A GENERALIZATION OF EDELMAN–GREENE INSERTION FOR SCHUBERT POLYNOMIALS

SAMi ASSAF

ABSTRACT. Edelman and Greene generalized the Robinson–Schensted–Knuth correspondence to reduced words in order to give a bijective proof of the Schur positivity of Stanley symmetric functions. Stanley symmetric functions may be regarded as the stable limits of Schubert polynomials, and similarly Schur functions may be regarded as the stable limits of Demazure characters for the general linear group. We modify the Edelman–Greene correspondence to give an analogous, explicit formula for the Demazure character expansion of Schubert polynomials. Our techniques utilize dual equivalence and its polynomial variation, but here we demonstrate how to extract explicit formulas from that machinery which may be applied to other positivity problems as well.

1. INTRODUCTION

Schur functions, the ubiquitous basis for symmetric functions with deep connections to representation theory and geometry, may be regarded as the generating functions for standard Young tableaux. In an analogous way, Stanley defined a generating function for reduced words that he proved was symmetric and conjectured was Schur positive. Edelman and Greene established a bijective correspondence between reduced words and ordered pairs of Young tableaux of the same partition shape such that the left is increasing with reduced reading word and the right is standard. Thus through this correspondence they proved Stanley’s conjecture and, moreover, gave an explicit formula for the Schur expansion as the number of such left tableaux that can appear in the correspondence.

Schubert polynomials were introduced by Lascoux and Schützenberger as polynomial representatives of Schubert classes for the cohomology of the flag manifold with nice algebraic and combinatorial properties. They can be defined as the generating polynomials of reduced words, and in the stable limit, they become the Stanley symmetric functions. Parallel to this, Demazure characters for the general linear group can be regarded as the generating polynomials for standard key tableaux, and in the stable limit, the key tableaux become Young tableaux and the Demazure characters become Schur functions. Lascoux and Schützenberger noticed that the Schubert polynomials expand nonnegatively into Demazure characters parallel to the nonnegative expansion of Stanley symmetric functions into Schur functions. The proof uses the same structure of partitioning reduced words into equivalence classes based on the Edelman–Greene right tableau, yet there is no direct formula for the coefficients given.

2010 Mathematics Subject Classification. Primary 05A05; Secondary 05A15, 05A19, 14N15.
Key words and phrases. Schubert polynomials, Demazure characters, key polynomials, RSK, Edelman–Greene insertion, reduced words.

Work supported in part by NSF DMS-1763336.
In this paper, we complete the analogy between the function and polynomial settings by providing a new bijective correspondence between reduced words and ordered pairs of key tableaux of the same weak composition shape such that the left is Yamanouchi with reduced reading word and the right is standard. Thus through this correspondence we prove the Demazure positivity of Schubert polynomials and, moreover, give an explicit formula for the Demazure expansion as the number of such left tableaux that can appear in the correspondence.

Our real purpose, in addition to this explicit result, is to provide a framework by which one can extract explicit Schur expansions of symmetric functions through the machinery of dual equivalence and explicit Demazure expansions of polynomials through the machinery of weak dual equivalence. Dual equivalence and its weak variant give universal methods for proving positivity results, but they do so indirectly without giving tractable formulas. Our techniques in this paper utilize dual equivalence to show how it can manipulated to give the desired formulas.

This paper is structured as follows. We begin in Section 2 with a review of definitions for reduced words for permutations. We develop parallel theories of the generating functions for reduced words, reviewing Stanley symmetric functions \[21\] in Section 2.1 and generating polynomials for reduced words, reviewing Schubert polynomials \[14\] in Section 2.2.

Questions of positivity arise in Section 3, where we consider the Coxeter–Knuth equivalence relations \[10\] on reduced words. Maintaining our parallel study, in Section 3.1, we review the machinery of dual equivalence \[4\] to see that the generating function of a Coxeter–Knuth equivalence class on reduced words is a Schur function, thus recovering the Schur positivity result of Edelman and Greene for Stanley symmetric functions \[10\]. In Section 3.2 we review the machinery of weak dual equivalence \[6\] to see that the generating polynomial of a Coxeter–Knuth equivalence class on reduced words is a Demazure character, thus recovering the Demazure positivity result of Lascoux and Schützenberger for Schubert polynomials \[15, 18\].

