Séminaire Lotharingien de Combinatoire **89B** (2023)
Article #33, 12 pp.

Proceedings of the 35th Conference on Formal Power Series and Algebraic Combinatorics (Davis)

$[\mathrm{Perm}_n]$ via QSym_n^+

Philippe Nadeau*[1] and Vasu Tewari†[2]

[1] *Univ Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, 43 blvd. du 11 novembre 1918, F-69622 Villeurbanne cedex, France*
[2] *Department of Mathematics, University of Hawaii at Manoa, Honolulu, HI 96822, USA*

Abstract. We give a combinatorial interpretation for the expansion of the cohomology class of the permutahedral variety $Perm_n$ in Schubert classes. This requires understanding Schubert polynomials modulo the ideal QSym_n^+ of positive degree quasisymmetric polynomials. We introduce a new basis for the polynomial ring that we call forest polynomials, together with a bijective correspondence. Both constructions are of independent interest.

Résumé. Nous donnons une interprétation combinatoire au développement de la classe de cohomologie de la variété permutaédrale $Perm_n$ sur les classes de Schubert. Pour cela, il faut comprendre la réduction des polynômes de Schubert modulo l'idéal QSym_n^+ des polynômes quasi-symétriques de degré positif. Nous introduisons une nouvelle base pour l'anneau de polynômes que nous appelons polynômes forestiers, ainsi qu'une correspondance bijective, toutes deux ayant un intérêt pour elles-mêmes.

Keywords: permutahedral variety, quasisymmetric polynomials, binary trees, linear extensions.

1 Introduction

Schubert calculus is the study of certain enumerative problems in algebraic geometry. These can be often phrased as finding intersection numbers of subvarieties subject to certain conditions. Littlewood–Richardson numbers arise for instance as generic triple intersection numbers of Schubert classes in the Grassmannian, and are known to count various families of objects. The same intersection problem in the flag variety has been the study of intensive research for at least the past forty years.

Here we study a particular problem in the flag variety, namely the intersection of the so-called permutahedral (sub)variety $Perm_n$ with a generic Schubert variety X^w. If $w \in \mathbb{S}_n$ has length $n-1$, the cardinality $a_w = |Perm_n \cap X^w|$ of intersection points is a well-defined nonnegative integer. Several properties of these numbers were obtained by the authors in [8], as well as their evaluation in special cases, all of which point to rich combinatorics.

*nadeau@math.univ-lyon1.fr. P.N. was supported by the project ANR19-CE48-011-01 (COMBINÉ).
†vvtewari@math.hawaii.edu. V.T. was supported by Simons Grant #855592.

In this work we complete these results with a combinatorial interpretation for a_w. Recall that $i_1 i_2 \dots i_k$ is a *reduced word* for a permutation w if $w = s_{i_1} s_{i_2} \dots s_{i_k}$ where s_i is the transposition $(i \leftrightarrow i+1)$, and k is minimal, given by the length $\ell(w)$. We denote the set of reduced words for w by $\mathrm{Red}(w)$. We also need a particular parking procedure in order to state our main result.

WF-parking procedure: Consider parking spots indexed by $\mathbb{Z}$, initially all empty. Cars $1, 2, \dots$ arrive successively, with car i having preferred spot v_i, and want to park at (empty) spots. Assume inductively that $i-1$ cars have already parked. If spot v_i is empty, then car i parks there. Otherwise, $v_i \in [a, b]$ an interval of occupied spots with spots $a-1$ and $b+1$ being free.[1] Define v_j to be the preferred spot of the car that parked last in $[a, b]$; that is, $j < i$ is maximal such that $v_j \in [a, b]$. Then, car i parks in $b+1$ if $v_i \geq v_j$, while it parks in $a-1$ if $v_i < v_j$.

After k cars have parked they occupy a k-subset $I(v_1 \cdots v_k) \subset \mathbb{Z}$. A preference word $v_1 \cdots v_k$ is called a *WF-parking word* if $I(v_1 \cdots v_k) = \{1, \dots, k\}$.

Theorem 1. *If $w \in \mathbb{S}_n$ has length $n-1$, then a_w is the number of reduced words of w^{-1} that are also WF-parking words.*

Example 2. *Consider $w = 21543 \in \mathbb{S}_5$ with $\ell(w) = 4$. To compute a_w, we need to compute WF-parking words in $\mathrm{Red}(w^{-1}) = \mathrm{Red}(w)$. These are given explicitly in (5.1), and only the first four of them are WF-parking words. It follows $a_{21543} = 4$.*

We will explain the main steps of the proof leading to this result, and stress in particular the rôle played by the *forest polynomials* that we introduce in Section 3. Let us first recall the geometrical context and previous results.

We will be brief on the geometry and refer to [8] for more information. The *complete flag variety* Fl_n is the set of complete flags $V_\bullet = (V_0 \subseteq V_1 \subseteq V_2 \subseteq \cdots \subseteq V_n)$ where each V_i is a linear subspace of $\mathbb{C}^n$ of dimension i. It is a smooth projective variety of dimension $\binom{n}{2}$ over $\mathbb{C}$. Inside Fl_n we have first and foremost the *Schubert varieties* X^w indexed by $w \in \mathbb{S}_n$, whose cohomology classes $\sigma_w := [X^w]$ form a linear basis of the cohomology ring $H^*(\mathrm{Fl}_n, \mathbb{Q})$. In this work we consider the *permutahedral variety* Perm_n. It is an ubiquitous toric variety, encoded by the normal fan of the standard permutahedron, and occurs in Fl_n as a generic torus orbit or a regular semisimple Hessenberg variety.

