Journal of Algebraic Combinatorics 2 (1993), 345–374

Some Combinatorial Properties of Schubert Polynomials

SARA C. BILLEY[1], WILLIAM JOCKUSCH[2], AND RICHARD P. STANLEY[3]
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139

Received November 2, 1992; Revised April 27, 1993

Abstract. Schubert polynomials were introduced by Bernstein et al. and Demazure, and were extensively developed by Lascoux, Schützenberger, Macdonald, and others. We give an explicit combinatorial interpretation of the Schubert polynomial $\mathfrak{S}_w$ in terms of the reduced decompositions of the permutation w. Using this result, a variation of Schensted's correspondence due to Edelman and Greene allows one to associate in a natural way a certain set $\mathcal{M}_w$ of tableaux with w, each tableau contributing a single term to $\mathfrak{S}_w$. This correspondence leads to many problems and conjectures, whose interrelation is investigated. In Section 2 we consider permutations with no decreasing subsequence of length three (or 321-avoiding permutations). We show for such permutations that $\mathfrak{S}_w$ is a *flag skew Schur function.* In Section 3 we use this result to obtain some interesting properties of the rational function $s_{\lambda/\mu}(1, q, q^2, \ldots)$, where $s_{\lambda/\mu}$ denotes a skew Schur function.

Keywords: divided difference operator, Schubert polynomial, reduced decomposition, Edelman-Greene correspondence, 321-avoiding permutation, flag skew Schur function, principal specialization

1. A combinatorial description of Schubert polynomials

Schubert polynomials were introduced by Bernstein et al. [3] and Demazure [4] (in the context of arbitrary root systems) and were extensively developed by Lascoux and Schützenberger. A treatment of this work, with much additional material, appears in [13] and will be our main reference on Schubert polynomials. We will also use some results from the theory of symmetric functions which can be found in [14] and [21]. Our main result in this section is a combinatorial interpretation of Schubert polynomials completely different from an earlier conjecture of Kohnert [13, (4.20)] and theorem of N. Bergeron [13 (B.1), p. 66]. A different proof of this result appears in [7]. Moreover, our result can be deduced from work of Lascoux-Schützenberger, as shown in [16, after Theorem 2]. In Section 2 we consider Schubert polynomials $\mathfrak{S}_w$ when w has no decreasing subsequence of

[1]Supported by the National Physical Science Consortium. Current address: Department of Mathematics, UCSD, La Jolla, CA 92093.

[2]Supported by an NSF Graduate Fellowship. Current address: Department of Mathematics, University of Michigan, Ann Arbor, MI 48109.

[3]Partially supported by NSF grants DMS-8901834 and DMS-9206374.

length three. Such Schubert polynomials have a number of interesting special properties; for instance, they are skew flag Schur (or multi-Schur) functions. In Section 3 we use our results on permutations with no decreasing subsequence of length three to obtain some new combinatorial properties of the rational function $s_{\lambda/\mu}(1, q, q^2, \ldots)$, where $s_{\lambda/\mu}$ denotes a skew Schur function.

We wish to thank Nantel Bergeron for helpful conversation, and Vic Reiner for carefully reading the proof of Theorem 1.1.

Let us begin with the definition of the Schubert polynomial $\mathfrak{S}_w = \mathfrak{S}_w(x) = \mathfrak{S}_w(x_1, x_2, \ldots, x_{n-1})$, where w is a permutation in the symmetric group $\mathcal{S}_n$. If f is a function of x and y (and possibly other variables), define the *divided difference* operator ∂_{xy} by

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

We also write $\partial_r = \partial_{x_r x_{r+1}}$. Let s_i for $1 \leq i \leq n-1$ denote the adjacent transposition $s_i = (i, i+1) \in \mathcal{S}_n$. For $w = w_1 w_2 \cdots w_n \in \mathcal{S}_n$ (where $w_i = w(i)$), write

$$l(w) = \#\{(i, j) : i < j,\ w_i > w_j\},$$

the *length* or *number of inversions* of w. A *reduced decomposition* of w is a sequence $(a_1, \ldots, a_p)$, $1 \leq a_i \leq n-1$, such that

$$w = s_{a_1} s_{a_2} \cdots s_{a_p},$$

where $p = l(w)$. (Permutations are multiplied left-to-right, so $s_1 s_2 = 231$, not 312.) Let w_0 denote the permutation $n, n-1, \ldots, 1$, i.e., the unique permutation in $\mathcal{S}_n$ of maximum possible length $\binom{n}{2}$. Let $R(w)$ denote the set of reduced decompositions of w, and suppose $(a_1, \ldots, a_p) \in R(w)$. Define the operator $\partial_w = \partial_{a_1} \cdots \partial_{a_p}$. If $a \in R(w)$, then we also say that a is a *reduced word.* It can be shown [13, (2.5)] that ∂_w is independent of the choice of reduced decomposition $(a_1, \ldots, a_p)$ and hence is well defined. Finally define

$$\mathfrak{S}_w = \partial_{w^{-1} w_0}(x_1^{n-1} x_2^{n-2} \cdots x_{n-1}).$$

Thus $\mathfrak{S}_w$ is a homogeneous polynomial of degree $l(w)$ in the variables $x_1, \ldots, x_{n-1}$. In particular, $\mathfrak{S}_{w_0} = x_1^{n-1} x_2^{n-2} \cdots x_{n-1}$.

It is a by no means obvious fact [13, (4.17)] that the coefficients of $\mathfrak{S}_w$ are *nonnegative*. A basic problem in the theory of Schubert polynomials is to give a combinatorial interpretation of these coefficients. The conjecture of Kohnert mentioned above gives a simple, albeit algorithmic, solution to this problem, while the result of Bergeron gives a more complicated algorithmic solution. We wish to give a nonalgorithmic solution to this problem.

If $a = (a_1, \ldots, a_p) \in R(w)$, then define a sequence $(i_1, \ldots, i_p) \in \mathbb{P}^p$ (where $\mathbb{P} = \{1, 2, \ldots\}$) to be *a-compatible* if

$$i_1 \leq i_2 \leq \cdots \leq i_p$$
$$i_j \leq a_j \text{ for } 1 \leq j \leq p$$
$$i_j < i_{j+1} \text{ if } a_j < a_{j+1}.$$

Let $K(a)$ denote the (finite) set of all a-compatible sequences. Also let

$$RK(w) = \{(a, i) : a \in R(w),\ i \in K(a)\}. \tag{1}$$

THEOREM 1.1. *Let* $w \in S_n$. *Then*

$$\mathfrak{S}_w = \sum_{a \in R(w)} \sum_{(i_1, \ldots, i_p) \in K(a)} x_{i_1} x_{i_2} \cdots x_{i_p}.$$

Example 1.1. Let $w = 2431 \in S_4$. The pairs $(a, i) \in RK(w)$ are (1323, 1223) and (3123, 1123). Hence $\mathfrak{S}_{2431} = x_1 x_2^2 x_3 + x_1^2 x_2 x_3$. There is one further element of $R(w)$, viz., $a = 1232$, but $K(1232) = \emptyset$.

We let $\mathfrak{R}_w$ denote the polynomial on the right-hand side of Theorem 1.1. We will prove Theorem 1.1 later in this section. First we will prove two additional formulas for these polynomials. In the proof, we will use lower-case letters to denote permutations and letters with tildes over them to denote reduced words for the corresponding permutation (e.g., $\tilde{w}$ denotes a reduced decomposition of w).

We define a **raising operator** $\uparrow$ on the space of polynomials in $x_1, x_2, \ldots$ by

$$\uparrow(x_i) = x_{i+1},$$

$\uparrow(XY) = \uparrow(X) \uparrow(Y)$, and $\uparrow(X + Y) = \uparrow(X) + \uparrow(Y)$ for polynomials X and Y. We write $\uparrow^j$ for $\uparrow$ iterated j times. Note that we have the identity $\uparrow(\partial_i(X)) = \partial_{i+1}(\uparrow(X))$. We say that a reduced word $\tilde{v} \in R(v)$ of length k is an **initial word** of a permutation w if there is a reduced word $\tilde{w}$ for w whose first k entries are the same as $\tilde{v}$, or, equivalently, if $l(w) = l(v) + l(v^{-1}w)$. Note that the permutation v is defined by $\tilde{v}$. We say that a word $\tilde{w}$ is **descending** if $\tilde{w}_{i+1} < \tilde{w}_i$ for all i. If w is a permutation with $w(1) = 1$, we write $\downarrow(w)$ for the permutation $w_2 - 1, \ldots, w_n - 1 \in S_{n-1}$. Similarly, if $\tilde{w} = (\tilde{w}_1, \ldots, \tilde{w}_k)$ is a reduced word, we use $\uparrow(\tilde{w})$ to denote the word $(\tilde{w}_1 + 1, \ldots, \tilde{w}_k + 1)$, and if $\tilde{w}_i > 1$ for all i, we use $\downarrow(\tilde{w})$ to denote the word $(\tilde{w}_1 - 1, \ldots, \tilde{w}_k - 1)$. We now prove the following recursion formula for the $\mathfrak{R}_w$:

$$\mathfrak{R}_w = \sum_{\substack{\text{descending initial words } \tilde{v} \in R(v) \text{ for } w \\ \text{with } v(1) = w(1)}} x_1^{|\tilde{v}|} \uparrow(\mathfrak{R}_{\downarrow(v^{-1}w)}) \tag{2}$$

Proof. Given $(\tilde{w}, i) \in RK(w)$ let k be the largest integer for which $i_k = 1$. (If $i_1 > 1$, let $k = 0$.) Let $\tilde{v}(\tilde{w}, i)$ consist of the first k entries of $\tilde{w}$, and let $\tilde{w}'(\tilde{w}, i)$ consist of the remaining entries of $\tilde{w}$. Finally, let $i'(\tilde{w}, i)$ consist of those entries of i which do not equal 1. Then the map

$$(\tilde{w}, i) \leftrightarrow (\tilde{v}(\tilde{w}, i), (\tilde{w}'(\tilde{w}, i), i'(\tilde{w}, i)))$$

is a bijection between $RK(w)$ and the set of ordered pairs $(\tilde{v}, (\tilde{w}', i'))$, where $\tilde{v} \in R(v)$ is a descending initial word for w, $v(1) = w(1)$, $(\tilde{w}', i') \in RK(v^{-1}w)$, and $i'_1 \geq 2$. Then the map $\phi : (\tilde{w}', i') \leftrightarrow (\downarrow(\tilde{w}'), \downarrow(i'))$ is a bijection between the set of $(\tilde{w}', i') \in RK(v^{-1}w)$ with $i'_1 \geq 2$ and $RK(\downarrow(v^{-1}w))$. Combining these, we have a bijection $\Phi : (\tilde{w}, i) \leftrightarrow (\tilde{v}, (\downarrow(\tilde{w}'), \downarrow(i')))$ between $RK(w)$ and the set of ordered triples $(\tilde{v}, \tilde{w}'', i'')$, where $\tilde{v} \in R(v)$ is a descending initial word for w with $v(1) = w(1)$ and $(\tilde{w}'', i'') \in RK(\downarrow(v^{-1}w))$. The recursion formula (2) is now established by checking that the term of $\mathfrak{R}_w$ corresponding to any $(\tilde{w}, i) \in RK(w)$ is the same as the contribution of $\Phi(\tilde{w}, i)$ to the right-hand side. □

Remark. Mark Shimozono points out that formula (7.6) in [13] is (when expanded) identical to the above recursion formula, except that $\mathfrak{R}_w$ is replaced by $\mathfrak{S}_w$. This gives an inductive proof of Theorem 1.1. However, as Macdonald's proof of formula (7.6) is rather long, we give our own proof of Theorem 1.1 below.

Our next goal is to give a decomposition formula for the Schubert polynomials and the polynomials $\mathfrak{R}_w$. This will be useful for doing induction in the main proof.

Given permutations $v = v_1, \ldots, v_j \in \mathcal{S}_j$ and $w = w_1, \ldots, w_k \in \mathcal{S}_k$, we let $v * w$ and $v \times w$ denote the permutations $v_1 + k, \ldots, v_j + k, w_1, \ldots, w_k \in \mathcal{S}_{j+k}$ and $v_1, \ldots, v_j, w_1 + j, \ldots, w_k + j \in \mathcal{S}_{j+k}$, respectively. We then have the following result.

