{S}_v$. The induction step will be to "*apply*" the operator ∂_r to the diagrams in $\Omega(v)$. To this end we need more tools.

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

$$\partial_r x^D = \partial_r \cdots x_r^a x_{r+1}^b \cdots = \begin{cases} \displaystyle\sum_{k=0}^{a-b-1} \cdots x_r^{a-k-1} x_{r+1}^{b+k} \cdots & \text{if } a > b, \\ 0 & \text{if } a = b, \\ \displaystyle -\sum_{k=0}^{b-a-1} \cdots x_r^{a+k} x_{r+1}^{b-k-1} \cdots & \text{if } a < b. \end{cases} \tag{1.3}$$

This suggests we define the operator ∂_r directly on the diagram D. For this we need only to concentrate our attention on the rows r and $r+1$ of D. Let $j(r, D) = (j_1, j_2, ..., j_p)$. Notice that in all columns $j < w_r$ of $D_{(r,r+1)}$ there are exactly two elements and in column $w_r = j_1$ of $D_{(r,r+1)}$ there is exactly one element in position (r, j_1). We shall now reduce the sequence of indices $j(r, D)$ according to the following rule. Let $J_{(0)} = (j_2, j_3, ..., j_p)$. Remove from $J_{(0)}$ all pairs j_k, j_{k+1} for which $(r, j_k) \in D$ and $(r+1, j_{k+1}) \in D$. Let us denote the resulting sequence by $J_{(1)}$. Repeat recursively this process on $J_{(1)}$ until no such pair can be found. Let us denote by $f(r, D) = (f_1, f_2, ..., f_q)$ the final sequence. From construction, the sequence $f(r, D)$ is such that if $(r, f_k) \in D$ then $(r, f_{k+1}) \in D$. Let $up(r, D)$ be the minimal k such that $(r, f_k) \in D$. If $(r+1, f_q) \in D$ then set $up(r, D) = q+1$. We are now in a position to define the operation of ∂_r on the diagram D. To this end let us first assume that $a > b$. This means that we have $a-b$ more elements in row r then in row $r+1$. Hence $q - up(r, D) + 1 \geqslant a - b - 1$ for q the length of $f(r, D)$. The equality holds if and only if $up(r, D) = 1$. In the case $a > b$ the operator ∂_r on the diagram D is defined by the map

$$\partial_r D \rightarrow \{D_0, D_1, D_2, ..., D_{a-b-1}\} \tag{1.4a}$$

where D_0 is identical to D except that we remove the element in position (r, w_r) and for $k = 1, 2, ..., a-b-1$ we successively set D_k to be identical to D_{k-1} except that the element $(r, f_{up(r,D)+k-1})$ is replaced by $(r+1, f_{up(r,D)+k-1})$. Now if $a < b$ we have $up(r, D) - 1 \geqslant b - a + 1$ (with equality iff $up(r, D) = q+1$). So $up(r, D) - 1 > b - a$. In this case the operator ∂_r on the diagram D is defined by the map

$$\partial_r D \rightarrow \{D_0, D_1, D_2, ..., D_{b-a-1}\} \tag{1.4b}$$

where D_0 is identical to D except that we remove the element in position (r, w_r) and the element $(r+1, f_{up(r,D)-1})$ is replaces by $(r, f_{up(r,D)-1})$. For $k = 1, 2, ..., b-a-1$ we successively set D_k to be identical to D_{k-1} except that the element $(r+1, f_{up(r,D)-k-1})$ is replaced by $(r, f_{up(r,D)-k-1})$. Finally, if $a = b$ then

$$\partial_r D \rightarrow \{\ \}. \tag{1.4c}$$

With this definition of ∂_r we have that

$$\partial_r x^D = \pm \sum_{D_i \in \partial_r D} x^{D_i}. \tag{1.5}$$

We shall now show the following proposition.