In Section 4, we embark on the quest to extract explicit formulas for these expansions by finding canonical representatives for the Coxeter–Knuth equivalence classes. In Section 4.1, we recover the explicit formula of Edelman and Greene for the Schur expansion Stanley symmetric functions \[10\] using simple techniques that avoid the subtlety of their insertion algorithm. In Section 4.2 we use similar techniques to arrive at our main result: an explicit combinatorial formula for the Demazure expansion of Schubert polynomials.

Finally, we return to the inspiration of this work in Section 5, where we present explicit insertion algorithms. In Section 5.1 we review the Edelman–Greene correspondence that associates to each reduced word a pair Young tableaux, and then in Section 5.2 we use results from Section 4 to modify this correspondence to associate to each reduced word a pair of key tableaux. In this way, we complete the parallel stories with satisfactory formulas for both cases.

### 2. Generating functions for reduced words

The symmetric group \(\mathcal{S}_n\) has generators \(s_i\), the simple transpositions interchanging \(i\) and \(i+1\), and relations \(s_i^2 = 1\) is the identity, \(s_is_j = s_js_i\) for \(|i-j| \geq 2\), and \(s_is_{i+1}s_i = s_is_{i+1}s_is_{i+1}\) for \(1 \leq i \leq n-2\).

An expression for a permutation \(w \in \mathcal{S}_n\) is a way of writing \(w\) in terms of these simple generators, i.e. \(w = s_{p_1} \cdots s_{p_k}\). The length of \(w\), denoted by \(\ell(w)\) is the
number of pairs \((i < j)\) such that \(w_i > w_j\). If an expression for \(w\) has exactly \(\ell(w)\) terms, then it is reduced. In this case, the sequence of indices \(\rho = (\rho_{\ell(w)}, \ldots, \rho_1)\) such that \(w = s_{\rho_{\ell(w)}} \cdots s_{\rho_1}\) is called a reduced word for \(w\).

For example, there are two reduced expressions for the permutation 321, namely \(s_1s_2s_1\) and \(s_2s_1s_2\), both of which have length 3 since there are 3 inversions in 321. Therefore we say that (1, 2, 1) and (2, 1, 2) are reduced words for 321. For a more elaborate example, Fig. 1 shows the 11 reduced words for the permutation 153264.

\[
(5, 3, 2, 3, 4) \quad (5, 2, 3, 2, 4) \quad (5, 2, 3, 4, 2) \quad (3, 5, 2, 3, 4) \quad (3, 2, 5, 3, 4) \quad (3, 2, 3, 5, 4) \\
(2, 5, 3, 4, 2) \quad (2, 3, 5, 4, 2) \quad (2, 5, 3, 2, 4) \quad (2, 3, 5, 2, 4) \quad (2, 3, 2, 5, 4)
\]

**Figure 1.** The set of reduced words for \(w = 153264\).

We consider the set \(R(w)\) of reduced words for a given permutation \(w\). Below we present two different generating functions for this set by assigning either a symmetric function or a polynomial to each reduced word.

### 2.1. Stanley symmetric functions.

Stanley [21] defined a family of symmetric functions in order to enumerate reduced words. These functions have since been realized to have important connections with geometry and representation theory.

A weak composition \(a = (a_1, \ldots, a_n)\) is a sequence of nonnegative integers. A composition \(\alpha = (\alpha_1, \ldots, \alpha_\ell)\) is a sequence of positive integers. A partition \(\lambda = (\lambda_1 \geq \cdots \geq \lambda_\ell)\) is a weakly decreasing sequence of positive integers. Given compositions \(\alpha, \beta\), we say \(\beta\) refines \(\alpha\) if there exist indices \(i_1 < \cdots < i_\ell\) such that

\[
\beta_1 + \cdots + \beta_{i_j} = \alpha_1 + \cdots + \alpha_j.
\]

For example, (1, 2, 2) refines (3, 2) but does not refine (2, 3).

Gessel introduced the *fundamental quasisymmetric functions* [11], indexed by compositions, that form an important basis for quasisymmetric functions.

**Definition 2.1.1 ([11]).** For \(\alpha\) a composition, the fundamental quasisymmetric function \(F_\alpha\) is

\[
F_\alpha(x_1, x_2, \ldots) = \sum_{\text{flat}(b) \text{ refines } \alpha} x_1^{b_1} x_2^{b_2} \cdots,
\]

where the sum is over weak compositions \(b\) for which the composition \(\text{flat}(b)\) obtained by removing all parts equal to 0 refines \(\alpha\).