Block decomposition formula:

$$\mathfrak{S}_{v*w} = (x_1 \cdots x_j)^k \mathfrak{S}_v \uparrow^j \mathfrak{S}_w.$$

Proof. Let L_m denote the longest permutation in $\mathcal{S}_m$. Let $\tilde{v}$ be a reduced word for $v^{-1}L_j$, and let $\tilde{w}$ be a reduced word for $w^{-1}L_k$. Then $\tilde{v} \uparrow^j \tilde{w}$ is a reduced word for $(v*w)^{-1}L_{j+k}$. Hence

$$\begin{aligned}\mathfrak{S}_{v*w} &= \partial_{\tilde{v}\uparrow(\tilde{w})} x_1^{j+k-1} \cdots x_{j+k-1} \\ &= \partial_{\tilde{v}} x_1^{j+k-1} \cdots x_j^k \partial_{\uparrow^j(\tilde{w})} x_{j+1}^{k-1} \cdots x_{j+k-1} \\ &= (x_1 \cdots x_j)^k \mathfrak{S}_v \uparrow^j \mathfrak{S}_w\end{aligned}$$

□

Remark. The block decomposition can be proven in more generality by referring to N. Bergeron's algorithmic method of finding Schubert polynomials [2]. If the diagram [13] of a permutation in $\mathcal{S}_n$ has block form, where the lower right $k \times j$ block is empty and the upper left $(n-k) \times (n-j)$ block is filled with balls (for some j, k with $j + k \geq n$), then the decomposition holds. In this case, $v \in \mathcal{S}_j$ is the permutation whose diagram has the same configuration as the upper right

$(n-k) \times j$ block and $w \in \mathcal{S}_k$ is the permutation whose diagram is the lower left $k \times (n-j)$ block.

In general the same block decomposition holds for the $\mathfrak{N}_w$ as well. We prove this only for the special cases $w = u * 1$. The following formula is the same as the previous one, except that it concerns the $\mathfrak{N}_w$ instead of the $\mathfrak{S}_w$, and we have set $v = u$, $j = n$ and $k = 1$.

Block decomposition formula:

$$\mathfrak{N}_{u*1} = (x_1 \cdots x_n)\mathfrak{N}_u$$

Proof. We define a bijective map $\Theta : RK(u) \to RK(u*1)$ which will induce a bijective map of monomials from $\mathfrak{N}_u$ into $\mathfrak{N}_{u*1}$ defined by multiplying each term in $\mathfrak{N}_u$ by $x_1 \cdots x_n$.

We can write $u*1$ as the product of two permutations, namely,

$$u * 1 = (1 \times u)(2, 3, \ldots, n, n+1, 1)$$

Note that

$$l(u*1) = l(u) + n = l(u) + l(2, 3, \ldots, n+1, 1).$$

Therefore for any $\widetilde{u} \in R(u)$, we have $\uparrow(\widetilde{u}) \cdot s_1 \cdots s_n \in R(u*1)$. □

Define Θ on $(\widetilde{u}, i) \in RK(u)$ as follows. For each $1 \le k \le n$, let $i_{k_1}, \ldots, i_{k_m}$ be all of the elements in i equal to k, and set $\widetilde{\mu}_k = (\widetilde{u}_{k_1} \cdots \widetilde{u}_{k_m})$. Each element of $\uparrow(\widetilde{\mu}_k)$ must be larger than k, so $1, \ldots, k-1$ commute with $\uparrow(\widetilde{\mu}_k)$. Therefore $\widetilde{w} = \uparrow(\widetilde{\mu}_1) \cdot 1 \cdot \uparrow(\widetilde{\mu}_2) \cdot 2 \cdots \uparrow(\widetilde{\mu}_n) \cdot n \in R(u*1)$. It is easy to see $\widetilde{w}$ has the compatible sequence $j = (1^{l(\widetilde{\mu}_1)}, 1, 2^{l(\widetilde{\mu}_2)}, 2, \ldots, n^{l(\widetilde{\mu}_n)}, n)$ (i.e., the sequence with $\widetilde{\mu}_1 + 1$ entries labeled 1, $\widetilde{\mu}_2 + 1$ entries labeled 2, etc). Hence we can define Θ by

$$\Theta\begin{pmatrix} \widetilde{u} & = & \widetilde{\mu}_1 & \widetilde{\mu}_2 & \cdots & \widetilde{\mu}_n \\ i & = & 1^{l(\widetilde{\mu}_1)} & 2^{l(\widetilde{\mu}_2)} & \cdots & n^{l(\widetilde{\mu}_n)} \end{pmatrix} = \begin{pmatrix} \widetilde{w} & = & \uparrow(\widetilde{\mu}_1) & 1 & \uparrow(\widetilde{\mu}_2) & 2 & \cdots & \uparrow(\widetilde{\mu}_n) & n \\ j & = & 1^{l(\widetilde{\mu}_1)} & 1 & 2^{l(\widetilde{\mu}_2)} & 2 & \cdots & n^{l(\widetilde{\mu}_n)} & n \end{pmatrix}$$

Conversely, for every $\widetilde{w} \in R(u*1)$, $\widetilde{w}$ we must contain the subsequence $1, 2, \ldots, n$. If j is $\widetilde{w}$-compatible then we can write $\widetilde{w}$ as the product of decreasing sequences ending in $1, 2, \ldots, n$, i.e. $\widetilde{w} = \alpha_1, 1, \alpha_2, 2, \ldots, \alpha_n, n$. Now we define the inverse of Θ on $(\widetilde{w}, j) \in RK(u*1)$ by

$$\Theta^{-1}\begin{pmatrix} w & = & \alpha_1 & 1 & \alpha_2 & 2 & \cdots & \alpha_n & n \\ j & = & 1^{l(\alpha_1)} & 1 & 2^{l(\alpha_2)} & 2 & \cdots & n^{l(\alpha_n)} & n \end{pmatrix} = \begin{pmatrix} \widetilde{u} & = & \downarrow(\alpha_1) & \downarrow(\alpha_2) & \cdots & \downarrow(\alpha_n) \\ i & = & 1^{l(\alpha_1)} & 2^{l(\alpha_2)} & \cdots & n^{l(\alpha_n)} \end{pmatrix}$$

Note that every element in α_k must be strictly larger than k; hence α_k commutes with $1, 2, \dots, k-1$. Therefore $(\alpha_1, \dots, \alpha_n)(1, 2, \dots, n) \in R(u*1)$, which implies $\downarrow(\alpha_1, \dots, \alpha_n) \in R(u)$. Furthermore, the sequence $(1^{l(\alpha_1)}, 2^{l(\alpha_2)}, \dots, n^{l(\alpha_n)})$ is compatible with $\downarrow(\alpha_1, \dots, \alpha_n)$ since each $\downarrow(\alpha_k)$ is decreasing and every element is greater than or equal to k. Therefore, we have shown that Θ is a bijective map.

From the definition of Θ, it is easy to see that if $\Theta(\tilde{u}, i) = (\tilde{w}, j)$, then the monomial corresponding to $(\tilde{u}, i)$ in $\mathfrak{R}_u$ will be x^i and the monomial corresponding to $(\tilde{w}, j)$ in $\mathfrak{R}_{u*1}$ will be $x^j = (x_1 x_2 \cdots x_n) \cdot x^i$ □

We now prove Theorem 1.1. We show by induction on n that the theorem holds for $\mathcal{S}_n$ i.e., that $\mathfrak{R}_w = \mathfrak{S}_w$ for any $w \in \mathcal{S}_n$. For $\mathcal{S}_1$, this is clear. For the remainder of the proof, we fix n and assume that the theorem has been proved for $\mathcal{S}_n$. We will show that the theorem is true in $\mathcal{S}_{n+1}$. We need the following lemma.

LEMMA 1.1. *Let $i > 1$; $w \in \mathcal{S}_{n+1}$. Then*

$$\partial_i \mathfrak{R}_w = \begin{cases} \mathfrak{R}_{ws_i} & \text{if } i \text{ is a descent of } w \\ 0 & \text{otherwise.} \end{cases}$$

Remark. The analogous formula for Schubert polynomials holds for all i. A proof can be found in [13, p. 45.]

Proof of Lemma. We have the following identities, which are explained below:

$$\partial_i \mathfrak{R}_w = \partial_i \left[\sum_{\substack{\text{descending initial words } \tilde{v} \in R(v) \text{ for } w \\ \text{with } v(1) = w(1)}} x_1^{|\tilde{v}|} \uparrow(\mathfrak{R}_{\downarrow(v^{-1}w)}) \right] \tag{3}$$

$$= \sum_{\substack{\text{descending initial words } \tilde{v} \in R(v) \text{ for } w \\ \text{with } v(1) = w(1)}} x_1^{|\tilde{v}|} \uparrow(\partial_{i-1} \mathfrak{R}_{\downarrow(v^{-1}w)}) \tag{4}$$

$$= \begin{cases} \displaystyle\sum_{\substack{\text{descending initial words} \\ \tilde{v} \in R(v) \text{ for } ws_i \\ \text{with } v(1) = ws_i(1)}} x_1^{|\tilde{v}|} \uparrow(\mathfrak{R}_{\downarrow(v^{-1}ws_i)}) & \text{if } i \text{ is a descent of } w \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

$$= \begin{cases} \mathfrak{R}_{ws_i} & \text{if } i \text{ is a descent of } w \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Equations (3) and (6) are both applications of the recursion formula for the $\mathfrak{R}_w$. Equation (4) follows from the formula $\partial_i \uparrow = \uparrow \partial_{i-1}$. We now prove (5).

We consider two cases. First, if i is an ascent of w, then i is also an ascent of $v^{-1}w$, so $i-1$ is an ascent of $\downarrow(v^{-1}w)$, which implies that $\partial_{i-1}\mathfrak{R}_{\downarrow(v^{-1}w)} = 0$; that is, every summand is zero, and in this case (5) follows.

On the other hand, suppose that i is a descent of w. Since $i > 1$, we know that $w(1) = ws_i(1)$, so the conditions $v(1) = w(1)$ and $v(1) = ws_i(1)$ are interchangeable. Recall that for a permutation $u \in \mathcal{S}_n$,

$$\partial_i(\mathfrak{S}_u) = \begin{cases} \mathfrak{S}_{us_i} & \text{if } i \text{ is a descent of } u \\ 0 & \text{otherwise.} \end{cases}$$

Since we are assuming that Theorem 1.1 has been proved for $\mathcal{S}_n$, we can freely interchange $\mathfrak{S}_w$ and $\mathfrak{R}_w$ for $w \in \mathcal{S}_n$. We also have the identity $\partial_{i-1}\mathfrak{S}_{\downarrow u} = \mathfrak{S}_{\downarrow(us_i)}$. Therefore,

$$\partial_{i-1}(\mathfrak{R}_{\downarrow(v^{-1}w)}) = \begin{cases} \mathfrak{R}_{\downarrow(v^{-1}ws_i)} & \text{if } i-1 \text{ is a descent of } \downarrow(v^{-1}w) \\ 0 & \text{otherwise.} \end{cases}$$

Every $\tilde{v}$ appearing in the right-hand side of (5) also appears on the right-hand side of (4) since $a \in ws_i$ implies $a \cdot s_i \in R(w)$. Hence, to complete the proof we need to show that for any initial word $\tilde{v}$ of w appearing in the index set on the right-hand side of (4), $i-1$ is a descent of $\downarrow(v^{-1}w)$ if and only if $\tilde{v}$ appears in the index set on the right-hand side of (5). Let $\tilde{v} \in R(v)$ be a descending initial word for w with v(1) = w(1). First, let's assume that $i-1$ is a descent of $\downarrow(v^{-1}w)$. Multiplying by s_i on the right transposes the elements in positions i and $i+1$ so i is a descent of $v^{-1}w$ implies i is an ascent of $v^{-1}ws_i$. Hence $l(v) + l(v^{-1}ws_i) = l(v) + l(v^{-1}w) - 1 = l(w) - 1 = l(ws_i)$, so $\tilde{v}$ is a descending initial word for ws_i and appears in the index set on the right-hand side of (5). On the other hand, if $i-1$ is an ascent of $\downarrow(v^{-1}w)$, similar logic shows that $l(v) + l(v^{-1}ws_i) = l(w) + 1$, so $\tilde{v}$ does not appear in the index set on the right-hand side of (5). This completes the proof of (5) and of the lemma. □

We now show that Theorem 1.1 holds for two classes of permutations in $\mathcal{S}_{n+1}$; as the union of the two classes is all of $\mathcal{S}_{n+1}$, this will complete the proof.