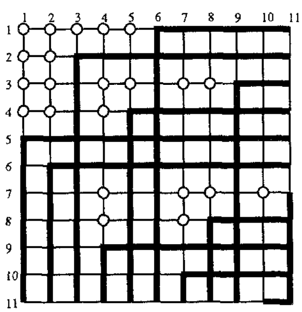
PROPOSITION 1.2.

$$\partial_r \textit{ maps } \Omega(v) \textit{ into } \Omega(w). \tag{1.6}$$

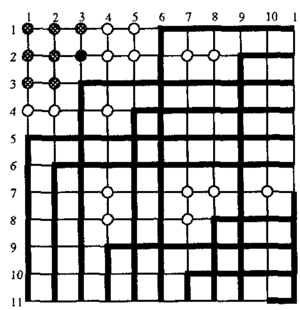
Proof. The reader shall notice that in $D(v)$ the rectangle defined by the rows $1, 2, .., r+1$ and the columns $1, 2, ..., w_r - 1$ is filled with elements.

None of these elements can B-move. Hence these elements are fixed in any diagram $D \in \Omega(v)$. The same applies to all elements in column w_r; they are packed in the smallest rows and there are no elements in the rows strictly greater than r. Now let D be a diagram in $\Omega(v)$ and assume that $\partial_r D = \{D_0, D_1, ..., D_m\}$ is non-empty. The remark above implies that the element in position (r, w_r) does not affect the sequence of B-moves from $D(v)$ to D. Hence we can apply the same sequence of B-moves to $D(v) - \{(r, w_r)\}$ and obtain D_0. Moreover, $D(v) - \{(r, w_r)\}$ is obtainable from $D(w)$ by a simple sequence of B-moves in rows r, $r+1$; for this one successively B-moves all the elements in row $r+1$ and columns given by $j(r, D(w))$. This gives that D_0 is obtainable from $D(w)$ by a sequence of B-moves, that is, $D_0 \in \Omega(w)$. Now from the construction of $\partial_r D$, D_k $(k > 0)$ is obtained from D_{k-1} by exactly one B-move. Note that in $D_{k-1} \in \Omega(w)$ an element in row r and column $j \geqslant w_r$ has a label $> r$. Hence $\partial_r D \subset \Omega(w)$. ∎

It is appropriate at this point to give an example. Let $w = (6, 3, 9, 5, 1, 2, 11, 8, 4, 7, 10)$. Hence $r = 2$ and $v = (6, 9, 3, 5, 1, 2, 11, 8, 4, 7, 10)$. We have depicted below the diagrams $D(w)$ and $D(v)$. In our example the fixed elements described above are colored in grey and the element in position (r, w_r) is colored black.



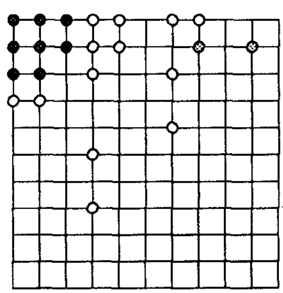
$D(w)$



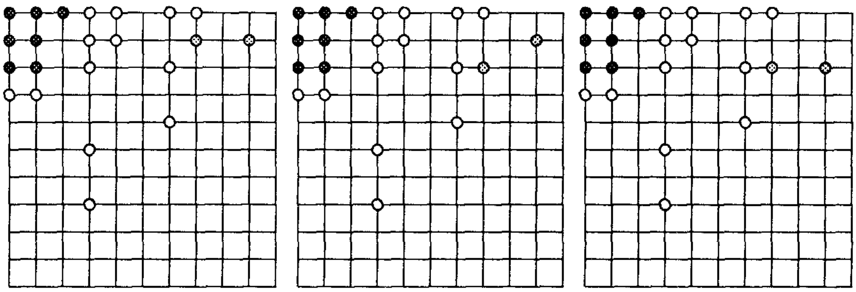
$D(v)$

Now let D be the following diagram of $\Omega(v)$.