Now by standard cohomology theory, the intersection numbers a_w defined above are known to be given by the coefficients in the expansion of the class $[\mathrm{Perm}_n]$ in terms of Schubert classes. Using a result of Anderson–Tymoczko [1], the authors were led in [8] to an algebraic expression for a_w that we now recall.

Henceforth, set $\mathbb{Q}[\mathbf{x}_n] = \mathbb{Q}[x_1, \dots, x_n]$. Let *divided symmetrization* (DS), introduced under that name by Postnikov [10, Section 3], be the linear form that takes a homogeneous

[1] Recall that in the classical parking algorithm, car i simply parks in spot $b+1$.

polynomial $f(x_1,\ldots,x_n) \in \mathbb{Q}[\mathbf{x}_n]$ of degree $n-1$ as input and outputs $\langle f\rangle_n \in \mathbb{Q}$ as:

$$\langle f\rangle_n := \sum_{w\in\mathbb{S}_n} w\cdot\left(\frac{f(x_1,\ldots,x_n)}{\prod_{1\le i\le n-1}(x_i-x_{i+1})}\right). \tag{1.1}$$

For $w\in\mathbb{S}_n$ with length $n-1$, we have [8, Theorem 3.2.1]

$$a_w = \langle\mathfrak{S}_w\rangle_n. \tag{1.2}$$

Here $\mathfrak{S}_w$ is a Schubert polynomial, see (2.1) for a definition. Before we explain how to go from the expression (1.2) to Theorem 1, note that the authors had previously obtained results on a_w using the algebra of $\mathbb{S}_n$-invariants of the cohomology ring $H^*(\mathrm{Perm}_n,\mathbb{Q})$. This was based on work of Klyachko [6] in a crucial way. The main result was an explicit formula for a_w as a positive *rational* linear combination of *mixed Eulerian numbers* [8, Theorem 5.1.2]. The formula is reminiscent of Macdonald's famous reduced word identity and this aspect is explored in [7].

A rational expression, even if positive, is unsatisfying since we know $a_w\in\mathbb{N}$. Theorem 1 thus provides a pleasant answer in that respect. Surprisingly the road to the interpretation therein goes through the quotient $\mathbb{Q}[\mathbf{x}_n]/\mathrm{QSym}_n^+$, first studied by Aval–Bergeron–Bergeron [4]. We quickly introduce the relevant notions (see Section 2 for QSym_n^+). Let $\mathcal{C}_n$ denote the set of weak compositions $(c_1,\ldots,c_n)$ such that for all $1\le j\le n$ we have $c_j+\cdots+c_n\le n-j$. We call its elements *ABB compositions*. We refer to the monomials $\mathbf{x}^c := x_1^{c_1}\cdots x_n^{c_n}$ where $c=(c_1,\ldots,c_n)\in\mathcal{C}_n$ as *ABB monomials*. Aval–Bergeron–Bergeron showed the direct sum decomposition:

$$\mathbb{Q}[\mathbf{x}_n] = \mathbb{Q}\{\mathbf{x}^c \mid c\in\mathcal{C}_n\}\oplus\mathrm{QSym}_n^+, \tag{1.3}$$

from which they inferred that $\dim(\mathbb{Q}[\mathbf{x}_n]/\mathrm{QSym}_n^+) = \mathrm{Cat}_n := \frac{1}{n+1}\binom{2n}{n}$. The following result connects the ABB decomposition with DS.

Theorem 3 ([9, Theorem 1.3])**.** *Express a degree $n-1$ homogeneous polynomial f as $f=g+h$ where $g\in\mathbb{Q}\{\mathbf{x}^c \mid c\in\mathcal{C}_n\}$ and $h\in\mathrm{QSym}_n^+$. Then $\langle f\rangle_n = g(1,1,\ldots,1)$.*

Given Formula (1.2), we are led to consider the decomposition of $\mathfrak{S}_w$ as in Theorem 3. The corresponding polynomial g will turn out to be a $\mathbb{N}$-combination of ABB monomials, thus showing $a_w\in\mathbb{N}$.

We define an $\mathbb{N}$-vector to be a sequence of nonnegative integers $(c_i)_{i\in\mathbb{Z}_{\ge 1}}$ where all but finitely many c_i are 0. The *support* of an $\mathbb{N}$-vector c is the set of indices $i\in\mathbb{Z}_{>0}$ such that $c_i>0$. We denote the set of $\mathbb{N}$-vectors by Codes. Occasionally we truncate an $\mathbb{N}$-vector to a finite sequence $(c_1,\ldots,c_n)$ for some positive integer n. It is then understood that $c_i=0$ for all $i>n$.

We propose a family of polynomials $\{\mathfrak{p}_c\}_{c\in\mathrm{Codes}}$ which we call *forest polynomials*; the reason will become clear in the sequel. They turn out to enjoy several nice combinatorial properties, and form a $\mathbb{Z}$-basis of the space of all polynomials. Our interest in them stems from the following theorem; its first part relies on the *Word-Forest Correspondence* that we introduce in Section 4.

Theorem 4. *The following hold.*

(i) *Every Schubert polynomial is a nonnegative integral combination of forest polynomials.*

(ii) *Let us restrict our attention to forest polynomials that only use variables x_1 through x_n. We then have $\mathfrak{p}_c$ modulo $\mathrm{QSym}_n^+ \neq 0$ if and only if $c \in \mathcal{C}_n$. In the latter case, $\mathfrak{p}_c$ is a nonnegative integral combination of ABB monomials.*

These two results, given in Sections 3 and 4 respectively, will help give a manifestly nonnegative integral expression for a_w that can eventually be interpreted as in Theorem 1; see Section 5. We close by mentioning some other applications in a final section.