PROPOSITION 1.1. *If* $w \in \mathcal{S}_{n+1}$ *and* $w(1) \neq 1$, *then* $\mathfrak{R}_w = \mathfrak{S}_w$.

Proof. First, we prove the proposition in the special case $w(n+1) = 1$. We then have $w = u * 1$ for some permutation u.

We now have

$$\begin{aligned}\mathfrak{R}_w &= x_1x_2\cdots x_n\mathfrak{R}_u \\ &= x_1x_2\cdots x_n\mathfrak{S}_u \\ &= \mathfrak{S}_w.\end{aligned}$$

The first of the above equations follows from the block formula for $\mathfrak{R}_w$; the second because Theorem 1.1 is assumed to be true for $\mathcal{S}_n$; and the third by the block decomposition for Schubert polynomials. This completes the proof of the special case. But now the proposition follows from the special case, the lemma, and the remark following the lemma, i.e., we apply a sequence of ∂_is which "move" the 1 to its position. □

PROPOSITION 1.2. *If* $w \in \mathcal{S}_{n+1}$ *and* $w(1) < w(2)$, *then* $\mathfrak{S}_w = \mathfrak{R}_w$.

Proof. We proceed by induction on $l(w)$. If $l(w) = 0$, then w is the identity permutation and $\mathfrak{R}_w = \mathfrak{S}_w = 1$. Now, let $w \in \mathcal{S}_{n+1}$ and suppose the proposition has been proven for every permutation in $\mathcal{S}_{n+1}$ of shorter length than w. If $1 < i \leq n$, then $\partial_i(\mathfrak{R}_w - \mathfrak{S}_w) = 0$ by the lemma, remark, and induction hypothesis. Hence $\mathfrak{R}_w - \mathfrak{S}_w$ is symmetric in the variables $x_2, \ldots, x_{n+1}$, which implies that it cannot depend on any of them since neither $\mathfrak{R}_w$ nor $\mathfrak{S}_w$ can depend on x_{n+1}. Hence all that remains to be shown is that $\mathfrak{R}_w$ and $\mathfrak{S}_w$ have the same coefficient of $x_1^{l(\overline{w})}$. But since $w(1) < w(2)$, $\mathfrak{S}_w$ is symmetric in x_1 and x_2. Therefore,

$$\mathfrak{S}_w|_{x_1^{l(w)}} = \mathfrak{S}_w|_{x_2^{l(w)}}.$$

From the lemma, we have $\partial_2(\mathfrak{R}_w - \mathfrak{S}_w) = 0$, so

$$\mathfrak{S}_w|_{x_2^{l(w)}} = \mathfrak{R}_w|_{x_2^{l(w)}}.$$

Note that $\mathfrak{R}_w|_{x_i^{l(w)}} = 0$ or 1 since there can be at most one decreasing reduced word for w. If $\mathfrak{R}_w|_{x_2^{l(w)}} = 1$ then there exists a decreasing word $a = (a_{l(w)}, \ldots, a_1) \in R(w)$ such that $a_1 \geq 2$. Hence $(1, 1, \ldots, 1)$ is compatible with a also. So

$$\mathfrak{S}_w|_{x_1^{l(w)}} = 1 = \mathfrak{R}_w|_{x_1^{l(w)}}.$$

If $\mathfrak{R}_w|_{x_2^{l(w)}} = 0$ then no decreasing reduced word for w exists that ends in 2 or higher. There cannot exist a decreasing reduced word ending in 1 when $w(1) < w(2)$; therefore $\mathfrak{R}_w|_{x_1^{l(w)}} = 0$ also. This completes the proof of the proposition and of Theorem 1.1. □

We now consider an interesting consequence of Theorem 1.1. Given a reduced decomposition $a = (a_1, \ldots, a_p)$, Edelman and Greene [5] describe a variation of the Robinson-Schensted correspondence which associates with a, a tableau T_a of some shape $\lambda \vdash p$. (An equivalent but less algorithmic description of T_a appears in [12].) If $w \in \mathcal{S}_n$, then the entries of T_a lie in $\{1, 2, \ldots, n-1\}$ and are strictly increasing in each row and column.

Briefly, the Edelman-Greene correspondence starts with the empty tableau, and builds up $T_{(a_1,\ldots,a_p)}$ by inserting a_1, then a_2, and so on, stopping after inserting

a_p. We insert a_i into the current tableau as follows. We start off by inserting a_i into row 1, using the algorithm described below. We then insert the "bumped" entry, if any, into row 2, continuing in this manner until there is no "bumped" entry. We now describe the algorithm for inserting a_i into a row R. If a_i and $a_i + 1$ are in R when we try to insert a_i, we leave R as it is and insert $a_i + 1$ (the "bumped" entry) into the next row. Otherwise, we insert a_i into R by replacing the smallest element larger than a_i, say x_i, with a_i and row inserting x_i (the "bumped" entry) into the next row. If no element in R is larger than a_i we add a_i to the right end of R, and nothing is bumped. An example follows the next definition.

With this definition, we see that if λ^j is the shape of $T_{(a_1, \dots, a_j)}$, then $\emptyset = \lambda^0 \subset \lambda^1 \subset \cdots \subset \lambda^j = \lambda$ with $|\lambda^j| = j$. Thus given $(a, i) = (a, (i_1, \dots, i_p)) \in RK(w)$, we can define an "insertion tableau" $I'(a, i)$ analogous to the usual Robinson-Schensted correspondence, as follows: Start with $I'_0 := I'(\emptyset, \emptyset) = \emptyset$. Once $I'_j := I'((a_1, \dots, a_j), (i_1, \dots, i_j))$ is defined, let I'_{j+1} be obtained by placing i_{j+1} into I'_j so that the resulting tableau has the same shape as $T_{(a_1, \dots, a_{j+1})}$. Finally let $I'(a, i) = I'_p(a, i)$, and let $I(a, i)$ be the transpose of $I'(a, i)$.

Example 1.2. Let $(a, i) = (3167687, 1135567)$. Write $T_j = T_{(a_1, \dots, a_j)}$. The tableaux, T_j and I'_j are given as follows:

T_j	I'_j
3	1
1 3	1 1
1 6 3	1 3 1
1 6 7 3	1 3 5 1

1 6 7   1 3 5
3 7   1 5

1 6 7 8   1 3 5 6
3 7   1 5

1 6 7 8   1 3 5 6
3 7 8   1 5 7

Hence

$$I(a, i) = \begin{matrix} 1 & 1 \\ 3 & 5 \\ 5 & 7 \\ 6 & \end{matrix}$$

It is easy to see from the definition of the Edelman-Greene correspondence that $I(a, i)$ is a column-strict reverse plane partition (or *semistandard tableau*, abbreviated SST), i.e., the rows weakly increase and the columns strictly increase. If a tableau I has m_j parts equal to j, then write $x^I = x_1^{m_1} x_2^{m_2} \cdots$. Thus $x^{I(a,i)} = x_{i_1} \cdots x_{i_p}$, where $i = (i_1, \ldots, i_p)$. We therefore have for each permutation w a multiset $\mathcal{M}_w = \{I(a, i) : (a, i) \in RK(w)\}$ of semistandard tableaux such that

$$\mathfrak{S}_w = \sum_{I \in \mathcal{M}_w} x^I. \tag{7}$$

It is natural to ask about the combinatorial structure of the multiset $\mathcal{M}_w$. In particular, when does it have a simple direct description avoiding the use of the Edelman-Greene correspondence?

To be more precise, we recall the definition of a flag (or flagged) skew Schur function (called a multi-Schur function in [13]). We give the combinatorial definition as in [23, p. 277]. In [23] there is a "left flag" a and "right flag" b. Here we will only have a right flag (which will be denoted by ϕ).

Let $X_1 \subseteq X_2 \subseteq \cdots \subseteq X_n$ be finite sets of variables, say $X_i = \{x_1, x_2, \ldots, x_{\phi_i}\}$, with the variables totally ordered by $x_1 < x_2 < \cdots < x_{\phi_n}$. We call the sequence $\phi = (\phi_1, \ldots, \phi_n)$ *a flag*, and write $\mathcal{X}_\phi = (X_1, \ldots, X_n)$, the *flag alphabet* corresponding to ϕ. Note that $\phi_1 \le \phi_2 \le \cdots \le \phi_n$. Now let $\lambda = (\lambda_1, \cdots, \lambda_n)$, $\mu = (\mu_1, \ldots, \mu_n)$ be partitions with $\mu \subseteq \lambda$ (i.e., $\mu_i \le \lambda_i$ for all i). Let $\mathcal{A} = \mathcal{A}_{\lambda/\mu}(\phi)$ be the set of all semistandard skew tableaux of shape λ/μ such that the elements in row i belong to X_i. Define the *flag skew Schur function*

$$s_{\lambda/\mu}(\mathcal{X}_\phi) = s_{\lambda/\mu}(X_1, \ldots, X_n) = \sum_{I \in \mathcal{A}} x^I. \tag{8}$$

For instance, if $\lambda = (3, 2)$, $\mu = (1, 0)$, $X_1 = \{x_1\}$, $X_2 = \{x_1, x_2\}$, then $\mathcal{A}$ is given by

$$\begin{matrix} & 1\ 1 \\ 1\ 2 & \end{matrix} \qquad \begin{matrix} & 1\ 1 \\ 2\ 2 & \end{matrix}$$

so $s_{\lambda/\mu}(X_1, X_2) = x_1^3 x_2 + x_1^2 x_2^2$. We have the determinantal formula [23, Theorem 3.5] (originally due to Gessel)

$$s_{\lambda/\mu}(X_1, \ldots, X_n) = \det(h_{\lambda_i - \mu_j - i + j}(X_i))_{1 \le i, j \le n}, \tag{9}$$

where $h_k(X_i)$ is the complete symmetric function [14, p. 14] in the variables X_i.

Definition 1.1. Let $f = f(x_1, x_2, \ldots)$ be a symmetric function (in infinitely many variables) and $\phi = (\phi_1, \ldots, \phi_n)$ a flag. Define the *flag symmetric function* $\tilde{f}(\mathcal{X}_\phi)$ by

$$\tilde{f}(\mathcal{X}_\phi) = \sum_\lambda \alpha_\lambda s_\lambda(\mathcal{X}_\phi),$$

where $f = \sum \alpha_\lambda s_\lambda$ is the expansion of f in terms of Schur functions.

Note: We have used the notation $\tilde{f}(\mathcal{X}_\phi)$ instead of $f(\mathcal{X}_\phi)$ to avoid confusion with flag skew Schur functions. In general, it is false that $s_{\lambda/\mu}(\mathcal{X}_\phi) = \tilde{s}_{\lambda/\mu}(\mathcal{X}_\phi)$. For instance, $s_{22/1}(\mathcal{X}_{(1,2)}) = x_1 x_2(x_1 + x_2)$, while $\tilde{s}_{22/1}(\mathcal{X}_{(1,2)}) = x_1^2 x_2$.

We now define two desirable properties that a Schubert polynomial can have. Given a permutation w and an integer $N \ge 0$, define the permutation $1_N \times w = 12 \cdots N(w_1 + N) \cdots (w_n + N)$, and set

$$G_w = \lim_{N \to \infty} \mathfrak{S}_{1_N \times w}.$$

(This limit is well defined, and G_w is a symmetric function in the variables $x_1, x_2, \ldots$ In the notation of [19] or [13, p. 101], we have $G_w = F_{w^{-1}}$.)

Definition 1.2. (a) A permutation w is *patriotic* if there exists a flag ϕ for which $\mathfrak{S}_w = \tilde{G}_w(\mathcal{X}_\phi)$.