D =



Here, $a_r(D)=7$, $a_{r+1}(D)=4$ and $j(r, D)=(3, 5, 7, 8, 10)$. The reduced sequence $f(r, D)$ is $(8, 10)$ and $up(r, D)=1$ ($e_{2,8}=7$ and $e_{2,10}=7$). Hence $\partial_r D=\{D_0, D_1, D_2\}$, where



D_0 D_1 D_2

To proceed in the induction step of Theorem 1.1 we first find a subset of $\Omega(v)$ such that when we operate with ∂_r we obtain $\Omega(w)$. To this end let

$$\Omega_0(v)=\{D\in\Omega(v): a_r(D)>a_{r+1}(D) \text{ and } up(r, D)=1\}.$$

We have

PROPOSITION 1.3.

$$\Omega(w)=\bigcup_{D\in\Omega_0(v)}\partial_r D \qquad (\textit{disjoint union}). \tag{1.7}$$

Proof. It is clear from construction that all the sets $\partial_r D$ are disjoint from each other for $D\in\Omega_0(v)$. From (1.6) we only have to prove that for any $D'\in\Omega(w)$ there is a $D\in\Omega_0(v)$ such that $D'\in\partial_r D$. To see that, reduce the sequence $j(r, D')=(j_1, ..., j_p)$ by removing recursively all pairs j_k, j_{k+1} for which $(r, j_k)\in D'$ and $(r+1, j_{k+1})\in D'$. Denote the final sequence by $f'(r, D')$. Let D be the diagram obtained from D' by adding an element in position (r, w_r) and successively B-moving all elements in positions $(r+1, f_i)\in D'$ to (r, f_i). We have that $D\in\Omega(v)$. To see this one applies to $D(v)$ the sequence of B-moves from $D(w)$ to $D-\{(r, w_r)\}$. Of course, one should ignore any B-move in rows r, $r+1$ performed on the original elements of $D(v)$ in row r. But by the choice of r, the other B-moves apply almost directly and the resulting diagram is precisely D. Moreover, since $f(r, D)=f'(r, D')$ and $up(r, D)=1$ we have $D\in\Omega_0(v)$ and $D'\in\partial_r D$. ∎

We now investigate the effect of ∂_r on $\Omega_1(v) = \Omega(v) - \Omega_0(v)$. More precisely we have

PROPOSITION 1.4.

$$\sum_{D \in \Omega_1(v)} \partial_r x^D = 0. \tag{1.8}$$

Proof. There are two classes of diagrams in $\Omega_1(v)$. The first class contains the diagrams D for which $a_r(D) = a_{r+1}(D)$. In this case it is trivial that $\partial_r x^D = 0$. The other class is formed by the diagrams D such that $a_r(D) \neq a_{r+1}(D)$ and $up(r, D) > 1$. In this case we shall construct an involution, $D \to D'$, such that $\partial_r x^D + \partial_r x^{D'} = 0$. Let $f(r, D) = (f_1, f_2, ..., f_q)$, $a = a_r(D)$ and $b = a_{r+1}(D)$. We first define the involution for the case $a > b$. Since $up(r, D) > 1$ we must have $q - up(r, D) + 1 \geqslant a - b$. So let D' be identical to D except that the elements in positions $(r, f_{up(r,D)})$, $(r, f_{up(r,D)+1})$, ..., $(r, f_{up(r,D)+a-b-1})$, are B-moved to the positions $(r+1, f_{up(r,D)})$, $(r+1, f_{up(r,D)+1})$, ..., $(r+1, f_{up(r,D)+a-b-1})$. It is clear that $D' \in \Omega(v)$. But $f(r, D') = f(r, D)$ and $up(r, D') > up(r, D) > 1$; hence $D' \in \Omega_1(v)$. Moreover, we have $a_r(D') = b$ and $a_{r+1}(D') = a$; hence $\partial_r x^D + \partial_r x^{D'} = 0$. The case $a < b$ is similar to the previous one. ∎

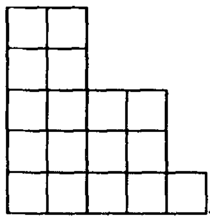
A proof of Theorem 1.1 is now completed by combining (1.2), (1.5), (1.7), and (1.8). More precisely, using the induction hypothesis, we have

$$\begin{aligned}
\mathcal{S}_w &= \partial_r \mathcal{S}_v && (1.2)\\
&= \sum_{D \in \Omega(v)} \partial_r x^D \\
&= \sum_{D \in \Omega_0(v)} \partial_r x^D && (1.8)\\
&= \sum_{D \in \Omega_0(v)} \sum_{D_i \in \partial_r D} x^{D_i} && (1.5)\\
&= \sum_{D' \in \Omega(w)} x^{D'} && (1.7).
\end{aligned}$$ ∎

Remark 1.5. If one considers (1.1) as the definition for the Schubert polynomials then the properties (P.1) through (P.6) are almost immediate! (Some of them are quite intricate to prove using the original definition.) The property (P.7) is the object of the next section.