2 Combinatorial preliminaries

Quasisymmetric polynomials. A *strong composition* $\alpha := (\alpha_1, \dots, \alpha_k)$ is a sequence of positive integers. We use $\ell(\alpha) := k$ to denote the number of *parts* in α. A polynomial $f \in \mathbb{Q}[\mathbf{x}_n]$ is *quasisymmetric* if for every strong composition α the coefficient of $x_1^{\alpha_1} \cdots x_k^{\alpha_k}$ equals that of $x_{i_1}^{\alpha_1} \cdots x_{i_k}^{\alpha_k}$ for every $1 \leq i_1 < \cdots < i_k \leq n$. Clearly every polynomial symmetric in x_1 through x_n is also quasisymmetric. A linear basis for the ring of quasisymmetric polynomials in x_1 through x_n is given by the *fundamental quasisymmetric polynomials* $\mathsf{L}_\alpha(x_1, \dots, x_n)$ with $\ell(\alpha) \leq n$. We are interested in the ideal QSym_n^+ generated by positive degree quasisymmetric polynomials in $\mathbb{Q}[\mathbf{x}_n]$. Equivalently, it is the ideal generated by all $\mathsf{L}_\alpha(x_1, \dots, x_n)$ for $\ell(\alpha) > 0$.

Schubert polynomials. Given $\mathbf{i} = i_1 \dots i_k \in \mathbb{Z}^*$, let $\mathrm{Comp}(\mathbf{i})$ denote the set of sequences $(a_1 \geq \cdots \geq a_k)$ satisfying $1 \leq a_j \leq i_j$ for all $j = 1, \dots, k$ and $a_j > a_{j+1}$ if $i_j > i_{j+1}$ for $j = 1, \dots, k-1$. Define $\mathfrak{F}(\mathbf{i})$ by

$$\mathfrak{F}(\mathbf{i}) = \sum_{(a_1, \dots, a_k) \in \mathrm{Comp}(\mathbf{i})} x_{a_1} \cdots x_{a_k}.$$

If the sum on the right is empty, then $\mathfrak{F}(\mathbf{i}) = 0$. Otherwise $\mathfrak{F}(\mathbf{i})$ can always be written as a *slide polynomial* $\mathfrak{F}_c$ [3].[2] As an example, $\mathfrak{F}(423) = x_2^2x_4 + x_2^2x_3 + x_1x_2x_4 + x_1x_2x_3 + x_1^2x_4 + x_1^2x_3 + x_1^2x_2$, which by the convention in [3] is $\mathfrak{F}_{(0,2,0,1)}$.

Now given $w \in \mathbb{S}_n$, let $\mathrm{Red}(w)$ denote the set of *reduced words* of w. We then have the following celebrated description of $\mathfrak{S}_w$ due to Billey–Jockusch–Stanley [5]:

$$\mathfrak{S}_w = \sum_{\mathbf{i} \in \mathrm{Red}(w^{-1})} \mathfrak{F}(\mathbf{i}). \tag{2.1}$$

[2] Assaf–Searles index slides by $\mathbb{N}$-vectors, and the latter can be read off from the exponent vector of the revlex leading monomial.

For example consider $w = 14253 \in \mathbb{S}_5$. Then $\mathrm{Red}(w^{-1}) = \{243, 423\}$. We just computed $\mathfrak{F}(423)$ above, and $\mathfrak{F}(243) = x_1x_2^2 = \mathfrak{F}_{(1,2)}$. Thus $\mathfrak{S}_{14253} = \mathfrak{F}(243) + \mathfrak{F}(423)$ is given by

$$\mathfrak{S}_{14253} = x_1x_2^2 + x_1^2x_2 + x_1^2x_3 + x_1x_2x_3 + x_2^2x_3 + x_1^2x_4 + x_1x_2x_4 + x_2^2x_4. \tag{2.2}$$

Indexed forests. An *indexed forest* F is a set of complete binary trees, each with leaves labeled from left to right by an interval in $\mathbb{Z}_{>0}$. These intervals must moreover partition $\mathbb{Z}_{>0}$, and all but a finite number of them are singletons, i.e. the corresponding tree is a leaf. See Figure 1 (ignoring red labels for now), where we identify leaves with $\mathbb{Z}_{>0}$.

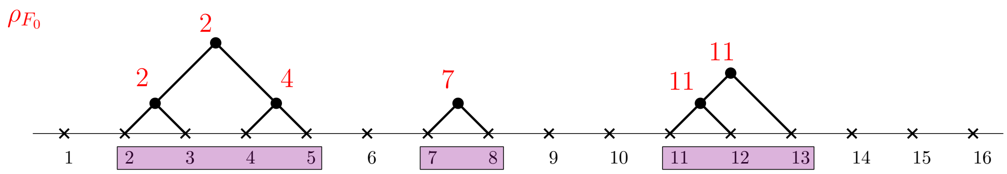

Figure 1: An indexed forest F_0 with size 6

Let $\mathrm{IN}(F)$ be the set of internal nodes of F, and let $|F|$ be its size. We call a node $v \in \mathrm{IN}(F)$ *terminal* if both its children are leaves. Let $\mathrm{Term}(F)$ denote the set of terminal nodes in F. The *support* $\mathrm{Supp}(F)$ of F is the set of labels of leaves attached to nontrivial trees. The *left support* $\mathrm{LSupp}(F)$ of F is the set of labels of left leaves attached to nontrivial trees. The indexed forest in Figure 1 has size 6, $\mathrm{Supp}(F) = \{2, 3, 4, 5, 7, 8, 11, 12, 13\}$, and $\mathrm{LSupp}(F) = \{2, 4, 7, 11\}$. Furthermore, it has 4 terminal nodes.