For example, restricting to three variables to make the expansion finite, we have

\[
F_{(3, 2)}(x_1, x_2, x_3) = x_2^3 x_3^2 + x_1^3 x_3^2 + x_1^3 x_2^2 + x_1 x_2 x_3 + x_1 x_2^2 x_3 + x_1^2 x_2 x_3^2.
\]

Stanley [21] defined a family of symmetric functions indexed by permutations that are the fundamental quasisymmetric generating functions for reduced words. To define this, we associate a composition to each reduced word.

**Definition 2.1.2.** The *run decomposition* of a reduced word \(\rho\) partitions \(\rho\) into increasing sequences \(\rho = (\rho^{(k)} \cdots | \rho^{(1)})\) of maximal length. The *descent composition* of \(\rho\), denoted by \(\text{Des}(\rho)\), is the composition \((|\rho^{(1)}|, \ldots, |\rho^{(k)}|)\).
For example, $\rho = (3, 6, 4, 7, 5, 2, 4)$ and $\sigma = (6, 7, 3, 4, 5, 2, 4)$, two reduced words for $w = 13625847$, have run decompositions

\[
\begin{align*}
\rho^{(4)} & = (3, 6) & \rho^{(3)} & = (3, 4) & \rho^{(2)} & = (5) & \rho^{(1)} & = (2, 4) \\
\sigma^{(3)} & = (6, 7) & \sigma^{(2)} & = (3, 4, 5) & \sigma^{(1)} & = (2, 4),
\end{align*}
\]

giving $\text{Des}(\rho) = (2, 1, 2, 2)$ and $\text{Des}(\sigma) = (2, 3, 2)$. Note the reversal of lengths.

We may visualize the descent composition via positive integer fillings of cell diagrams as follows.

**Definition 2.1.3.** The *descent tableau* of a reduced word $\rho$, denoted by $\mathbb{D}(\rho)$, is the filling of unit cells in the first quadrant constructed as follows. Place $\rho_{\ell(w)}$ into the first column of row $|\text{Des}(\rho)|$. For $i = \ell(w) - 1, \ldots, 2, 1$, place $\rho_i$ immediately right of $\rho_{i+1}$ if $\rho_{i+1} < \rho_i$; otherwise in the first column of the next row down.

For example, $\rho = (3, 6, 4, 7, 5, 2, 4)$ is inserted as shown on the left side of Figure 2 and $\sigma = (6, 7, 3, 4, 5, 2, 4)$ inserts as shown on the right side of Figure 2.

![Figure 2](image)

**Figure 2.** Constructing the descent tableaux for reduced words $\rho = (3, 6, 4, 7, 5, 2, 4)$ (left) and $\sigma = (6, 7, 3, 4, 5, 2, 4)$ (right).

By construction, rows of $\mathbb{D}(\rho)$ are increasing and the descent composition for $\rho$ is given by the lengths of the rows of the descent tableau for $\rho$, read bottom to top.

**Definition 2.1.4 ([21]).** For $w$ a permutation, the *Stanley symmetric function* $S_w$ is

\[
S_w(X) = \sum_{\rho \in R(w)} F_{\text{Des}(\rho)}(X).
\]

For example, from Fig. 1 we compute

\[
S_{135264}(X) = F_{(3,1,1)}(X) + 2F_{(2,2,1)}(X) + 2F_{(1,3,1)}(X) + F_{(3,2)}(X) + 2F_{(1,1,3)}(X) + F_{(2,1,2)}(X) + F_{(2,3)}(X).
\]

2.2. **Schubert polynomials.** Lascoux and Schützenberger [14] introduced *Schubert polynomials* as a basis for the polynomial ring that gives polynomial representatives of Schubert classes for the cohomology of the flag manifold with nice algebraic and combinatorial properties.

Assaf and Searles [1] introduced *fundamental slide generating polynomials* as a generalization of the fundamental quasisymmetric functions that form a basis for the full polynomial ring. They showed that Schubert polynomials are the fundamental slide generating polynomials for reduced words.
Definition 2.2.1. For a weak composition \( \rho \) of length \( n \), the fundamental slide polynomial \( \mathcal{F}_\rho = \mathcal{F}_\rho(x_1, \ldots, x_n) \) is
\[
\mathcal{F}_\rho = \sum_{b_1 + \cdots + b_k \geq a_1 + \cdots + a_k \forall k \text{ flat}(b)} x_1^{b_1} \cdots x_n^{b_n},
\]
where \( \text{flat}(a) \) denotes the composition obtained by removing all zero parts.