2. Schur Functions and Diagrams

A *partition* λ of an integer n is a decreasing sequence of integers $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_k \geqslant 0$ such that $n = \lambda_1 + \lambda_2 + \cdots + \lambda_k$. A *Ferrers diagram* of shape λ is the set of lattice points (i, j) such that $1 \leqslant i \leqslant k$ and $1 \leqslant j \leqslant \lambda_i$. We usually represent these diagrams using boxes in the positions (i, j). For example, if $\lambda = (5, 4, 4, 2, 2)$, the Ferrers diagram is



A *column strict tableau* of shape λ is a filling of the Ferrers diagram of shape λ with positive integers such that they are increasing from left to right in the rows of λ and strictly increasing from bottom to top in the columns. For example, the following is a column strict tableau of shape $\lambda = (5, 4, 4, 2, 2)$:

$$\begin{array}{|c|c|c|c|c|} \hline 7 & 7 \\ \cline{1-2} 6 & 6 \\ \cline{1-4} 4 & 5 & 6 & 6 \\ \cline{1-4} 2 & 3 & 3 & 5 \\ \hline 1 & 1 & 1 & 2 & 4 \\ \hline \end{array}. \tag{2.1}$$

For t a column strict tableau, let us denote by x^t the monomial $x_1^{\mu_1} x_2^{\mu_2} x_3^{\mu_3} \cdots$, where μ_i is the number of occurrences of the integer i in t. For example, if t is the tableau in (2.1) then x^t is

$$x_1^3 x_2^2 x_3^2 x_4^2 x_5^2 x_6^4 x_7^2.$$

With this notation one may define the *Schur function* indexed by λ as [15]

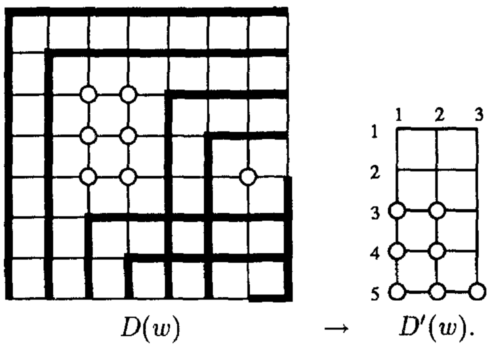
$$S_\lambda(x_1, ..., x_r) = \sum_t x^t, \tag{2.2}$$

where the sum is taken over all column strict tableaux of shape λ, filled with integers between 1 and r.

We shall show, using Theorem 1.1, that if $w \in S_n$ is grassmannian of shape λ with descent at r, we have

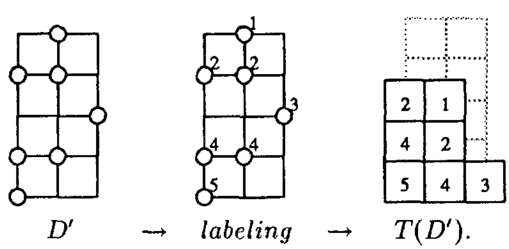
$$\mathscr{S}_w = S_\lambda(x_1, x_2, ..., x_r). \tag{2.3}$$

To this end let us fix $w \in S_n$ grassmannian of shape $\lambda = (\lambda_1, \lambda_2, ..., \lambda_k)$ with descent at r ($r \geqslant k$). Note that the diagram $D(w)$ has λ_i elements in the row $r-i+1$. Moreover, for $2 \leqslant i \leqslant k$, if $(j_1, j_2, ..., j_{\lambda_i})$ and $(j'_1, j'_2, ..., j'_{\lambda_{i-1}})$ are the columns where $D(w)$ contains an element in row $r-i+1$ and $r-i+2$, respectively, then $j_s = j'_s$ for $1 \leqslant s \leqslant \lambda_i$. Hence we can consider, without loss of generality, the diagram $D(w)$ as a diagram $D'(w)$ included in $[1 ... r] \times [1 ... \lambda_1]$. For example, if $w = (1, 2, 5, 6, 8, 3, 4, 7)$ then



Denote by $\Omega'(w)$ the set of diagrams obtainable from $D'(w)$ by any sequence of B-moves. It is clear that we have a one-to-one correspondence between the diagrams $D \in \Omega(w)$ and $D' \in \Omega'(w)$. Moreover, $x^D = x^{D'}$.