(b) Let $G_w = \sum_{\lambda \in M} s_\lambda$, where λ ranges over some multiset $M = M_w$. (There exists such a multiset by [13, (7.18)].) We say that the permutation w is *heroic* if there exists a flag ϕ such that

$$\mathcal{M}_w = \bigcup_{\lambda \in M_w} \mathcal{A}_\lambda(\phi).$$

Note that by (7) a heroic permutation is patriotic. We had conjectured the converse, i.e., every patriotic permutation is heroic. This conjecture was later proved by V. Reiner (private communication):

PROPOSITION 1.3. (Reiner[16]). *Every patriotic permutation is heroic.*

Let us note that if there exists a flag ϕ such that

$$\mathcal{M}_w = \bigcup_{\lambda \in N} \mathcal{A}_\lambda(\phi) \tag{10}$$

for *some* multiset N, then $N = M_w$. This follows from Theorem 1.1 in the limiting case $\lim_{n\to\infty} 1_n \times w$ (which itself is essentially in [13, pp. 101–102]) and the result of Edelman-Greene [5, §8] that the shapes λ of the tableaux T_a, where $a \in R(w)$, are exactly the shapes in M_w. In fact, if λ has multiplicity m in M_w, then exactly mf^λ of the tableaux $\{T_a : a \in r(w)\}$ have shape λ, where f^λ is the number of standard young tableaux of shape λ. Hence, $N \subseteq M_w$. To see that $M_w \subseteq N$, let $I(\lambda)$ be the semistandard tableau of shape λ with only is in the ith row. Using the Edelman-Greene correspondence, we can partition $R(w)$ into Coxeter-Knuth classes, each represented by a row and column strict tableau T. For each such class, the multiset $\{I(a, i) : T_a = T\}$ will contain $I(\lambda(P))$ exactly once corresponding to the column-word of P (i.e., reading the columns of P bottom up, left to right). Therefore the multiset of tableaux $I(\lambda)$ appearing in $\mathcal{M}_w$ is stable under $\lim_{n\to\infty} 1_n \times w$.

Examples of permutations which are not patriotic (so not heroic) are $w = 215364$, $w = 231645$, $w = 21543$, and $w = 254163$. We will see in the next section (Theorem 2.2) that the first two of these have an alternative nice property which leads to a simple combinatorial interpretation of $\mathfrak{S}_w$, while the first three satisfy a more general property (Proposition 2.3), for which a simple combinatorial description still holds.

If $u = u_1u_2\cdots u_m \in \mathcal{S}_m$, then we say that $w = w_1w_2\cdots w_n \in \mathcal{S}_n$ is *u-avoiding* if there does not exist $1 \le i_1 < i_2 < \cdots < i_m \le n$ such that the following condition holds for all $1 \le j < k \le m$: $w_{i_j} < w_{i_k}$ if and only if $u_j < u_k$. In other words, w has no subsequence in the same relative order as u. For instance, 321-avoiding means no decreasing subsequence of length three. A permutation is *vexillary* if and only if it is 2143-avoiding [13, (1.27)]. It follows from [13, (4.9)] that vexillary permutations are patriotic (and hence heroic). There are numerous examples of heroic permutations which are not vexillary, such as 2143.

There is a simple necessary condition for a permutation w to be heroic. Let $a(w) = (a_1, \ldots, a_p)$ be that reduced decomposition of w for which the sequence $a_p \cdots a_1$ is *last* in lexicographic order. In other words, a_p is the largest descent of w, then a_{p-1} is the largest descent of ws_{a_p}, etc. Let $i(w)$ be that $a(w)$-compatible sequence which is last in lexicographic order. (It is easy to see that $K(w(a)) \neq \emptyset$, so $i(w)$ exists.) For instance, if $w = 21543$, then $a(w) = (1, 4, 3, 4)$ and $i(w) = (1, 3, 3, 4)$.

Definition 1.3. A permutation w is *saturated* if the tableau $I = I(a(w), i(w))$ has the following two properties:

(a) if I' is obtained from I by increasing an entry of I which is not the last element of a row, then I' is not semistandard.
(b) Let the last entry of row s be i_s. Suppose $i_r < i_{r-1}$, and I' is obtained from I by changing the last entry i_r of row r to $i_r + 1$. Then I' is not semistandard.

For instance, if $w = 21543$ then I is given by

$$\begin{matrix} 1 & 3 \\ 3 & \\ 4 & \end{matrix}$$

We can increase the 1 to 2 without destroying semistandardness. Hence w is not saturated. Similarly, let $w = 231645$. Then I is given by

$$\begin{matrix} 1 & 4 \\ 2 & \\ 4 & \end{matrix}$$

We have $i_2 = 2 < i_1 = 4$. We can replace i_2 by 3 without destroying semistandardness. Hence w is not saturated.

There is a simple direct description of the tableau $I = I(a(w), i(w))$. Let $c(w) = (c_1(w), \ldots, c_n(w))$ be the code of w [13, p. 9], i.e.,

$$c_i(w) = \#\{j : j > i,\ w(j) < w(i)\}.$$

Then the kth column of I consists of the entries $r_1 < r_2 < \cdots < r_t$, where $r_1, \ldots, r_t$ are such that $c_{r_j} \geq k$ for $1 \leq j \leq t$. For instance, if $w = 4217635$ then $c(w) = (3, 1, 0, 3, 2, 0, 0)$, and one reads off directly that

$$I = \begin{matrix} 1 & 1 & 1 \\ 2 & 4 & 4 \\ 4 & 5 & \\ 5 & & \end{matrix}$$

PROPOSITION 1.4. *Every heroic permutation is saturated.*

Proof. Let $I = I(a(w), i(w))$. One easily sees that x^I is the "leading term" of the Schubert polynomial $\mathfrak{S}_w$, i.e., if we list all monomials appearing in $\mathfrak{S}_w$ in lexicographic order, then x^I comes last. Suppose w is heroic, and that

λ = shape(I). Let i_r be the last entry in row r of I. Since $I \in \mathcal{M}_w$ and w is heroic, we have $\mathcal{A}_\lambda(\phi) \subseteq \mathcal{M}_w$ for some flag $\phi = (\phi_1, \ldots, \phi_m)$ with $\phi_r \geq i_r$.

Now assume that condition (a) of Definition 1.3 fails. Then $I' \in \mathcal{A}_\lambda(\phi)$ and $x^{I'}$ follows x^I in lexicographic order, a contradiction.

Similarly assume that Definition 1.3(b) fails. Since $i_r < i_{r-1}$ and ϕ is weakly increasing, we have $\phi_r \geq i_{r-1} \geq i_r + 1$. Thus again $I' \in \mathcal{A}_\lambda(\phi)$ and $x^{I'}$ follows x^I in lexicographic order, another contradiction. □

We don't know whether the converse to Proposition 1.4 holds, i.e., whether every saturated permutation is heroic.

2. 321-avoiding permutations

In this section we consider permutations w that are 321-avoiding, i.e., for no $i < j < k$ do we have $w(i) > w(j) > w(k)$. In particular, we show that although 321-avoiding permutations need not be patriotic or heroic, their Schubert polynomials satisfy an alternative simple combinatorial property: They are flag *skew* Schur functions.

First let us consider a certain equivalence relation on the set $R(w)$ of all reduced decompositions of a permutation w. It is well known [15, 22] that given any two $a, b \in R(w)$, we can convert a to b by successive applications of the Coxeter relations

$$C_1: \quad s_i s_j = s_j s_i, \; |i - j| \geq 2$$

$$C_2: \quad s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1}.$$

If a can be converted to b only using C_1, then we call a and b *C_1-equivalent* and write $a \sim b$. Clearly $\sim$ is an equivalence relation. For instance, if $w = 4132$ then there are two equivalence classes, viz., {2321} and {3231, 3213}.

THEOREM 2.1. *Let $w \in S_n$. The following four conditions are equivalent.*

(a) w is 321-avoiding.

(b) Any two reduced decompositions of w are C_1-equivalent.

(c) Let $c(w) = (c_1, c_2, \ldots, c_n)$ be the code of w. Suppose $i < j$, $c_i > 0$, $c_j > 0$, and $c_{i+1} = c_{i+2} = \cdots = c_{j-1} = 0$. Then $j - i > c_i - c_j$.

(d) Let $D(w)$ denote the diagram of w, as defined in [13, p. 8]. If (h, i), $(j, k) \in D(w)$ with $h < j$ and $k < i$, then $(j, i) \in D(w)$. (It's always true that $(h, k) \in D(w)$.)

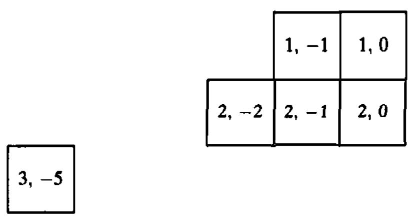
Note: Condition (d) is equivalent to the following: If we reflect $D(w)$ about a vertical line and remove all empty rows and columns, then we obtain the diagram of a *skew partition* or *skew shape* λ/μ. The exact embedding of λ/μ

in $\mathbb{Z} \times \mathbb{Z}$ is not clear from our above description, so we make it more precise as follows. Suppose $c(w) = (c_1, \ldots, c_n)$ and $\{j_1, \ldots, j_\ell\} = \{j : c_j > 0\}$, with $j_1 < \cdots < j_\ell$. Then λ/μ is embedded in $\mathbb{Z} \times \mathbb{Z}$ so that the last element of the kth nonempty row is $(k, k - j_k)$. We then have that

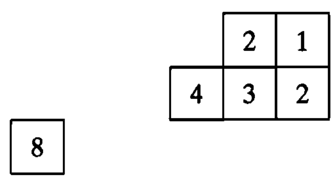
$$\lambda_k - \mu_k = c_{j_k},$$
$$\lambda/\mu = \{(k, h) : 1 \le k \le \ell,\ k - j_k - c_{j_k} + 1 \le h \le k - j_k\}. \tag{11}$$

Denote the skew shape λ/μ of (11) by $\sigma(w)$. If $(k, h) \in \sigma(w)$, then we say that (k, h) lies in the kth row of $\sigma(w)$. We will also need a certain labeling $\omega = \omega(w) : \sigma(w) \to \mathbb{Z}$, defined by $\omega(k, h) = k - h$. (Thus $\omega(k, h) = -\kappa(k, h)$, where κ denotes *content* as in [14, Example 3, p. 10] and [21, Definition 15.1].) It follows that the labels of $\sigma(w)$ decrease by one along rows, increase by one along columns, and that the last (rightmost) label in row k is just j_k.

For instance, let $w = 351246798 \in S_9$. Then $c(w) = (2, 3, 0, 0, 0, 0, 0, 1, 0)$, $(j_1, j_2, j_3) = (1, 2, 8)$, and $\sigma(w)$ (with square (h, k) labeled by h, k) is given by:



The labeling $\omega(w)$ is given by:



Proof of Theorem 2.1. (a) $\Rightarrow$ (b) Assume (b) is false. Then there is a reduced decomposition $a = (a_1, \ldots, a_p) \in R(w)$ such that for some j and t we have $a_{t-1} = j$, $a_t = j+1$, $a_{t+1} = j$. Let $v = s_{a_1} \cdots s_{a_{t+1}}$. Then $v(j) > v(j+1) > v(j+2)$. If we perform a succession of adjacent transpositions each of which increases the length of the permutation, then we can never change the relative order of $v(j)$, $v(j+1)$, $v(j+2)$. Hence in $w = v s_{a_{t+2}} \cdots s_{a_p}$, the elements $v(j)$, $v(j+1)$, $v(j+2)$ form a decreasing subsequence of length three, contradicting (a).

(b) $\Rightarrow$ (a) Assume (a) is false. Thus $w(i) > w(j) > w(k)$ for some $i < j < k$. Regarding j as fixed, we may assume that k is such that $j < k$ and $w(k)$ is minimized, while i is such that $i < j$ and $w(i)$ is maximized. Thus by a sequence of adjacent transpositions, each decreasing length, we can reach a permutation v with $v(j-1) = w(i)$, $v(j) = w(j)$, $v(j+1) = w(k)$. Thus v has a reduced decomposition $(\ldots, j-1, j, j-1)$, and so w has a reduced decomposition $(\ldots, j-1, j, j-1, \ldots)$, contradicting (b).