For example, we have
\[
\mathcal{F}_{(3,0,2)}(x_1, x_2, x_3) = x_1^3 x_2^2 + x_1^3 x_2 x_3,
\]
which is not equal to \( F_{(3,2)}(x_1, x_2, x_3) \) computed earlier. However, we do have \( \mathcal{F}_{(3,2)} = F_{(3,2)}(x_1, x_2, x_3) \). Moreover, Assaf and Searles proved fundamental slide polynomials stabilize \( \mathcal{F}(Theorem 4.5). \)

Proposition 2.2.2. For a weak composition \( \rho \), we have
\[
\lim_{m \to \infty} \mathcal{F}_{0^m \times \rho}(x_1, \ldots, x_m, 0, \ldots, 0) = F_{\text{flat}(\rho)}(X),
\]
where \( 0^m \times \rho \) is the weak composition obtained by prepending \( m \)'s to \( \rho \).

We generalize the descent composition of a reduced word to a weak composition as defined in \( \mathcal{F}(Definition 3.2). \)

Definition 2.2.3. For a reduced word \( \rho \), define the weak descent composition of \( \rho \), denoted by \( \text{des}(\rho) \), as follows. Let \( (\rho^{(k)}) \cdots (\rho^{(1)}) \) be the run decomposition of \( \rho \), that is, each \( \rho^{(i)} \) is increasing and as long as possible. Set \( r_i = \min(\rho^{(i)}) \) for \( i = 1, \ldots, k \). Set \( \hat{r}_k = r_k \), and for \( i < k \), set \( \hat{r}_i = \min(r_i, \hat{r}_{i+1} - 1) \). If \( r_1 \leq 0 \), then define \( \text{des}(\rho) = \emptyset \); otherwise, set the \( \hat{r}_i \)th part of \( \text{des}(\rho) \) to be \( \text{des}(\rho)_{\hat{r}_i} = |\rho^{(i)}| \) and set all other parts to 0.

Revisiting our previous examples, for \( \rho = (3, 6 | 4, 7 | 5 | 2, 4) \), we have \( r_4 = 3 \), \( \hat{r}_3 = \min(4, 3 - 1) = 2 \), \( \hat{r}_2 = \min(5, 2 - 1) = 1 \), and finally \( \hat{r}_1 = \min(2, 1 - 1) = 0 \) giving \( \text{des}(\rho) = \emptyset \) since \( r_1 = 0 \). For \( \sigma = (6, 7 | 3, 4, 5 | 2, 4) \), we have \( s_3 = 6 \), \( s_2 = \min(3, 6 - 1) = 3 \), and \( \hat{s}_2 = \min(2, 3 - 1) = 2 \), giving \( \text{des}(\sigma) = (0, 2, 3, 0, 0, 2) \).

We may visualize Definition 2.2.3 via a simple insertion algorithm as follows.

Definition 2.2.4. The weak descent tableau of a reduced word \( \rho \), denoted by \( \mathcal{D}(\rho) \), is the following filling of unit cells in the right half plane. Place \( \rho_{\ell(w)} \) into the first column of row \( \rho_{\ell(w)} \). For \( i = \ell(w) - 1, \ldots, 2, 1 \), place \( \rho_i \) immediately right of \( \rho_{i+1} \) if \( \rho_{i+1} < \rho_i \), or in the first column of the lower of row \( \rho_i \) or the row below \( \rho_{i+1} \).

For example, \( \rho = (3, 6, 4, 7, 5, 2, 4) \) is inserted as shown on the left side of Figure 8 and \( \sigma = (6, 7, 3, 4, 5, 2, 4) \) inserts as shown on the right side of Figure 8.

Notice \( \text{des}(\rho) = \emptyset \) if and only if there is an occupied row of \( \mathcal{D}(\rho) \) with non-positive index, and otherwise \( \text{des}(\rho)_i \) is the number of entries in row \( i \) of \( \mathcal{D}(\rho) \).