Choose $D' \in \Omega'(w)$. Label the elements (i, j) of D' with the integer "i." Then let all element of D' (along with their labels) fall back into their positions in $D'(w)$. We obtain in this manner a Ferrer diagram of shape λ filled with integers between 1 and r. Let us denote by $T(D')$ the resulting labeled diagram. For example,



THEOREM 2.1. *If w is grassmannian of shape λ then for $D' \in \Omega'(w)$*

the labeled diagram $T(D')$ is strictly decreasing in the columns of λ and decreasing in the rows of λ.

Proof. Let us first notice that if we re-label $T(D')$ by exchanging the integer i with $r-i+1$ then the theorem states that the resulting labeled diagram is a column strict tableau. So for the sake of the proof let $T'(D')$ be the labeled diagram defined with the labeling $r-i+1$ in row i.

Let $D' \in \Omega'(w)$. For every position $(i,j) \in [1 \dots r] \times [1 \dots \lambda_1]$ let $\alpha_{i,j}(D')$ denote the number of elements of D' in column j and in rows $i, i+1, \dots, r$. It is clear that $\alpha_{i,j}(D') \geqslant \alpha_{i+1,j}(D')$ for all (i,j). Now if $\alpha_{i,j}(D') \geqslant \alpha_{i,j+1}(D')$ for all (i,j) then we say that D' is *α-decreasing*. Using a result in [1] we have

LEMMA 2.2. *$T'(D')$ is column strict if and only if D' is α-decreasing.*

Let D' be α-decreasing. perform on D' a B-move and denote the resulting diagram by D''. We claim that D'' is also α-decreasing. To see this, suppose first that $(e,f) \in D'$ is B-moved to $(e+1,f) \in D''$. We have $\alpha_{i,j}(D'') = \alpha_{i,j}(D')$ for all $(i,j) \neq (e+1,f)$ and $\alpha_{e+1,f}(D'') = \alpha_{e+1,f}(D') + 1 = \alpha_{e,f}(D')$. The only inequality we have to check is "$\alpha_{e+1,f-1}(D'') \geqslant \alpha_{e+1,f}(D'')$." But,

$$\alpha_{e+1,f-1}(D'') = \alpha_{e+1,f-1}(D') \geqslant \alpha_{e,f-1}(D') \geqslant \alpha_{e,f}(D') = \alpha_{e+1,f}(D'').$$

Next suppose instead that $(e,f) \in D'$ is B-moved to $(e-1,f) \in D''$. Here $\alpha_{i,j}(D'') = \alpha_{i,j}(D')$ for all $(i,j) \neq (e,f)$ and $\alpha_{e,f}(D'') = \alpha_{e,f}(D') - 1 = \alpha_{e+1,f}(D')$. The only inequality we have to check is "$\alpha_{e,f}(D'') \geqslant \alpha_{e,f+1}(D'')$." We have three cases to check: First if $(e,f+1) \notin D'$ then

$$\alpha_{e,f}(D'') = \alpha_{e+1,f}(D') \geqslant \alpha_{e+1,f+1}(D') = \alpha_{e,f+1}(D') = \alpha_{e,f+1}(D'').$$

Second, if $(e,f+1) \in D'$ and $(e-1,f+1) \in D'$ then

$$\begin{aligned}\alpha_{e,f}(D'') + 1 = \alpha_{e-1,f}(D'') &= \alpha_{e-1,f}(D') \\ &\geqslant \alpha_{e-1,f+1}(D') = \alpha_{e,f+1}(D') + 1 = \alpha_{e,f+1}(D'') + 1.\end{aligned}$$

Finally, the case $(e,f+1) \in D'$ and $(e-1,f+1) \notin D'$ does not occur by definition of the B-move.