Define ρ_F on $\mathrm{IN}(F)$ by $\rho_F(v) :=$ the leaf label reached by following left edges down from v; see red labels in Figure 1. Define c_i to be the number of nodes labeled i by ρ_F to get a vector $c(F) = (c_1, c_2, \dots) \in \mathrm{Codes}$.

Lemma 5. *$F \mapsto c(F)$ is a bijection between indexed forests and* Codes*.*

For F in Figure 1 we get $c(F) = (0, 2, 0, 1, 0, 0, 1, 0, 0, 0, 2, 0, 0, 0, 0, \dots)$.

Words. Given any alphabet $\mathcal{A}$, we denote the set of words in $\mathcal{A}$ by $\mathcal{A}^*$. We denote the *length* of any word $w \in \mathcal{A}^*$ by $\ell(w)$. We let ϵ denote the *empty word*. The set of *injective words* in $\mathcal{A}^*$ consists of words comprising solely distinct letters. We denote this set by $\mathrm{Inj}(A)$. Our ordered alphabet of interest is obtained by "augmenting" $\mathbb{Z}$: we define the alphabet $\overline{\mathbb{Z}}$ of letters $i^{[j]}$ where $i \in \mathbb{Z}$ and $j \in \mathbb{Z}_{>0}$. These letters extend the linear order on $\mathbb{Z}$ by $i < i^{[1]} < i^{[2]} < i^{[3]} < \cdots < i+1$ for all i.

3 Forest polynomials

We are ready to introduce forest polynomials. Given the bijection $F \mapsto c(F)$, we can index them by either $\mathbb{N}$-vectors or by indexed forests. We will do this without prior warning, omitting parentheses, commas, and trailing zeros in writing c.

Definition 6. *The* forest polynomial $\mathfrak{p}_F \in \mathbb{Z}[x_1, x_2, \ldots]$ *is defined as*

$$\mathfrak{p}_F = \sum_{\kappa} \prod_{v \in \mathrm{IN}(F)} x_{\kappa(v)}$$

where the sum is over all labelings $\kappa : \mathrm{IN}(F) \to \mathbb{Z}_{>0}$ *that are bounded above by* ρ_F*, weakly increasing down left edges and strictly increasing down right edges.*

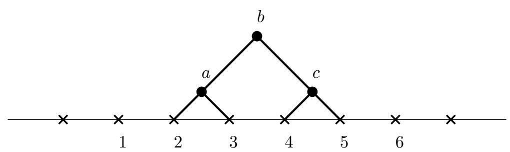

Figure 2: An indexed tree T_0 with $c(T_0) = (0, 2, 0, 1, 0, \ldots)$

Example 7. *Consider the indexed tree* T_0 *in Figure 2. Then*

$$\mathfrak{p}_{T_0} = \sum_{\substack{2 \geq a \geq b \\ 4 \geq c > b}} x_a x_b x_c = x_2^2 x_4 + x_1 x_2 x_4 + x_1^2 x_4 + x_2^2 x_3 + x_1 x_2 x_3 + x_1^2 x_3 + x_1^2 x_2 + x_1 x_2^2. \quad (3.1)$$

which coincides with the Schubert polynomial $\mathfrak{S}_{14253}$*, cf.* (2.2)*. Note that in general* $\mathfrak{p}_F$ *is the product of* $\mathfrak{p}_T$ *over all indexed trees* T *composing* F*. Thus for* F_0 *in Figure 1, we get* $\mathfrak{p}_{F_0} = \mathfrak{p}_{T_0}(\sum_{i=1}^{7} x_i)(\sum_{1 \leq i \leq j \leq 11} x_i x_j)$*.*

Example 8. (Fundamental quasisymmetric polynomials L_α are forest polynomials) *Let* F *be an indexed forest with* $\mathsf{LSupp} = [a, b]$ *for some* $a \leq b$*. A moment's thought should convince the reader that such an* F *must in fact be a "linear tree". In fact this condition is equivalent to* $\mathrm{Supp}(c(F))$ *being an interval in* $\mathbb{Z}_{>0}$*. Figure 3 shows two instances. The one on the left gives* $\mathfrak{p}_F = \mathsf{L}_{121}(x_1, \ldots, x_4)$ *and that on the right* $\mathfrak{p}_F = \mathsf{L}_2(x_1, x_2)$*. In general, let* $c(F) = (c_1, c_2, \ldots)$*. Then* $\mathfrak{p}_F = \mathsf{L}_\alpha(x_1, \ldots, x_b)$ *with* $\alpha = (c_b, \ldots, c_a)$*.*

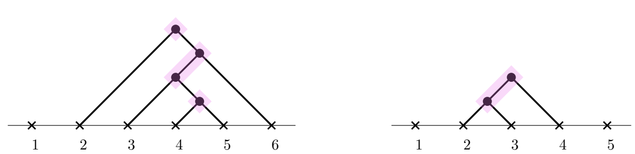

Figure 3: Two indexed forests with left support an interval

The next proposition is a consequence of the observation that the revlex leading term in $\mathfrak{p}_F$ is $\mathbf{x}^{c(F)}$. It is interesting to compare it with the fact that Schubert polynomials indexed by permutations with last descent in position at most n (respectively, in $\mathbb{S}_n$) form a basis of $\mathbb{Z}[\mathbf{x}_n]$ (respectively, of $\mathbb{Q}\{\mathbf{x}^c \mid c_i \leq n - i\}$).