We say that \( \rho \) is virtual if \( \text{des}(\rho) = \emptyset \). To facilitate virtual objects, set
\[
\mathcal{F}_\emptyset = 0.
\]

Billey, Jockusch, and Stanley \( \mathcal{F}(Theorem 1.1) \) gave a combinatorial definition for the monomial expansion of Schubert polynomials in terms of compatible sequences for reduced words. Assaf and Searles \( \mathcal{F}(Theorem 3.13) \) refined this to give a combinatorial model for the expansion in terms of fundamental slide polynomials. The re-formulation of the latter given below appears in \( \mathcal{F}(Theorem 3.3) \), and we take this as our definition.
Figure 3. Constructing the weak descent tableaux for reduced words \( \rho = (3, 6, 4, 7, 5, 2, 4) \) (left) and \( \sigma = (6, 7, 3, 4, 5, 2, 4) \) (right).

**Definition 2.2.5.** For \( w \) any permutation, the Schubert polynomial \( S_w \) is

\[
S_w = \sum_{\rho \in R(w)} S_{\text{des}(\rho)},
\]

where the sum may be taken over non-virtual reduced words \( \rho \).

For example, the seven non-virtual reduced words for \( R(153264) \) give

\[
S_{153264} = S_{(0,3,1,0,1)} + S_{(2,2,0,0,1)} + S_{(1,3,0,0,1)} + S_{(0,3,2,0,0)}
+ S_{(2,2,1,0,0)} + S_{(1,3,1,0,0)} + S_{(2,3,0,0,0)}.
\]

Macdonald [16](7.18) showed that Schubert polynomials stabilize and that their stable limits are precisely the Stanley symmetric functions. This follows directly from Proposition 2.2.2 by Definitions 2.2.5 and 2.1.4 as well.

**Proposition 2.2.6** ([16]). For \( w \) a permutation, we have

\[
\lim_{m \to \infty} S_{1^m \times w}(x_1, \ldots, x_m, 0, \ldots, 0) = S_w(X),
\]

where \( 1^m \times w \) is the permutation obtained by adding \( m \) to each \( w \), and then prepending \( 12 \cdots m \).

3. Equivalence relations

We consider simple involutions based on the Coxeter relations for the simple transpositions that generate the symmetric group.

Given \( \rho \in R(w) \), for \( 1 \leq j < \ell(w) \), let \( c_j \) denote the commutation relation that acts by exchanging \( \rho_j \) and \( \rho_{j+1} \) if \( |\rho_j - \rho_{j+1}| > 1 \) and the identity otherwise.

Given \( \rho \in R(w) \), for \( 1 < j < \ell(w) \), let \( b_j \) denote the braid relation that acts by sending \( \rho_j \rho_{j+1} \rho_j \) to \( \rho_{j+1} \rho_j \rho_{j+1} \) if \( \rho_{j+1} = \rho_{j+1} \) and the identity otherwise.

Any reduced words in the same equivalence class under \( \{c_j, b_j\} \) are called Coxeter equivalent. A classical result of Tits [22] states that each set \( R(w) \) is a single Coxeter equivalence class. For examples, see Fig. 4.

Knuth [12] considered relations on permutations that characterize when two permutations give rise to the same Schensted insertion tableau [20]. Analogously, Edelman and Greene [10] characterize when two reduced words give rise to the same Edelman–Greene insertion tableau using elementary Coxeter-Knuth relations.
Definition 3.0.1. For \( 1 < i < \ell(w) \), the elementary Coxeter-Knuth relation \( \mathfrak{d}_i \) acts on a reduced word \( \rho \in R(w) \) by

\[
\mathfrak{d}_i(\rho) = \begin{cases} 
\mathfrak{b}_i(\rho) & \text{if } \rho_{i+1} = \rho_{i-1} = \rho_i \pm 1 \\
\mathfrak{c}_{i-1}(\rho) & \text{if } \rho_{i-1} > \rho_{i+1} > \rho_i \text{ or } \rho_{i-1} < \rho_{i+1} < \rho_i, \\
\mathfrak{c}_i(\rho) & \text{if } \rho_{i+1} > \rho_{i-1} > \rho_i \text{ or } \rho_{i+1} < \rho_{i-1} < \rho_i, \\
\rho & \text{otherwise,}
\end{cases}
\]

where \( \mathfrak{c}_j \) denotes a commutation relation and \( \mathfrak{b}_j \) denotes a braid relation.