From this claim we deduce that any diagram in $\Omega'(w)$ is α-decreasing, since $D'(w)$ is α-decreasing. Hence by Lemma 2.2 we have that $T'(D')$ is column strict for any $D' \in \Omega'(w)$. Hence the theorem follows. ∎

The converse of Theorem 2.1 is also true if we restrict ourselves to a tableau filled with integers between 1 and r. To see this we start with such a tableau t of shape λ, decreasing in the rows and strictly decreasing in the columns. Denote by $D(t)$ the diagram included in $[1 \dots r] \times [1 \dots \lambda_1]$ obtained by reversing the above process (i.e., the diagram such that $T(D(t)) = t$). The diagram $D(t)$ is in the set $\Omega'(w)$, since we can successively B-move the elements of $D'(w)$ from right to left and from top to bottom to produce $D(t)$. All B-moves are legal, since t is weakly decreasing in the rows and strictly decreasing in the columns. This shows

COROLLARY 2.3. *For w grassmannian of shape λ and descent at r,*

$$S_\lambda(x_r, x_{r-1}, \dots, x_1) = \mathscr{S}_w. \tag{2.4}$$

The relation (2.3) (or P.7)) follows from the fact that $S_\lambda(x_1, x_2, \dots, x_r)$ is a symmetric function.

Remark 2.5. For w a permutation of S_n one may define the notion of the *flag* of w [16, 17]: $\phi(w) = (\phi_1, \phi_2, \dots, \phi_k)$, where $\phi_1 \leqslant \phi_2 \leqslant \cdots \leqslant \phi_k$. More precisely, let $d_i = \min\{j-1 : j > i \text{ and } w_j < w_i\}$. Then $\phi(w)$ is the sequence of integers $d_1, d_2, \dots, d_k$ arranged in increasing order. In [17] it is shown that for w a vexillary permutation,

$$\mathscr{S}_w = \sum_t x^t, \tag{2.5}$$

where the t run over all column strict tableaux of shape $\lambda(w)$ such that no integer i is in a row greater than ϕ_i. The right-hand side of (2.5) may be taken as the definition of multi-Schur functions. A proof of (2.5) in the spirit of Theorem 2.1 may be of some interest. The work of A. Kohnert [4] is precisely in this vein.

More generally can one find a simple notion of a "tableau of shape $D(w)$" that constructs the Schubert polynomials? The diagrams in $\Omega(w)$ are certainly a good start!

3. KOHNERT'S CONSTRUCTION

Let D be any diagram. Choose $(i, j) \in D$ such that $(i, j') \notin D$ for all $j' > j$. Let us suppose that there is a point $(i', j) \notin D$ with $i' < i$. Then let $h < i$ be the largest integer such that $(h, j) \notin D$ and let D_1 denote the diagram obtained from D by replacing (i, j) by (h, j). We say that D_1 is obtained from D by a "*K-move*." Now let $K(D(w))$ denote the set of all diagrams (including D itself) obtainable from D by any sequence of K-moves. Kohnert's conjecture states that for any permutation w we have

$$\mathscr{S}_w = \sum_{D \in K(D(w))} x^D. \tag{3.1}$$

A. Kohnert [4] has proved (3.1) for the case where w is a vexillary permutation using techniques similar to Section 2, but the general case was still open. For the interested reader, here is a sketch of how one may try to prove (3.1).

We have noticed by computer that $\Omega(w) = K(D(w))$. The idea then is to show both inclusions by induction. The inclusion $K(D(w)) \subset \Omega(w)$ is the easiest one. We only have to show that any K-move of an element (i, j) to (h, j) can be simulated using B-moves. For this we proceed by induction on $i-h$. If $i-h=1$ then the K-move is simply one B-move. Now if $i-h>1$, we first perform the sequence of B-moves in row h, $h+1$ necessary to B-move the element $(h+1, j)$ to (h, j). Then using the induction hypothesis we can K-move (i, j) to $(h+1, j)$. Finally, we reverse the first sequence of B-moves in rows h, $h+1$. That shows $K(D(w)) \subset \Omega(w)$.