Proposition 9. *Given $n \in \mathbb{Z}_{>0}$, the set $\{\mathfrak{p}_F \mid \mathrm{LSupp}(F) \subseteq [n]\}$ is a basis of $\mathbb{Z}[\mathbf{x}_n]$, while the set $\{\mathfrak{p}_F \mid \mathrm{Supp}(F) \subseteq [n]\}$ is a basis of $\mathbb{Q}\{\mathbf{x}^c \mid c \in \mathcal{C}_n\}$.*

Definition 6 essentially describes forest polynomials as a generating function for certain *flagged* (P, ω, ρ)*-partitions* in the sense of Assaf–Bergeron [2]. Let $\mathrm{Dec}(F)$ denote the set of *decreasing forests* with shape F, that is bijections $\mathrm{IN}(F) \to \{1, \ldots, |F|\}$ decreasing down branches. For a fixed F, one can identify $L \in \mathrm{Dec}(F)$ with a word $\rho_F(v_1) \cdots \rho_F(v_{|F|})$ where v_i is the node with label $L(v_i) = i$. The next expansion is a special case of results in *loc. cit.*

$$\mathfrak{p}_F = \sum_{L \in \mathrm{Dec}(F)} \mathfrak{F}(L). \tag{3.2}$$

Thus any forest polynomial is a sum of slide polynomials. For T_0 in Example 7 and Figure 2, the expansion in (3.2) specializes to $\mathfrak{F}(242) + \mathfrak{F}(422) = \mathfrak{F}_{(1,2)} + \mathfrak{F}_{(0,2,0,1)}$, which as expected coincides with the one defining $\mathfrak{S}_{14253} (= \mathfrak{p}_{T_0})$.

Reduction modulo QSym_n^+. Henceforth, given $f \in \mathbb{Q}[\mathbf{x}_n]$, we take $f \bmod \mathrm{QSym}_n^+$ to mean the unique g satisfying $f = g + h$ with $h \in \mathrm{QSym}_n^+$ and $g \in \mathbb{Q}\{\mathbf{x}^c \mid c \in \mathcal{C}_n\}$. Our aim is to give a straightforward description for $\mathfrak{p}_F \bmod \mathrm{QSym}_n^+$; see Theorem 12. Despite the simplicity of the statement, the proof is technical (and omitted in this abstract). It relies on the recurrence in Lemma 10.

We discuss two operations on indexed forests that will allow us to describe this recurrence. The first is the *shift* map τ which, given an indexed forest F, simply shifts $\mathrm{Supp}(F)$ one unit to the right. We abuse notation and define $\tau^i(\mathfrak{p}_F) = \mathfrak{p}_{\tau^i(F)}$ for $i \in \mathbb{Z}$, making note that if $\mathrm{Supp}(\tau^i(F)) \not\subseteq \mathbb{Z}_{>0}$, then $\mathfrak{p}_{\tau^i(F)}$ equals 0.

The second operation is *trimming*, which corresponds to removing a terminal node. Explicitly, consider $v \in \mathrm{Term}(F)$. It lies in one of the indexed trees T composing F, say with support $[a, b]$. We obtain $\mathrm{trim}_v(F)$ from F by replacing T with $T' = T \setminus \{v\}$ with new support $[a, b-1]$, and leaving the other trees unaltered. If $a = b - 1$, then T' has no internal nodes (i.e it is trivial), and we omit it.

Figure 4: The two indexed forests resulting from trimming T_0 in Figure 2

Lemma 10. *Given an indexed forest F we have*

$$\mathfrak{p}_F = \mathfrak{p}_{\tau^{-1}(F)} + \sum_{v \in \mathrm{Term}(F)} x_{\rho_F(v)}\, \mathfrak{p}_{\mathrm{trim}_v(F)}. \tag{3.3}$$

One can rewrite this recurrence by indexing forest polynomials by $\mathbb{N}$-vectors. For the tree T_0 in Figure 2, using Lemma 10 we get: $\mathfrak{p}_{0201} = \mathfrak{p}_{201} + x_2\,\mathfrak{p}_{011} + x_4\,\mathfrak{p}_{02}$. Now, $\mathfrak{p}_{201} = x_1^2x_2 + x_1^2x_3$, $\mathfrak{p}_{011} = e_2(x_1,x_2,x_3)$, and $\mathfrak{p}_{02} = h_2(x_1,x_2)$, and substituting these gives the same expansion as in (3.1).

Lemma 10 has several consequences; we record two that concern us most. First, we have a nice interpretation for the sum of coefficients of $\mathfrak{p}_F$ in a special case.

Proposition 11. *If T is an indexed tree with support $\{1,\dots,|T|\}$, we have*

$$\mathfrak{p}_T(1,\dots,1) = |\mathrm{Dec}(T)|. \tag{3.4}$$

One can extend this to a result for all $\mathfrak{p}_F$ but we will not need it here. The second consequence of Lemma 10 is Theorem 4*(ii)* which we reformulate here.