We partition \( R(w) \) by stating any reduced words in the same equivalence class under \( \{\mathfrak{d}_i\} \) are Coxeter-Knuth equivalent. For example, see Fig. 5.

Inverting history, a natural question to ask is whether this partitioning can be realized on the level of symmetric functions by decomposing the Stanley symmetric functions or on the level of polynomials by decomposing Schubert polynomials.

3.1. Dual equivalence. Based on the explicit elementary dual equivalence involutions on standard Young tableaux, Assaf [3, 4] defined an abstract notion of dual equivalence that can be used to prove that a given fundamental quasisymmetric generating function is symmetric and Schur positive.
A **Young diagram** is the set of unit cells in the first quadrant with \( \lambda_i \) cells in row \( i \) for some partition \( \lambda \). A **Young tableau** is a filling of a Young diagram with positive integers. A Young tableau is *increasing* if it has strictly increasing rows (left to right) and columns (bottom to top). A Young tableau is *standard* if it is increasing and uses each integer \( 1, 2, \ldots, n \) exactly once. For example, Figure 6 shows the standard Young tableaux of shape \( (3, 2) \).

![Figure 6](image-url) The standard Young tableaux for \( \lambda = (3, 2) \) and their descent compositions.

For a standard Young tableau \( T \), say \( i \) is a descent of \( T \) if \( i + 1 \) lies weakly left of \( i \). The **descent composition** of \( T \), denoted by \( \text{Des}_T \), is the strong composition given by maximal lengths of runs of the word \( 1 2 \ldots n \) crossing no descents. For example, the descent compositions for the tableaux in Figure 6.

Schur functions may be defined combinatorially as the fundamental quasisymmetric generating functions for standard Young tableaux. This follows from the classical definition (see [17]) by results of Gessel [11].

**Definition 3.1.1.** For \( \lambda \) a partition, the **Schur function** \( s_\lambda \) is

\[
s_\lambda(X) = \sum_{T \in \text{SYT}(\lambda)} F_{\text{Des}_T}(X).
\]

For example, from Figure 6 we have

\[
s_{(3,2)}(X) = F_{(2,3)}(X) + F_{(1,2,2)}(X) + F_{(1,3,1)}(X) + F_{(3,2)}(X) + F_{(2,2,1)}(X).
\]

A **dual equivalence** for a set of objects endowed with a descent statistic is a family of involutions \( d_2, \ldots, d_{n-1} \) such that \( d_id_j = d_jd_i \) for \( |i-j| \geq 3 \) and for which each restricted equivalence classes under \( d_i, \ldots, d_j \) for \( j - i \leq 3 \) is a single Schur function. The main theorem for dual equivalence [4](Corollary 4.4) states that this local Schur positivity implies global Schur positivity.

**Theorem 3.1.2 ([4]).** The fundamental quasisymmetric generating function of any dual equivalence class is a single Schur function.

Stanley proved that \( S_w \) is symmetric [21](Theorem 2.1) and conjectured that it is, in fact, **Schur positive**, meaning the expansion into Schur functions has only nonnegative coefficients. For example,

\[
S_{153264}(X) = s_{(3,2)}(X) + s_{(3,1,1)}(X).
\]

Edelman and Greene [10] proved this by generalizing the Robinson–Schensted–Knuth insertion algorithm [19, 20, 12] on permutations.

**Theorem 3.1.3 ([10]).** For \( w \) a permutation, we have

\[
S_w = \sum_{\rho \in R(w)} s_{\text{Des}(\rho)},
\]

where \( \rho \) is increasing, \( \text{row}(T_\rho) = \rho \).
where row(T) is the row reading word (left to right along rows from the top) of T.

For example, from Fig. 4 the two reduced words for 153264 that are the row reading words of increasing tableaux are (3, 5 | 2, 3, 4) and (5 | 3 | 2, 3, 4), corresponding to the Schur expansion of $S_{153264}$ given above.

The Edelman–Greene correspondence is an elegant solution to the Schur positivity conjecture, but the arguments involved in the proof require intricate analysis of bumping paths with many separate cases. Thus one can hope to find a simpler proof that avoids much of this subtlety.

Edelman and Greene \[10\] (Corollary 6.15) relate Coxeter–Knuth equivalence with dual equivalence through the Edelman–Greene recording tableaux. Implicit in their work and explicit in \[6\] (Theorem 2.10), the Coxeter–Knuth involutions give a dual equivalence on reduced words.