The other inclusion needs a lot more work. For $D \in K(D(w))$ and i any row of D let $B_i(D)$ denote the set of all diagrams (including D) obtainable from D by any sequence of B-moves in the rows i, $i+1$ only. It is clear that if i is big enough then $B_i(D) \subset K(D(w))$. We may then proceed by reverse induction on i. Now for a fixed i, notice that $B_i(D(w))$ is obtainable from $D(w)$ using only K-moves. Let Ω_0 denote the set of all diagrams obtainable from $B_i(D(w))$ by any sequence of K-moves for which no element crosses the border between the rows $i+1$ and $i+2$. A simple inductive algorithm may be used here to show that for any $D \in \Omega_0$ we have $B_i(D) \subset \Omega_0$. Next let Ω_k denote the set of all diagrams of $K(D(w))$ which have k more elements than $D(w)$ in the rows $1, 2, \ldots, i+1$. For almost all the cases it is fairly easy to show (using induction on k and the induction hypothesis on i) that for $D \in \Omega_k$ we have $B_i(D) \subset \Omega_k$. But some of the cases are really hard to formalize! Now this completed would show that $\Omega(w) \subset K(D(w))$, since $K(D(w)) = \bigcup \Omega_k$.

Acknowledgments

The author thanks I. G. Macdonald for introducing the problem and for his invaluable suggestions. He would also like to thank M. Shimozono and S. Biley for all the fruitfull discussions which generate much off the contents of this paper.

References

1. N. Bergeron and A. M. Garsia, Sergeev's formula and the Littlewood–Richardson rule, *Linear and Multilinear Algebra* **27** (1990).
2. I. N. Bernstein, I. M. Gelfand, and S. I. Gelfand, Schubert celles and cohomology of the spaces G/P, *Russian Math. Surveys* **28** (1973), 1–26.
3. M. Demazure, Désingularisation des variétés de Schubert généralisées, *Ann. Sci. École Norm. Sup. (4)* **7** (1974), 53–58.
4. A. Kohnert, Weintrauben, polynome, tableaux, *Bayreuther Math. Schriften* **38** (1990), 1–97.
5. A. Lascoux, Puissances extérieures, déterminants et cycles de Schubert, *Bull. Soc. Math. France* **102** (1974), 161–179.
6. A. Lascoux, Classes de Chern des variétés de drapeaux, *C.R. Acad. Sci. Paris* **295** (1982), 393–398.
7. A. Lascoux, Anneau de Grothendieck de la variété des drapeaux, preprint, 1988.
8. A. Lascoux and M.-P. Schützenberger, Polynômes de Schubert, *C.R. Acad. Sci. Paris* **294** (1982), 447–450.
9. A. Lascoux and M.-P. Schützenberger, Structure de Hopf de l'anneau de cohomologie et de l'anneau de Grothendieck d'une variété de drapeaux, *C.R. Acad. Sci. Paris* **295** (1982), 629–633.
10. A. Lascoux and M.-P. Schützenberger, Symmetry and flag manifolds, *in* "Lecture Notes in Math.," Vol. 996, pp. 118–144, Springer-Verlag, New York/Berlin, 1983.
11. A. Lascoux and M.-P. Schützenberger, Schubert polynomials and the Littlewood–Richardson rule, *Lett. Math. Phys.* **10** (1985), 111–124.
12. A. Lascoux and M.-P. Schützenberger, Interpolation de Newton à plusieurs variables, *in* "Sém. d'Algèbre M. P. Malliavin 1983–84," Lecture Notes in Math., Vol. 1146, pp. 161–175, Springer-Verlag, New York/Berlin, 1985.
13. A. Lascoux and M.-P. Schützenberger, Symmetrization operators on polynomial rings, *Funkt. Anal.* **21** (1987), 77–78.
14. A. Lascoux and M.-P. Schützenberger, Schubert and Grothendieck polynomials, preprint, 1988.
15. I. G. Macdonald, "Symmetric Functions and Hall Polynomials," Oxford Univ. Press, Oxford, 1979.
16. I. G. Macdonald, Notes on Schubert Polynomials, Publication du LACIM, Université du Québec à Montréal, No. 6, 1991.
17. M. L. Wachs, Flagged Schur functions, Schubert polynomials and symmetrizing operators, *J. Combin. Theory Ser. A* **40** (1985), 276–289.