Theorem 12. *Let F be an indexed forest satisfying $\mathsf{LSupp}(F) \subseteq [n]$. Then*

$$\mathfrak{p}_F \mod \mathrm{QSym}_n^+ = \begin{cases} 0 & \text{if } \mathrm{Supp}(F) \not\subseteq [n] \text{ (equivalently, } c(F) \notin \mathcal{C}_n) \\ \mathfrak{p}_F & \text{if } \mathrm{Supp}(F) \subseteq [n] \text{ (equivalently, } c(F) \in \mathcal{C}_n). \end{cases}$$

As an example, note that Theorem 12 says that the forest polynomial $\mathfrak{p}_{0201}$ in (3.1) is in QSym_4^+ as $\mathrm{Supp}(T_0) \not\subseteq [4]$. We invite the reader to check that

$$\mathfrak{p}_{0201} = \mathsf{L}_{(2,1)}(x_1,\dots,x_4) + \mathsf{L}_{(1,2)}(x_1,\dots,x_4) - (x_3+x_4)\mathsf{L}_{(1,1)}(x_1,\dots,x_4)$$

and thus clearly belongs in QSym_4^+. On the other hand, $\tau^{-1}(T_0)$ has support exactly $[4]$, and $\mathfrak{p}_{2010} = x_1^2x_2 + x_1^2x_3$; each monomial here is an ABB monomial.

4 Word-Forest correspondence

We now proceed to describe the main insertion algorithm in this work, the *Word-Forest correspondence*. We will need labeled versions of indexed forests. A *labeling* of F is a map from $\mathrm{IN}(F)$ to some alphabet $\mathcal{A}$. *Furthermore, we allow our indexed forests to be supported on $\mathbb{Z}$ and not just $\mathbb{Z}_{>0}$. When we consider forest polynomials subsequently, only those that are supported on $\mathbb{Z}_{>0}$ shall matter.*

The input is an injective word $W \in \mathsf{Inj}(\overline{\mathbb{Z}})$, while the output is a pair $\mathcal{WF}(W) = (\mathsf{P},\mathsf{Q})$ of labeled indexed forests where the underlying unlabeled forests are the same. Following conventions, we call P (respectively Q) the *insertion forest* (respectively *recording forest*). We define $\mathcal{WF}(W)$ inductively via an insertion procedure.

A convenient bookkeeping device while executing our algorithm is the *root list* $\mathsf{rl}(W)$ for W. It is a totally ordered subset of $\mathbb{Z} \sqcup \overline{\mathbb{Z}}$ that records the list of root labels in P: trivial trees are encoded by their index in $\mathbb{Z}$, other root labels are certain letters in W.

The WF-correspondence. Let $W \in \mathsf{Inj}(\overline{\mathbb{Z}})$. If W is empty, P and Q are trivial forests. We thus set $\mathsf{rl}(\epsilon) = \mathbb{Z}$. Now suppose $W = W'b$ and $\mathcal{WF}(W') = (\mathsf{P}',\mathsf{Q}')$ by induction. Then $\mathcal{WF}(W)$ is determined as follows.

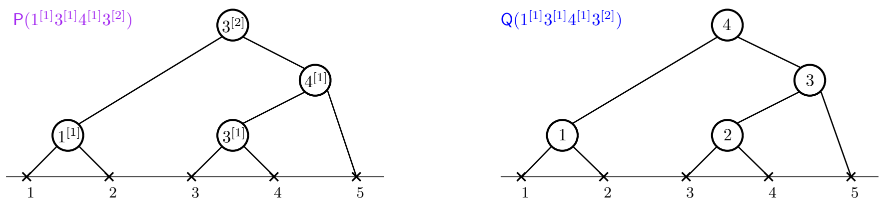

Figure 5: The insertion and recording forest for $W = 1^{[1]}3^{[1]}4^{[1]}3^{[2]}$

1. Let a (respectively c) be the greatest (respectively least) elements in $\mathsf{rl}(W')$ such that $a < b < c$.[3]
2. Merge the two trees in P' with roots labeled by a and c into a single tree by adding a new root with label b as their parent. This gives us $\mathsf{P}(W)$. The new root list is then $\mathsf{rl}(W) = (\mathsf{rl}(W') \setminus \{a, c\}) \cup \{b\}$.
3. $\mathsf{Q}(W)$ records this merging procedure by performing a corresponding merge in Q' and having $|W|$ as the new root label.

See Figure 5 for the P and Q that result when one applies the preceding procedure. Note that internal nodes in P are labeled by letters in W whereas those in Q are labeled by $\{1, \ldots, |W|\}$. One can say more— P is a *local binary search forest*, i.e. the label of any node is strictly greater (lesser) than the label of its left (respectively right) child, while Q is a *decreasing forest* as already defined.

Fix a finite set $B \subseteq \overline{\mathbb{Z}}$. Denote by $\mathbb{S}_B$ the set of all permutations of B (in one-line notation). Clearly $|\mathbb{S}_B| = |B|!$ and $\mathbb{S}_B \subset \mathrm{Inj}(\overline{\mathbb{Z}})$. Denote by $\mathcal{PF}(B)$ the set of all ordered pairs (P, Q) where P,Q have the same underlying complete indexed forest and additionally

- internal nodes in P are bijectively labeled by letters in B so that the result is a local binary search forest;
- internal nodes in Q are bijectively labeled by letters in $\{1, \ldots, |B|\}$ so that the result is a decreasing forest.

We are ready for our main procedure that governs the combinatorics.