(a) $\Rightarrow$ (c) Assume (c) is false, and let i and j satisfy $i < j$, $c_i > 0$, $c_j > 0$, $c_{i+1} = c_{i+2} = \cdots = c_{j-1} = 0$, but $j - i \leq c_i - c_j$. Then in the word $w = w(1)w(2)\cdots$, we have that $w(i)$ appears to the left of c_i elements $u_1 < u_2 < \cdots < u_{c_i}$ all less than $w(i)$. Since $c_{i+1} = c_{i+2} = \cdots = c_{j-1} = 0$, we have $w(i+1) = u_1$, $w(i+2) = u_2, \ldots, w(j-1) = u_{j-i-1}$. Since $j - i \leq c_i - c_j$ we have $w(j) = u_{j-i+c_j}$. (Note that $j - i + c_j \leq c_i$ so u_{j-i+c_j} is defined.) Thus u_{j-i} appears to the right of $w(j)$ in w. If $w(k) = u_{j-1}$, then we have $i < j < k$ and $w(i) > w(j) = u_{j-i+c_j} > w(k) = u_{j-i}$, contradicting (a).

(c) $\Rightarrow$ (a) Assume (a) is false. Let i be the largest integer for which there exists $i < j < k$ with $w(i) > w(j) > w(k)$. Given such i, let j be the smallest integer for which such $i < j < k$ exists.

If $i < r < j$ then $w(i) > w(r)$, else $r < j < k$ and $w(r) > w(j) > w(k)$, contradicting the definition of i. If $i < r < s$ and $r < j$, then $w(r) < w(s)$, else $i < r < s$ and $w(i) > w(r) > w(s)$, contradicting the definition of j. These conditions show that $c_{i+1} = c_{i+2} = \cdots = c_{j-1} = 0$, while clearly $c_i > 0$ and $c_j > 0$.

Now since $w(i) > w(r)$ for $i < r \leq j$, there are exactly $c_i - (j - i)$ values of $s > j$ for which $w(i) > w(s)$. There are also c_j values of $t > j$ for which $w(j) > w(t)$. Since $w(i) > w(j)$, every t is also an s. Hence $c_j \leq c_i - (j - i)$, contradicting (c).

(a) $\Rightarrow$ (d) Assume (d) is false. Thus we have (h, i), $(j, k) \in D(w)$ such that $h < j$, $k < i$, and $(j, i) \notin D(w)$ (and always $(h, k) \in D(w)$). Assume that such h, i, j, k are chosen so that $i - k$ is minimized. Since (h, i), $(h, k) \in D(w)$, we have that i and k appear to the right of $w(h)$ in w, and that $w(h) > i$, $w(h) > k$. Since $(j, k) \in D(w)$, we have that k appears to the right of $w(j)$, and that $w(j) > k$.

Case 1: $w(j) > i$. Then i cannot appear to the right of $w(j)$ in w, so i appears to the left of k. Hence $w(h), i, k$ is a decreasing subsequence of w of length three, contradicting (a).

Case 2: $k < w(j) < i$. Then $(h, w(j)) \in D(w)$, since $w(h) > i > w(j)$ and $j = w^{-1}(w(j)) > h$. But then $h, w(j), j, k$ gives a "bad" quadruple with $w(j) - k < i - k$, contradicting the minimality of $i - k$. Hence case 2 cannot occur, so (a) is always false when (d) is false.

(d) $\Rightarrow$ (a) Assume (a) is false. Let $h < j < m$ and $w(h) > w(j) > w(m)$. Let $k = w(m)$ and $i = w(j)$. Then $h < j$, $k < i$, $(j, k) \in D(w)$, $(h, i) \in D(w)$, but $(j, i) \notin D(w)$, contradicting (d) and completing the proof. □

Note: There is a connection between Theorem 2.1 and the *Temperley-Lieb algebra* $A_{\beta,n}$ (over the field K). For the basic properties of this algebra, see [9, p. 33 and §2.8]. For any $\beta \in K^*$, the algebra $A_{\beta,n}$ has generators $\epsilon_1, \ldots, \epsilon_{n-1}$

and relations

$$\epsilon_i^2 = \epsilon_i$$
$$\beta\epsilon_i\epsilon_j\epsilon_i = \epsilon_i, \text{ if } |i-j| = 1$$
$$\epsilon_i\epsilon_j = \epsilon_j\epsilon_i, \text{ if } |i-j| \geq 2.$$

It follows easily that $A_{\beta,n}$ has a K-basis consisting of monomials $u_1, u_2, \ldots$ in the ϵ_i's. Using the above relations we can reduce each monomial u_k to one of minimal length, up to scalar multiplication. Suppose then each u_k is of minimal length, and say $u_k = \epsilon_{a_1}\epsilon_{a_2}\cdots\epsilon_{a_p}$. From Theorem 2.1 it is easily seen that then $(a_1, a_2, \ldots, a_p) \in R(w)$ for some 321-avoiding $w \in \mathcal{S}_n$. Moreover, if V_u is the set of all monomials of length p equal to a scalar multiple of u_k, then $V_u = \{\epsilon_{b_1}\epsilon_{b_2}\cdots\epsilon_{b_p} : (b_1, b_2, \ldots, b_p) \in R(w)\}$. Conversely, for any 321-avoiding w and $(a_1, \ldots, a_p) \in R(w)$, the monomial $\epsilon_{a_1}\cdots\epsilon_{a_p}$ is a scalar multiple of some u_j. Hence if $T_n = \{w \in \mathcal{S}_n : w \text{ is 321-avoiding}\}$, and $a(w) = (a_1(w), \ldots, a_p(w))$ is any fixed reduced decomposition of w, then the set $\{\epsilon_{a_1}(w)\cdots\epsilon_{a_p}(w) : w \in T_n\}$ is a K-basis for $A_{\beta,n}$. Thus $\dim A_{\beta,n}$ is the number of 321-avoiding $w \in \mathcal{S}_n$. It is in fact well known that this number is the Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$. In [9, Proposition 2.8.1] there is a simple combinatorial proof that $\dim A_{\beta,n} = C_n$. This proof can be rewritten in terms of the code of 321-avoiding w and yields a simple direct argument that there are C_n such permutations. (For prior proofs of this result, see [10, pp. 63–64] [17]. The latter reference gives a bijection between 321-avoiding $w \in \mathcal{S}_n$ and 132-avoiding (i.e., dominant) $w \in \mathcal{S}_n$.)

The proof based on [9] goes as follows. Suppose $c(w) = (c_1, \ldots, c_n)$ and $\{j_1, \ldots, j_\ell\} = \{j : c_j > 0\}$, with $j_1 < \cdots < j_\ell$. Define a lattice path from $(0, 0)$ to (n, n) in $\mathbb{R} \times \mathbb{R}$ as follows: Walk horizontally from $(0, 0)$ to $(c_{j_1} + j_1 - 1, 0)$, then vertically to $(s_{j_1} + j_1 - 1, j_1)$, then horizontally to $(c_{j_2} + j_2 - 1, j_1)$, then vertically to $(c_{j_2} + j_2 - 1, j_2)$, etc. The last part of the path is a vertical line from $(c_{j_\ell} + j_\ell - 1, j_{\ell-1})$ to $(c_{j_\ell} + j_\ell - 1, j_\ell)$, then (if needed) a horizontal line to $(c_{j_\ell} + j_\ell - 1, n)$, and finally a vertical line to (n, n). This establishes a bijection between 321-avoiding permutations in $\mathcal{S}_n$ and lattice paths from $(0, 0)$ to (n, n) with steps $(1, 0)$ and $(0, 1)$ which never rise above the diagonal $y = x$. Such paths are standard objects counted by Catalan numbers, as desired.

Let us consider once again the equivalence relation $\sim$ defined earlier on $R(w)$. Suppose $\mathcal{E}$ is an equivalence class, and let $a = (a_1, \ldots, a_p) \in \mathcal{E}$. Define a partial order P_a on the symbols $u_1, \ldots, u_p$ as follows: If $i < j$ and $s_{a_i}s_{a_j} \neq s_{a_j}s_{a_i}$ (i.e., $a_i = a_j \pm 1$), then let $u_i < u_j$; and let P_a be the transitive and reflexive closure of these relations. Moreover, we define $\omega = \omega_a : P_a \to \mathbb{Z}$ by $\omega(u_i) = a_i$. The pair (P_a, ω) may be regarded as a *labeled poset*. For instance, if $a = (3, 2, 1, 3)$ then (P_a, ω) is given in Figure 1.

It is easily seen (and is closely related to work of Cartier-Foata; see [18, Exercise 3.48]) that if also $b \in \mathcal{E}$, then (P_a, ω) and (P_b, ω) are isomorphic as labeled posets, i.e., there is a poset isomorphism $f : P_a \to P_b$ such that $\omega(u) = \omega(f(u))$ for all $u \in P_a$. Moreover, if we associate with a linear extension $\epsilon = (y_1, y_2, \ldots, y_p)$ of

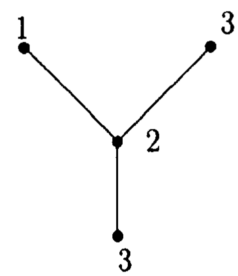


Figure 1. A labeled poset (P_a, W).

P the sequence $\omega(\epsilon) = (\omega(y_1), \ldots, \omega(y_p))$, then we obtain a bijection between the set $\mathcal{L}(P)$ of linear extensions of P and the equivalence class $\mathcal{E}$. For instance, if (P_a, ω) is the labeled poset of Figure 1, then its "labeled linear extensions" $\omega(\epsilon)$ are (3, 2, 1, 3) and (3, 2, 3, 1), which are just the elements of the class $\mathcal{E}$ containing a.

We may therefore write $P_{\mathcal{E}}$ for P_a and $\omega_{\mathcal{E}}$ for ω_a. When w is 321-avoiding, then by Theorem 2.1 there is a single class $\mathcal{E}$, and we can write $P_w = P_{\mathcal{E}}$, $\omega_w = \omega_{\mathcal{E}}$. Then

$$R(w) = \omega\mathcal{L}(P_w) := \{\omega(\epsilon) : \epsilon \in \mathcal{L}(P_w)\}.$$

Let w be 321-avoiding. We wish to describe (P_w, ω) more explicitly. The skew shape $\lambda/\mu = \sigma(w) \subset \mathbb{Z} \times \mathbb{Z}$ has a standard poset structure inherited from $\mathbb{Z} \times \mathbb{Z}$, viz., $(i, j) \leq (i', j')$ in $\sigma(w)$ if $i \leq i'$ and $j \leq j'$.

PROPOSITION 2.1. *Let w be a 321-avoiding permutation. Then there is an isomorphism $(P_w, \omega_w) \cong (\sigma(w), \omega(w))$ of labeled posets.*

Proof. Consider the reduced decomposition $a = (a_1, \ldots, a_p) \in R(w)$ which comes last in lexicographic order, i.e., a_1 is as large as possible, then after that a_2 is as large as possible, etc. First note that if $a_k < a_{k+1}$ then $a_{k+1} = a_k + 1$, since otherwise we could interchange a_k and a_{k+1}, thus obtaining a lexicographically later reduced decomposition. Hence a has the form

$$\begin{aligned} a_1 < a_1 + 1 < \cdots < a_1 + j_1 - 1 > a_{j_1+1} < a_{j_1+1} + 1 < \cdots < a_{j_1} + j_2 - 1 \\ > a_{j_1+j_2+1} < a_{j_1+j_2+1} + 1 < \cdots. \end{aligned} \tag{12}$$

Next note that $a_{j_1} + 1 < a_1$, since otherwise we could continually interchange a_{j_1+1} with elements to the left of it until reaching a reduced decomposition containing the consecutive elements a_{j_1+1}, $a_{j_1+1} + 1$, a_{j_1+1}. This violates the condition of Theorem 2.1(b). Similarly $a_{j_1+j_2+1} < a_{j_1+1}$, etc., so $a_1 > a_{j_1+1} > a_{j_1+j_2+1} > a_{j_1+j_2+j_3+1} > \cdots$. From this it follows that (P_a, ω) is a skew tableau with j_1 squares in the first column, labeled $a_1, a_1+1, \ldots, a_{j_1-1}$ from top to bottom, then j_2 squares in the second column, labeled $j_1+1, \ldots, j_2$ from top to bottom, etc. the columns are arranged so that the labels in each row increase by one from right to left.