Theorem 13 (Word-Forest correspondence)**.** *The map $\mathcal{WF} : \mathbb{S}_B \to \mathcal{PF}(B)$ is a bijection.*

Proof (sketch). The forward direction produces an element in $\mathcal{PF}(B)$ as discussed earlier. The inverse correspondence simply consists in reading the labels of P according to the order determined by Q, and this is injective as well. The claim follows. □

The WF-correspondence specializes to the "Sylvester correspondence" when B has the form $\{1^{[1]}, 2^{[1]}, \ldots, n^{[1]}\}$, but does not reduce to it in general. Our main use for this correspondence will be to group together all words that have the same labeled forest P.

The WF-equivalence. We say that words $W_1, W_2 \in \mathrm{Inj}(\overline{\mathbb{Z}})$ are WF-*equivalent* if $\mathsf{P}(W_1) = \mathsf{P}(W_2)$. We denote this by $W_1 \underset{\mathrm{WF}}{\equiv} W_2$. For instance the reader can check that $3^{[1]}1^{[1]}4^{[1]}3^{[2]}$

[3]Recall that our inputs are injective words

and $1^{[1]}3^{[1]}4^{[1]}3^{[2]}$ produce the same P (shown in Figure 5), and thus $3^{[1]}1^{[1]}4^{[1]}3^{[2]} \underset{\mathrm{WF}}{\equiv} 1^{[1]}3^{[1]}4^{[1]}3^{[2]}$. It is helpful to compare the Q symbol for these two words as well. Observe that two *incomparable* nodes swap their labels. We can actually give a 'local' description to WF-equivalence which generates $\underset{\mathrm{WF}}{\equiv}$ as its transitive closure. We say that $W_1 \underset{\mathrm{WF}}{\sim} W_2$ if we can write $W_1 = UabV$, $W_2 = UbaV$ and there are at least two elements c, d of in the rootlist $\mathsf{rl}(U)$ such that

$$\min(a, b) < c, d < \max(a, b). \tag{4.1}$$

Note that this entails that W_1 and W_2 are in the same *commutation class*: if we denote $a = i^{[\cdot]}, b = j^{[\cdot]}$, then necessarily $|j - i| \geq 2$.

Lemma 14. *$\underset{\mathrm{WF}}{\equiv}$ is the equivalence relation generated by $\underset{\mathrm{WF}}{\sim}$.*

5 a_w by expanding Schuberts into forests

We now discuss the decomposition of Schubert polynomials into forest polynomials. To this end we will need to apply the WF-correspondence to reduced words. We thus need to transform words in $\mathbb{Z}^*$ as as elements in $\mathrm{Inj}(\overline{\mathbb{Z}})$.

For $w \in \mathbb{Z}^*$, define the word $\mathrm{stan}(w)$ in $\overline{\mathbb{Z}}^*$ by labeling the occurrences of the letter i in w from left to right by $i^{[1]}, i^{[2]}, \ldots,$ for any $i \in \mathbb{Z}$. For instance $\mathrm{stan}(1221625) = 1^{[1]}2^{[1]}2^{[2]}1^{[2]}6^{[1]}2^{[3]}5^{[1]}$. It is clear $w \mapsto \mathrm{stan}(w)$ embeds $\mathbb{Z}^* \in \mathrm{Inj}(\overline{\mathbb{Z}})$. We will implicitly apply constructions of the previous section to words via these standardized versions.

Definition 15. *Given a WF-equivalence class C, we define its* shape $\mathrm{sh}(\mathsf{C})$ *to be the underlying unlabeled indexed forest F of $\mathsf{P}(W)$ for any element in $W \in \mathsf{C}$. We define* $\mathrm{Supp}(\mathsf{C})$ *to be equal to* $\mathrm{Supp}(F)$.

Recall that $\mathrm{Red}(w)$ is the set of reduced words for a permutation w. For instance, if $w = 21543$ in one-line notation, then

$$\mathrm{Red}(w) = \{1343, 3143, 3413, 3431, 1434, 4134, 4314, 4341\}. \tag{5.1}$$

Those with the same color form an equivalence class under $\underset{\mathrm{WF}}{\equiv}$. The corresponding shapes from left to right are shown in Figure 6. In particular we note that $\mathrm{Red}(21543)$ is closed under $\underset{\mathrm{WF}}{\equiv}$. This is in fact true for any set $\mathrm{Red}(w)$: more generally, commutation classes are always closed under $\underset{\mathrm{WF}}{\equiv}$ by Lemma 14 and the remark before it.

We are ready to make Theorem 4(i) explicit.

Theorem 16. *Fix a permutation w. Let $\mathrm{Red}(w^{-1})/\underset{\mathrm{WF}}{\equiv}$ denote the set of WF-equivalence classes decomposing $\mathrm{Red}(w^{-1})$, and let C_1 through C_k be those satisfying $\mathrm{Supp}(\mathsf{C}_i) \subset \mathbb{Z}_{>0}$. Then*

$$\mathfrak{S}_w = \sum_{1 \leq i \leq k} \mathfrak{p}_{\mathrm{sh}(\mathsf{C}_i)}. \tag{5.2}$$

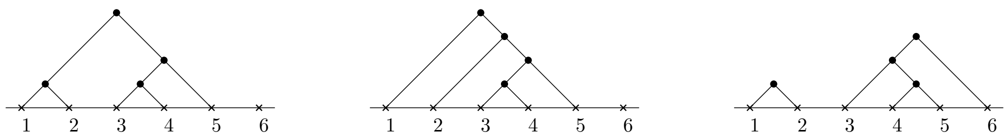

Figure 6: Shapes of equivalence classes for $\mathrm{Red}(21543)$

Proof (sketch). By definition of $\mathfrak{S}_w$, it equals the sum over $\mathsf{C} \in \mathrm{Red}(w^{-1})/_{\underset{\mathrm{WF}}{\equiv}}$ of the polynomials $\mathfrak{p}(\mathsf{C}) := \sum_{W \in \mathsf{C}} \mathfrak{F}(W)$. It is easy to show that all terms $\mathfrak{F}(W)$ vanish in $\mathfrak{p}(\mathsf{C})$ when $\mathrm{Supp}(\mathsf{C}) \nsubseteq \mathbb{Z}_{>0}$. Then a key lemma is to show that one always has $\mathfrak{p}(\mathsf{C}) = \mathfrak{p}_{\mathrm{sh}(\mathsf{C})}$, which we omit in this abstract. □