We must verify that the skew shape P_w coincides with $\sigma(w)$, and that the labelings ω_w and $\omega(w)$ agree. This is a straightforward verification whose details we omit. □

Proposition 2.1 shows that if w is 321-avoiding, then the labeled poset (P_w, ω) has the properties that (a) P_w is a skew shape (namely $\sigma(w)$), and (b) P_w can be embedded as a convex subset of $\mathbb{Z} \times \mathbb{Z}$ such that $\omega(i, j) = i - j$. The next result establishes a converse.

PROPOSITION 2.2. *Let P be a finite convex subset of $\mathbb{Z} \times \mathbb{Z}$ such that if $(i, j) \in P$ then $i > j$. Then there is a unique 321-avoiding permutation w such that the labeled poset (P_w, ω) has the property that there is an isomorphism $f : P \to P_w$ satisfying $i - j = \omega(f(i, j))$ for all $(i, j) \in P$.*

Proof. Regard P as a skew shape in "English notation," so $(i, j+1)$ is to the right of (i, j) and $(i+1, j)$ is below (i, j). Let $a_1, a_2, \ldots, a_p$ be the numbers $i - j$, as (i, j) ranges over P_w, moving from top to bottom in the first column, then top to bottom in the second column, etc. Thus $a_1, \ldots, a_p$ has the form (12), where the nonempty column lengths are $j_1, j_2, \ldots$ Moreover, since P is a skew shape we have $a_1 > a_{j_1+1} > a_{j_1+j_2+1} > \cdots$. It's then easy to see (as in the proof of Proposition 2.1) that $(a_1, a_2, \ldots, a_p) \in R(w)$ for some permutation w, and that no application of commuting Coxeter relations can ever create consecutive terms, $k, k+1, k$ or $k+1, k, k+1$. Hence by Theorem 2.1, w is 321-avoiding. The proof of Proposition 2.1 then shows that $P \cong P_w$, and that defining $\omega(i, j) = i - j$ for $(i, j) \in P$ agrees with the labeling ω_w of P_w. Uniqueness of w is clear since the labeled poset (P_w, ω) determines the reduced decompositions of w and therefore also w itself. □

We thus have established a one-to-one correspondence between 321-avoiding permutations w and labeled skew shapes $(\lambda/\mu, \omega)$, where the labeling is obtained from an embedding $\lambda/\mu \subset \mathbb{Z} \times \mathbb{Z}$ by the rule $\omega(i, j) = i - j > 0$. (Two skew shapes λ/μ, $\alpha/\beta \subset \mathbb{Z} \times \mathbb{Z}$ are regarded as the same if there is an order-preserving bijection $f : \lambda/\mu \to \alpha/\beta$ which preserves labels, i.e., if $f(i, j) = (k, h)$ then $i - j = k - h$. Equivalently, α/β is obtained from λ/μ by translating by a "diagonal vector" (m, m).) Clearly we have $w \in S_n$ if and only if every label $i - j$ is less than n. For instance, there are fourteen 321-avoiding $w \in S_4$. The corresponding fourteen labeled skew shapes are:

$$\emptyset \quad 1 \quad 2 \quad 21 \quad \begin{matrix} 1 \\ 2 \end{matrix} \quad 3 \quad 32$$

$$321 \quad \begin{array}{c}2\\3\end{array} \quad \begin{array}{c}1\\2\\3\end{array} \quad \begin{array}{cc}&1\\3&\end{array} \quad \begin{array}{cc}2&1\\3&\end{array} \quad \begin{array}{cc}&1\\3&2\end{array} \quad \begin{array}{cc}2&1\\3&2\end{array}$$

The following corollary to Proposition 2.1 will be significantly generalized in Corollary 2.4.

COROLLARY 2.1. *Let w be 321-avoiding, with $\sigma(w) = \lambda/\mu$. Then the number $r(w)$ of reduced decompositions of w is given by*

$$r(w) = f^{\lambda/\mu},$$

the number of standard tableaux of shape λ/μ (i.e., the entries $1, 2, \ldots, p$ appear once each and increase in every row and column).

Proof. We know that $r(w) = e(P_w)$, the number of linear extensions of P_w. By Proposition 2.1, $P_w \cong \lambda/\mu$. But there is an obvious (and well-known) bijection between linear extensions of λ/μ and standard tableaux of shape λ/μ, and the proof follows. □

Let w be 321-avoiding with code $(c_1, \ldots, c_n)$, where the indices i for which $c_i \neq 0$ are given by $\hat{\phi}_1 < \hat{\phi}_2 < \cdots < \hat{\phi}_k$. Define the *flag* $\hat{\phi}(w)$ of w by $\hat{\phi}(w) = (\hat{\phi}_1, \hat{\phi}_2, \ldots, \hat{\phi}_k)$.

Note. The above definition of flag differs from the two definitions $\phi(w)$ and $\phi^*(w)$ in [13, p. 14]. Theorem 2.2 below is still valid with $\hat{\phi}$ replaced by ϕ^*, but not by ϕ.

The next theorem is the main result of this section.

THEOREM 2.2. *Let w be a 321-avoiding permutation, with skew shape $\sigma(w) = \lambda/\mu$ and flag $\hat{\phi} = \hat{\phi}(\sigma)$. Then*

$$\mathfrak{S}_w = s_{\lambda/\mu}(\mathcal{X}_{\hat{\phi}}),$$

the flag skew Schur function of shape λ/μ and flag $\hat{\phi}$.

Proof. Let $RK(w)$ and $\mathcal{A} = \mathcal{A}_{\lambda/\mu}(\phi)$ be as in equation (1) and before equation (8), respectively. By Theorem 1.1 it suffices to find a bijection $\psi : RK(w) \to \mathcal{A}$ such that if $\psi(a, i) = T$, then $x^i = x^T$. By Proposition 2.1, we have $R(w) = \omega\mathcal{L}(P_w)$. If $(a, i) \in RK(w)$ and $a = \omega(y_1, \ldots, y_p)$, then define $\psi(a, i)$ to be the tableaux T of shape λ/μ obtained by placing the number i_j in position y_j, for $1 \leq i \leq p$. (We are identifying the poset P_w with the skew shape λ/μ). By the theory of (P, ω)-partitions [20, Theorem 6.2], the conditions $i_1 \leq i_2 \leq \cdots \leq i_p$ and $a_j < a_{j+1} \Rightarrow i_j < i_{j+1}$ are equivalent to T being semistandard. Now the labels $\omega(y_i)$ are strictly decreasing in each row, and the last element of row j is just $\hat{\phi}_j$. Hence the condition $i_j \leq a_j$ is equivalent to T satisfying the flag condition that all entries in row j cannot exceed $\hat{\phi}_j$. Thus ψ is the desired bijection. □

An alternative description (implicit in the above proof) of those T enumerated by $\mathfrak{S}_w = s_{\lambda/\mu}(\mathcal{X}_{\hat{\phi}})$ is given by the following result.

COROLLARY 2.2. *Let w be 321-avoiding, with $(\sigma(w), \omega(w)) = (\lambda/\mu, \omega)$. Then*

$$\mathfrak{S}_w = \sum_T x^T,$$

where T ranges over all SST of shape λ/μ satisfying $T_{kh} \leq \omega(k, h) = k - h$ for all $(k, h) \in \lambda/\mu$.

Proof. The labels decrease in each row of λ/μ, and the last label in row k is $\hat{\phi}_k$. Hence if $T_{kh} \leq \hat{\phi}_k$ then $T_{kh} \leq \omega(k, h)$, and the proof follows from Theorem 2.2. □

Let us call a permutation w for which $\mathfrak{S}_w$ is a flag skew Schur function $s_{\lambda/\mu}(\mathcal{X}_\phi)$ *a skew vexillary* permutation of *shape* λ/μ. Thus Theorem 2.2 asserts that 321-avoiding permutations are skew vexillary. An obvious problem suggests itself at this point, namely, the classification of skew vexillary permutations. To this end we state the following result, which had been conjectured by one of us and which was proved by V. Reiner (private communication) based on the results of [11].

PROPOSITION 2.3. *A permutation w is skew vexillary of shape λ/μ if one can obtain the skew shape (or diagram) λ/μ from the diagram $D(w)$ of w by the following operations: (a) Any permutation of columns. (b) Interchange two consecutive rows if the top row is "contained" in the bottom row, i.e., if the top row has a square in column j, then so does the bottom row.*

We don't know whether the sufficient condition given above for w to be skew vexillary is also necessary. One could also ask if every heroic (or patriotic) permutation is skew vexillary (the converse is false, e.g., $w = 32154$ is not patriotic, but $\mathfrak{S}_w$ is skew vexillary by Proposition 2.3).

A related problem, suggested by Corollary 2.2 but which we have not looked at, is the following. For what permutations w does there exist a skew shape λ/μ and a labeling $\omega : \lambda/\mu \to \mathbb{P}$ such that

$$\mathfrak{S}_w = \sum_T x^T, \tag{13}$$

where T ranges over all SST of shape λ/μ satisfying $T_{hk} \leq \omega(h, k)$ for all $(h, k) \in \lambda/\mu$? Of course these w include the skew vexillary permutations, where we can take ω to be constant along rows and weakly increasing along columns.

We now give several corollaries to Theorem 2.2.

COROLLARY 2.3. *Let w be 321-avoiding with $\sigma(w) = \lambda/\mu$ and $\hat{\phi}(w) = (\hat{\phi}_1, \ldots, \hat{\phi}_k)$. Let $X_i = (x_1, x_2, \ldots, x_{\hat{\phi}_i})$. Then*

$$\mathfrak{S}_w = \det(h_{\lambda_i-\mu_j-i+j}(X_i))_{1\leq i,j\leq k},$$

with h_r as in (9).

Proof. Immediate from (9) and Theorem 2.2. □

COROLLARY 2.4. *Let w be 321-avoiding with $\sigma(w) = \lambda/\mu$, and let $G_w = \lim_{N\to\infty} \mathfrak{S}_{1_N\times w}$ as in Section 1. Then $G_w = s_{\lambda/\mu}$, the skew Schur function of shape λ/μ.*

Proof. If w is 321-avoiding with $\sigma(w) = \lambda/\mu$, then $1_N \times w$ is also 321-avoiding with $\sigma(1_N \times w) = \{(i, j - N) : (i, j) \in \lambda/\mu\}$, i.e., $\sigma(1^N \times w)$ is obtained by translating $\sigma(w)\, N$ units to the left (using the English coordinate system). Hence if $\hat{\phi}(w) = (\hat{\phi}_1, \ldots, \hat{\phi}_k)$ then $\hat{\phi}(1_N \times w) = (\hat{\phi}_1 + N, \ldots, \hat{\phi}_k + N)$, so

$$\begin{aligned} G_w &= \lim_{N\to\infty} s_{\lambda/\mu}(\mathcal{X}_{(\hat{\phi}_1+N,\ldots,\hat{\phi}_k+N)}) \\ &= s_{\lambda/\mu}(x_1, x_2, \ldots). \end{aligned}$$

□

The proof of Corollary 2.4 shows that not only does a 321-avoiding permutation w satisfy

$$\mathfrak{S}_w = s_{\lambda/\mu}(\mathcal{X}_\phi)$$

for some skew shape λ/μ and flag $\phi = (\phi_1, \ldots, \phi_n)$, but also

$$\mathfrak{S}_{1_N\times w} = s_{\lambda/\mu}(\mathcal{X}_{\phi+N}), \tag{14}$$

where $N \geq 0$ and $\phi + N = (\phi_1 + N, \ldots, \phi_n + N)$. The same is true for vexillary permutations by [13, (4.9)] and the definition of the flag ϕ appearing there. We may thus define a permutation w to be *strongly skew vexillary* if there is a skew shape λ/μ and flag ϕ such that for all $N \geq 0$, equation (14) holds. Letting $N \to \infty$ in (14) yields $G_w = s_{\lambda/\mu}$. Hence a necessary condition that w be strongly skew vexillary is that G_w be a skew Schur function. The converse is false, e.g., $w = 241653$ is not even skew vexillary, but $\mathfrak{S}_w = s_{3221/11}$. We don't know whether every skew vexillary permutation is strongly skew vexillary. An example of a permutation w for which $G_w \neq s_{\lambda/\mu}$ is given by $w = 246153$.