We now return to our motivating question— describing a_w combinatorially via Theorem 1. We call a WF-equivalence class C a *tree class* if $\mathrm{sh}(\mathsf{C})$ is an indexed tree T supported on an initial interval $[n]$ for some n. Going back to Figure 6, we see that $\{1343, 3143, 3413\}$ and $\{3431\}$ are tree classes, whereas $\{1434, 4134, 4314, 4341\}$ is not.

Proof of Theorem 1. Theorem 16 gives us the expansion for $\mathfrak{S}_w$ into forest polynomials. Here w is a permutation in S_n of length $n-1$, which implies that any forest polynomial $\mathfrak{p}_F$ that appears in (5.2) necessarily satisfies $|F| = n-1$ and $\mathrm{LSupp}(F) \subseteq [n]$. It follows from Theorem 12 that only forest polynomials indexed by trees with support $[n]$ survive when reducing modulo QSym_n^+. We thus obtain

$$\mathfrak{S}_w \mod \mathrm{QSym}_n^+ = \sum_{\mathsf{C}} \mathfrak{p}_{\mathrm{sh}(\mathsf{C})},$$

where C runs through tree classes in $\mathrm{Red}(w^{-1})/_{\underset{\mathrm{WF}}{\equiv}}$. Now by Theorem 3 we know that

$$a_w = \sum_{\mathsf{C}} \mathfrak{p}_{\mathrm{sh}(\mathsf{C})}(1, \ldots, 1).$$

Proposition 11 says that each summand $\mathfrak{p}_{\mathrm{sh}(\mathsf{C})}(1, \ldots, 1)$ equals $|\mathsf{C}|$, so we have that a_w is the number of words in $\mathrm{Red}(w^{-1})$ belonging to a tree class.

To conclude, one needs to compare the recursive descriptions of the WF-parking procedure and the WF-correspondence. A careful analysis shows that the former is but a "shadow" of the latter, and the set $I(v_1 \cdots v_k)$ at the end of the WF-parking procedure is exactly the common support of $\mathsf{P}(v_1 \cdots v_k)$ and $\mathsf{Q}(v_1 \cdots v_k)$.

In particular a word in $\mathbb{Z}^*$ belongs to a tree class if and only if it is a WF-parking word, which completes the proof. □

6 Other applications

We have defined several notions in this work in order to solve a particular problem, but these notions are of independent interest. We close with a non-exhaustive list of some notable properties.

- The strategy developed to determine a_w can be used to compute the DS of other polynomials. It can for instance give automatically a combinatorial interpretation for the remixed Eulerian numbers introduced in [7].
- Forest polynomials enjoy a positive multiplication rule which can be determined combinatorially via the WF-correspondence.
- The WF-correspondence and WF-parking deserve further study. For instance, while the combinatorial interpretation in Theorem 1 implies several of the results from [8], it is not obvious to the authors why $a_w = a_{w^{-1}}$, though this is indeed known to hold [*loc. cit.*, Proposition 5.1.1].

References

[1] D. Anderson and J. Tymoczko. "Schubert polynomials and classes of Hessenberg varieties". *J. Algebra* **323**.10 (2010), pp. 2605–2623.

[2] S. Assaf and N. Bergeron. "Flagged $(\mathcal{P}, \rho)$-partitions". *European J. Combin.* **86** (2020), pp. 103085, 17. DOI.

[3] S. Assaf and D. Searles. "Schubert polynomials, slide polynomials, Stanley symmetric functions and quasi-Yamanouchi pipe dreams". *Adv. Math.* **306** (2017), pp. 89–122. DOI.

[4] J.-C. Aval, F. Bergeron, and N. Bergeron. "Ideals of quasi-symmetric functions and super-covariant polynomials for S_n". *Adv. Math.* **181**.2 (2004), pp. 353–367.

[5] S. C. Billey, W. Jockusch, and R. P. Stanley. "Some combinatorial properties of Schubert polynomials". *J. Algebraic Combin.* **2**.4 (1993), pp. 345–374.

[6] A. A. Klyachko. "Orbits of a maximal torus on a flag space". *Funktsional. Anal. i Prilozhen.* **19**.1 (1985), pp. 77–78.

[7] P. Nadeau and V. Tewari. "A q-analogue of an algebra of Klyachko and Macdonald's reduced word identity". 2021. arXiv:2106.03828.

[8] P. Nadeau and V. Tewari. "The permutahedral variety, mixed Eulerian numbers, and principal specializations of Schubert polynomials". *Int. Math. Res. Not. IMRN* (2021, to appear).

[9] P. Nadeau and V. Tewari. "Divided symmetrization and quasisymmetric functions". *Selecta Math. (N.S.)* **27**.4 (2021), Paper No. 76, 24 pp. DOI.

[10] A. Postnikov. "Permutohedra, associahedra, and beyond". *Int. Math. Res. Not. IMRN* 6 (2009), pp. 1026–1106.