We have the following analog of Proposition 2.3 for the symmetric functions G_w. The proof, which will not be given here, is a simple consequence of the results in [11].

PROPOSITION 2.4. *A permutation w satisfies $G_w = s_{\lambda/\mu}$ if the diagram $D(w)$ of w can be transformed into the skew shape λ/μ by arbitrary row and column permutations.*

As was the case for Proposition 2.3, we don't know whether the converse to Proposition 2.4 holds.

It is a simple consequence of the Littlewood-Richardson rule [14, Chapter 1.9] that $s_{\lambda/\mu}$ is an ordinary Schur function if and only if either λ/μ or its 180° rotation $(\lambda/\mu)^r$ is an ordinary shape. Since G_w (for any permutation w) is a Schur function if and only if w is vexillary [13, (7.24)(iii)], we see that a 321-avoiding permutation w is vexillary if and only if either $\sigma(w)$ or $\sigma(w)^r$ is an ordinary shape. This observation has the following curious consequence.

COROLLARY 2.5. *Let $g(n)$ be the number of permutations in S_n which are both 2143-avoiding and 321-avoiding. Then*

$$g(n) = 2^{n+1} - \binom{n+1}{3} - 2n - 1.$$

Proof. Let $g_1(n)$ (respectively, $g_2(n)$) be the number of 321-avoiding $w \in S_n$ such that $\sigma(w)$ (respectively, $\sigma(w)^r$) is an ordinary shape. Let $g_3(n)$ be the number of 321-avoiding $w \in S_n$ such that both $\sigma(w)$ and $\sigma(w)^r$ are ordinary shapes, i.e., $\sigma(w)$ is a rectangle. Thus $g(n) = g_1(n) + g_2(n) - g_3(n)$. Reflecting an ordinary shape about a diagonal from northeast to southwest which fixes the upper right-hand square of $\sigma(w)$ yields a skew diagram $\sigma(w)^r$ whose labels $\omega(i, j) = i - j$ are preserved. Hence $g_1(n) = g_2(n)$, so $g(n) = 2g_1(n) - g_3(n)$.

Fix integers $a, b \geq 1$, and suppose $\sigma(w) = \lambda$ with $\lambda_1 = a$ and $\lambda_1' = b$. In order that $\omega(i, j) \leq n - 1$ for all $(i, j) \in \lambda$, we must have $a + b \leq n$. There are $n + 1 - a - b$ labelings ω of λ with $\omega(i, j) \leq n - 1$ for all $(i, j) \in \lambda$, and there are $\binom{a+b-2}{a-1}$ possible partitions λ. We can also have $\lambda = \emptyset$, so

$$\begin{aligned} g_1(n) &= 1 + \sum_{\substack{a+b\leq n \\ a,b\geq 1}} (n + 1 - a - b)\binom{a+b-2}{a-1} \\ &= 1 + \sum_{c=0}^{n-2}(n-1-c)\sum_{d=0}^{c}\binom{c}{d} \\ &= 1 + \sum_{c=0}^{n-2}(n-1-c)2^c \\ &= 2^n - n. \end{aligned}$$

If λ is an $a \times b$ rectangle with $a, b \geq 1$, then $a + b \leq n$ and there are $n + 1 - a - b$ labelings ω of λ. We can also have $\lambda = \emptyset$, so

$$g_2(n) = 1 + \sum_{\substack{a+b\leq n \\ a,b\geq 1}} (n + 1 - a - b)$$

$$= 1 + \sum_{c=0}^{n-2}(n-1-c)\sum_{d=0}^{c} 1$$

$$= 1 + \sum_{c=0}^{n-2}(n-1-c)(c+1)$$

$$= 1 + \binom{n+1}{3}.$$

Hence

$$\begin{aligned} g(n) &= 2g_1(n) - g_2(n) \\ &= 2(2^n - n) - 1 - \binom{n+1}{3} \\ &= 2^{n+1} - \binom{n+1}{3} - 2n - 1, \end{aligned}$$

as claimed. □

Our final result of this section is a Schubert analogue of the well-known formula

$$s_1^p = \sum_{\lambda \vdash p} f^\lambda s_\lambda \tag{15}$$

occurring in the theory of symmetric functions (e.g., [14, p. 62]). For convenience write

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

PROPOSITION 2.5. *Let* $1 \leq a_1 < a_2 < \cdots < a_p$. *Then*

$$T_{a_1}T_{a_2}\cdots T_{a_p} = \sum_w \mathfrak{S}_w, \tag{16}$$

where w *ranges over all distinct permutations (necessarily 321-avoiding) whose reduced decompositions are permutations of* $(a_1, a_2, \ldots, a_p)$ *(i.e., are of the form* $(a_{\pi(1)}a_{\pi(2)}, \ldots, a_{\pi(p)})$, *for* $\pi \in S_p$).

Proof. First note that since the a_is are distinct, any permutation of $(a_1, a_2, \ldots, a_p)$ is a reduced decomposition of some 321-avoiding permutation w. Since all reduced decompositions of such w are permutations of one another, it follows that the set $S(a_1, a_2, \ldots, a_p)$ of all permutations of $(a_1, a_2, \ldots, a_p)$ is a disjoint union of $R(w)$s for certain 321-avoiding w's. Thus by Theorem 1.1 it suffices to show that

$$T_{a_1}T_{a_2}\cdots T_{a_p} = \sum_{(b,i)} x_{i_1}x_{i_2}\cdots x_{i_p}, \tag{17}$$

where (b, i) ranges over all pairs $b = (b_1, \dots, b_p) \in \mathcal{S}(a_1, a_2, \dots, a_p)$ and $i \in K(b)$. A typical monomial in the expansion of $T_{a_1}T_{a_2}\cdots T_{a_p}$ is $x_{j_1}x_{j_2}\cdots x_{j_p}$, where $1 \le j_k \le a_k$. We can identify this monomial with the function $f : \{a_1, \dots, a_p\} \to \mathbb{P}$ defined by $f(a_k) = j_k$. By the theory of P-partitions [20, Theorem 6.2; 18, Lemma 4.5.1], there is a unique permutation $\pi = (b_1, \dots, b_p) \in \mathcal{S}(a_1, \dots, a_p)$ such that $f(b_1) \le f(b_2) \le \cdots \le f(b_p)$, and $f(b_i) < f(b_{i+1})$ if $b_i < b_{i+1}$. (The results in [20, 18] are stated in a dual but equivalent form.) Hence we have a bijection between the terms $x_{i_1}\cdots x_{i_p}$ in the sum (17) and the monomials $x_{j_1}\cdots x_{j_p}$ appearing in $T_{a_1}\cdots T_{a_p}$ as desired. □

Note: Proposition 2.5 can be proved by several other means, e.g., by use of Monk's rule.

Suppose we replace each $T_{a_i} = \mathfrak{S}_{s_{a_i}}$ in (16) by its "stabilization" $G_{s_{a_i}} = x_1 + x_2 + \cdots$ (the Schur function s_1). Similarly in the right-hand side replace each $\mathfrak{S}_w$ by G_w. Although in general the operation $\mathfrak{S}_w \mapsto G_w$ does not commute with multiplication (i.e., does not extend to a ring homomorphism), in the present case it can be checked that the two limits agree. Equivalently,

$$\sum_w G_w = \sum_{\lambda \vdash p} f^\lambda s_\lambda,$$

where w ranges over the same set as in (16). Thus (15) is a "stable" version of (16).

It would be interesting to generalize (16) to arbitrary sequences $1 \le a_1 \le a_2 \le \cdots \le a_p$, but we have been unable to find such a result. If for $1 \le a_1 \le a_2 \le \cdots \le a_p$ we write

$$T_{a_1}T_{a_2}\cdots T_{a_p} = \sum_{\substack{w \\ \ell(w)=p}} c_w \mathfrak{S}_w,$$

then it is no longer true that $\sum c_w G_w = s_1^p$ (unless $a_1 < a_2 < \cdots < a_p$). For instance, for $T_1T_2^2T_3$ we get $\sum c_w G_w = s_4 + 3s_{31} + 2s_{22} + s_{211}$. It would be interesting to find an explicit description of $\sum c_w G_w$.

3. Principal specialization

Define the *stable principal specialization* of a Schubert polynomial $\mathfrak{S}_w$ (or more generally of any power series $f(x_1, x_2, \cdots)$) to be the polynomial $\mathfrak{S}_w(1, q, q^2, \cdots)$ respectively, power series $f(1, q, q^2, \cdots)$) in the variable q. If S is a finite subset of $\mathbb{Z}$, then we write $\sum S = \sum_{i \in S} i$. The following formula was conjectured by Macdonald [13,(6.11$_q$?), p.91] and is proved in [7]. For any permutation w, we have

$$\mathfrak{S}_w(1, q, q^2, \dots) = \frac{1}{[p]!} \sum q^{\mathrm{comaj(a)}} [a_1][a_2]\cdots[a_p], \tag{18}$$

where the sum ranges over all $a = (a_1, \ldots, a_p) \in R(w)$, $[b] = 1 - q^b$, $[p]! = [1][2]\cdots[p]$, and $\mathrm{comaj}(a) = \sum\{i : 1 \le i \le p-1,\ a_i < a_{i+1}\}$ (the *comajor index* of a). Here we will deduce some properties of the power series $s_{\lambda/\mu}(1, q, q^2, \ldots)$ from (18) in the case when w is 321-avoiding. (It is also possible to prove (18) for 321-avoiding w using properties of skew Schur functions $s_{\lambda/\mu}$, but there is no point in doing so since (18) is now known to hold for all w.)

LEMMA 3.1. *Let w be 321-avoiding of length p, with $\sigma(w) = \lambda/\mu$. Then*

$$s_{\lambda/\mu}(1, q, q^2, \ldots) = \frac{\sum\limits_{a \in R(w)} q^{\mathrm{comaj}(a)}}{[p]!}.$$

Proof. Regard $(\lambda/\mu, \omega) = (\sigma(w), \omega(w))$ as a labeled poset (P, ω), as in the proof of Theorem 2.2. Let ϖ be the labeling of P defined by $\varpi(s) = -\omega(s)$ for all $s \in P$. A *(P, ϖ)-partition* π, as defined in [20, p. 1], is an order-reversing map $\pi : P \to \mathbb{N}$ such that if $s < t$ in P and $\varpi(s) > \varpi(t)$, then $\pi(s) > \pi(t)$. Hence π is obtained by inserting nonnegative integers into the squares of λ/μ which are weakly decreasing along rows and strictly decreasing along columns. By the standard combinatorial interpretation of skew Schur functions [21, Definition 12.1] we have

$$s_{\lambda/\mu}(1, q, q^2, \ldots) = \sum_{\pi} q^{|\pi|},$$

where π ranges over all (P, ϖ)-partitions and $|\pi| = \sum_{s \in P} \varpi(s)$.

It follows from [20, Theorem 7.2] that we have

$$s_{\lambda/\mu}(1, q, q^2, \ldots) = \frac{\sum\limits_{\epsilon} q^{\mathrm{maj}(\varpi(\epsilon))}}{[p]!}$$

where ϵ ranges over all linear extensions $(\epsilon_1, \ldots, \epsilon_p)$ of P, $\varpi(\epsilon) = (\varpi(\epsilon_1), \ldots, \varpi(\epsilon_p))$, and $\mathrm{maj}(b_1, \ldots, b_p) = \sum\{1 \le i \le p-1 : b_i > b_{i+1}\}$. Clearly $\mathrm{maj}(\varpi(\epsilon)) = \mathrm{comaj}(\omega(\epsilon))$. By Theorem 2.1 and Proposition 2.1 the sequences $\omega(\epsilon)$ are just the reduced decompositions of w, and the proof follows. □

For a tableau $T = (T_{ij})$, write $|T| = \sum T_{ij}$.

THEOREM 3.1. *Let $\lambda/\mu \subset \mathbb{Z} \times \mathbb{Z}$ be a skew shape, embedded in $\mathbb{Z} \times \mathbb{Z}$ so that if $(i, j) \in \lambda/\mu$ then $i - j > 0$. Let*

$$t_{\lambda/\mu}(q) = \sum_{T} q^{|T|}, \tag{19}$$

summed over all SST (allowing 0 as a part) $T = (T_{ij})$ of shape λ/μ such that $T_{ij} < i - j$ for all $(i, j) \in \lambda/\mu$. (Equivalently, $T_{ij} < \widehat{\phi}_i$, where $\widehat{\phi}_i = i - j$ for (i, j) the rightmost square of T in row i.) Then

$$s_{\lambda/\mu}(1, q, q^2, \ldots) = \frac{t_{\lambda/\mu}(q)}{\prod_{(i,j)\in\lambda/\mu} [i-j]}. \tag{20}$$

Proof. Let w be 321-avoiding with $\sigma(w) = \lambda/\mu$ (exists by Proposition 2.2). For all $(a_1, \ldots, a_p) \in R(w)$ we have

$$[a_1]\cdots[a_p] = \prod_{(i,j)\in\lambda/\mu} [\omega(i, j)],$$

by Proposition 2.1. Hence by (20) and Lemma 3.1, we have

$$s_{\lambda/\mu}(1, q, q^2, \ldots) = \frac{\mathfrak{S}_w(1, q, q^2, \ldots)}{\prod_{(i,j)\in\lambda/\mu} [\omega(i, j)]}.$$

But $\mathfrak{S}_w(1, q, q^2, \ldots) = t_{\lambda/\mu}(q)$ by Theorem 2.2, and the proof follows. □

Example 3.1. Let $(\lambda/\mu, \omega)$ be given by

$$\begin{matrix} & 2 & 1 \\ 4 & 3 & \end{matrix}$$

The tableaux enumerated by $t_{\lambda/\mu}(q)$ are

$$\begin{matrix} 00 & 00 & 00 & 00 & 00 \\ 01 & 11 & 02 & 12 & 22 \end{matrix}$$

Hence

$$s_{\lambda/\mu}(1, q, q^2, \ldots) = \frac{q + 2q^2 + q^3 + q^4}{[1][2][3][4]}.$$

Let us consider the special case $\mu = \emptyset$ of Theorem 3.1. Suppose $\lambda = \{(i, j) : 1 \le i \le \ell(\lambda), 1 \le j \le \lambda_i\}$. It is well known [21, Theorem 15.3; 14, Example 1, p. 27] that for any $m \ge 1$,

$$s_\lambda(1, q, \ldots, q^{m-1}) = H_\lambda(q) \prod_{(i,j)\in\lambda} [j - i + m], \tag{21}$$

where $H_\lambda(q)$ is independent of m. (An explicit formula exists for $H_\lambda(q)$, but this is irrelevant here.) Letting $m \to \infty$ in (21) yields $s_\lambda(1, q, q^2, \ldots) = H_\lambda(q)$, so

$$s_\lambda(1, q, q^2, \ldots) = \frac{s_\lambda(1, q, \ldots, q^{m-1})}{\prod_{(i,j)\in\lambda} [j - i + m]}. \tag{22}$$

Equation (22) is similar to (20). We can change the embedding $\lambda \subset \mathbb{Z} \times \mathbb{Z}$ so that the exponent $j - i + m$ in (22) becomes simply $j - i$. However, in (20) the corresponding exponent is $i - j$. We can reconcile the two formulas by appealing to the *reciprocity theorem* for (P, ω)-partitions [20 Theorem 10.1], which implies that if $s_{\lambda/\mu}(1, q, q^2, \ldots) = G_{\lambda/\mu}(q)$, then

$$G_{\lambda/\mu}(1/q) = (-1)^{|\lambda/\mu|} q^{|\lambda/\mu|} G_{(\lambda/\mu)'}(q)$$

(as rational functions of q). (See also [20, Proposition 11.1].) It then follows easily from (20) that if λ is embedded in $\mathbb{Z}\times\mathbb{Z}$ so that the main diagonal elements (i, j) satisfy $i - j = m$, then

$$q^{\eta(\lambda)} s_\lambda(1, q, q^2, \ldots) = \frac{s_{\lambda'}(1, q, \ldots, q^{m-1})}{\prod_{(i,j)\in\lambda} [i - j]},$$

where $\eta(\lambda) = \sum \binom{\lambda_i'}{2} - \sum \binom{\lambda_i}{2}$. It is easy to give a direct combinatorial proof that

$$t_\lambda(q) = q^{\eta(\lambda)} s_{\lambda'}(1, q, \ldots, q^{m-1}),$$

thereby proving (20) directly from known results when $\mu = \emptyset$. But for general μ, it seems to be new that the rational function $s_{\lambda/\mu}(1, q, q^2, \ldots)$ has (not necessarily least) denominator $\Pi[i - j]$ as in (20). Moreover, the numerator coefficients are then nonnegative. Note that the left-hand side of (20) does not depend on the actual embedding of λ/μ in $\mathbb{Z} \times \mathbb{Z}$, while the right-hand side does depend on this embedding. It follows that the denominator of $s_{\lambda/\mu}(1, q, q^2, \ldots)$ can be chosen to be the greatest common divisor of $\Pi[i - j]$, over *all* embeddings $\lambda/\mu \subset \mathbb{Z} \times \mathbb{Z}$ (for which $i - j > 0$ for all $(i, j) \in \lambda/\mu$). For instance, suppose $\lambda/\mu = (4, 4, 3)/(2, 1)$. One embedding gives denominator $[1][2]^2[3][4]^2[5][6]$, while another gives $[2][3]^2[4][5]^2[6][7]$. Hence $s_{\lambda/\mu}(1, q, q^2, \ldots)$ has denominator $[1]^3[2][3][4][5][6]$. (No further embeddings lead to a smaller denominator.)

Let us mention one further special case of Theorem 3.1. A *border strip* (or *rim hook*) [14, Example 11, p. 31] is a skew shape λ/μ such that for any two consecutive rows there is exactly one column which intersects both rows. Such a shape can be embedded in $\mathbb{Z} \times \mathbb{Z}$ so that the labels $i - j$ consist simply of $1, 2, \ldots, p$ (once each). Assume such an embedding has been chosen. By Theorem 3.1 we have

$$s_{\lambda/\mu}(1, q, q^2, \ldots) = \frac{t_{\lambda/\mu}(q)}{[p]!}.$$

On the other hand, it is a well-known consequence of the theory of (P, ω)-partitions (see [20, Corollary 7.2]) that

$$s_{\lambda/\mu}(1, q, q^2, \ldots) = \frac{\sum_T q^{\mathrm{maj}(T)}}{[p]!},$$

where T ranges over all skew SYT of shape λ/μ, and where

$$\mathrm{maj}(T) = \sum\{i : i+1 \text{ appears in a lower row of } T \text{ than } i\}.$$

Hence we have the curious formula

$$t_{\lambda/\mu}(q) = \sum_T q^{\mathrm{maj}(T)} \tag{23}$$

where λ/μ is a border strip and $t_{\lambda/\mu}(q)$ is given by (19). R. Simion has observed that a bijective proof of (23) is a simple consequence of a well-known bijection $f : \mathcal{S}_p \to \mathcal{S}_p$ of Foata [6] satisfying $\ell(w) = \mathrm{maj}(f(w))$.

4. Open problems

For the reader's convenience we list all open problems and conjectures appearing in this paper.

- Find a direct description of the multiset $\mathcal{M}_w$. (after Example 1.2)
- Classify patriotic (= heroic) permutations. In particular, is every saturated permutation heroic? (after Proposition 1.4)
- Is every heroic permutation skew vexillary? (after Corollary 2.2)
- Classify all skew vexillary permutations. In particular, does the converse to Proposition 2.3 hold? (after Proposition 2.3)
- For what permutations w does $\mathfrak{S}_w = \sum_T x^T$, as defined by equation (13)?
- Is every skew vexillary permutation strongly skew vexillary? (after Corollary 2.4)
- When is G_w a skew Schur function? In particular, is the converse to Proposition 2.4 valid?
- Expand $T_{a_1} T_{a_2} \cdots T_{a_p}$ in terms of Schubert polynomials for any $a_1 \leq a_2 \leq \cdots \leq a_p$. (after Proposition 2.5)
- In the previous problem if $T_{a_1} T_{a_2} \cdots T_{a_p} = \sum c_w \mathfrak{S}_w$, then find a simple explicit description of the symmetric function $\sum c_w G_w$. (after Proposition 2.5)

Added in proof. V. Reiner has observed that 214365 is saturated but not heroic, and that 246153 is heroic but not skew vexillary. Moreover, he has solved the last open problem of Section 4.

References

1. E. Bender and D. Knuth, "Enumeration of plane partitions," *J. Combin. Theory* **13** (1972), 40–54.
2. N. Bergeron, "A combinatorial construction of the Schubert polynomials," *J. Combin. Theory Series A*, **60** (1992), 168-182.
3. I.N. Bernstein, I.M. Gelfand, and S.I. Gelfand, "Schubert cells and cohomology of the spaces G/P," *Russian Math, Surveys* **28**, 3, (1973), 1–26. Translated from *Uspekhi Mat. Nauk* **28** (1973), 3–26.
4. M. Demazure, "Désingularisation des variétés Schubert généralisées," *Ann. Sc. E.N.S.(4)* **7** (1974), 53–88.
5. P. Edelman and C. Greene, "Balanced tableaux," *Adv. Math.* **63** (1987), 42–99.
6. D. Foata, "On the Netto inversion number of a sequence," *Proc. Amer. Math. Soc.* **19** (1968), 236–240.
7. S. Fomin and R. Stanley, "Schubert polynomials and the nilCoxeter algebra," *Adv. Math.*, to appear.
8. F.R. Gantmacher, *The Theory of Matrices*, Vol. 1, Chelsea, New York, 1959.
9. F.M. Goodman, P. de la Harpe, and V.F.R. Jones, *Coxeter Graphs and Towers of Algebras*, Springer-Verlag, New York, 1989.
10. D.E. Knuth, *The Art of Computer Programming*, vol. 3, Addison-Wesley, Reading, MA, 1973.
11. W. Kraśkiewicz and P. Pragacz, "Schubert polynomials and Schubert functors," preprint.
12. 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. Sc. Paris* **295** (1982), 629–633.
13. I.G. Macdonald, *Notes on Schubert Polynomials*, Laboratoire de combinatoire et d'informatique mathématique (LACIM), Université du Québec à Montréal, Montreal, 1991.
14. I.G. Macdonald, *Symmetric Functions and Hall Polynomials*, Oxford University Press, Oxford, 1979.
15. H. Matsumoto, "Générateurs et relations des groupes de Weyl généralisées," *C. R. Acad. Sci. Paris* **258** (1964), 3419–3422.
16. V. Reiner and M. Shimozono, "Key polynomials and a flagged Littlewood-Richardson rule", preprint.
17. R. Simion and F.W. Schmidt, "Restricted permutations," *European J. Combin.* **6** (1985), 383–406.
18. R. Stanley, *Enumerative Combinatorics*, Vol. 1, Wadsworth & Brooks/Cole, Belmont, CA, 1986.
19. R. Stanley, "On the number of reduced decompositions of elements of Coxeter groups," *European J. Combin.* **5** (1984), 359–372.
20. R. Stanley, "Ordered structures and partitions," *Mem. Amer. Math. Soc.* **119**, 1972.
21. R. Stanley, "Theory and application of plane partitions," Parts 1 and 2, *Studies in Appl. Math.* **50** (1971), 167–188, 259–279.
22. J. Tits, "Le problème des mots dans les groupes de Coxeter," *Symposia Math.*, Vol. 1, Academic Press, London, 1969, 175–185.
23 M.L. Wachs, "Flagged Schur functions, Schubert polynomials, and symmetrizing operators," *J. Combin. Theory Series A* **40** (1985), 276–289.