**Theorem 3.1.4** \[6\]. The Coxeter–Knuth involutions $\varnothing_i$ give a dual equivalence for $R(w)$, thus Stanley symmetric functions are symmetric and Schur positive.

That is, the Coxeter–Knuth relations $\varnothing_i$ partition reduced words for a given permutation into dual equivalence classes, each of which has fundamental quasisymmetric generating function equal to a single Schur function. However, while the proof of Theorem 3.1.4 is simple, the resulting formula requires computing each equivalence class in its entirety, falling short of the explicit formula in Theorem 3.1.3.

### 3.2. Weak dual equivalence

The Demazure characters, introduced by Demazure \[9\], originally arose as characters of Demazure modules for the general linear group \[8\]. These polynomials were studied combinatorially by Lascoux and Schützenberger \[15\], who call them standard bases, and more extensively by Reiner and Shimozono \[18\] who call them key polynomials. We use the key tableau model \[6\] based on ideas of Kohnert \[13\] developed further by Assaf and Searles \[2\].

A key diagram is a collection of left-justified unit cells in the right half place with $a_i$ cells in row $i$ for some weak composition $a$. A key tableau is a filling of a key diagram with positive integers. The definition for standard key tableaux \[6\] (Definition 3.10) is more subtle than for standard Young tableaux.

**Definition 3.2.1** \[6\]. A standard key tableau is a bijective filling of a key diagram with $\{1, 2, \ldots, n\}$ such that rows decrease (left to right) and if some entry $i$ is above and in the same column as an entry $k$ with $i < k$, then there is an entry right of $k$, say $j$, such that $i < j$.

```latex
\begin{center}
\begin{tabular}{cccc}
5 & 4 & 3 & 2 \\
3 & 2 & 1 & |
\end{tabular} & 
\begin{tabular}{cccc}
5 & 3 & 4 & 2 \\
4 & 2 & 1 & |
\end{tabular} & 
\begin{tabular}{cccc}
5 & 1 & 4 & 2 \\
1 & 3 & 2 & |
\end{tabular} & 
\begin{tabular}{cccc}
3 & 1 & 5 & 4 \\
\ | & \ | & |
\end{tabular} & 
\begin{tabular}{cccc}
2 & 1 & 5 & 4 \\
\ | & \ | & |
\end{tabular}
\end{center}
```

(0, 3, 0, 2) (2, 2, 0, 1) (1, 3, 0, 1) $\emptyset$ (2, 3, 0, 0)

**Figure 7.** Standard key tableaux of shape (0, 3, 0, 2) and their weak descent compositions.
For a standard key tableau $T$, say $i$ is a **descent of** $T$ if $i + 1$ lies weakly right of $i$ in $T$. Note that this is the reverse of the concept of descents for standard Young tableaux. Next we define a **weak descent composition** \([6](Definition\ 3.12)\).

**Definition 3.2.2** \((\text{[6]}\)\). For a standard key tableau $T$, define the **weak descent composition** of $T$, denoted by $\text{des}(T)$, as follows. Let $(\tau(1)|\tau(2)|\cdots|\tau(k))$ be the run decomposition of $n\cdot \ldots \cdot 21$ based on descents of $T$, that is, each $\tau(i)$ has no descents between adjacent letters and is as long as possible. Set $t_i = \min(\tau(i))$ for $i = 1, \ldots, k$. Set $\hat{t}_k = r_k$, and for $i < k$, set $\hat{t}_i = \min(\hat{t}_i, \hat{t}_{i+1} - 1)$. If $t_1 \leq 0$, then define $\text{des}(\rho) = \emptyset$; otherwise, set the $\hat{t}_i$th part of $\text{des}(T)$ to be $\text{des}(T)_{\hat{t}_i} = |\tau(i)|$ and set all other parts to 0.

For example, the standard key tableaux of shape $(0, 3, 0, 2)$ shown in Figure 7 have weak descent compositions given beneath.

Using this notion, we have the following reformulation of Demazure characters given in \([6](Corollary\ 3.16)\) that we take as our definition.

**Definition 3.2.3.** Given a weak composition $a$, we have

\[(3.4)\]

\[
\kappa_a = \sum_{T \in \text{SKT}(a)} \mathfrak{F}_{\text{des}(T)}.
\]

For example, from Figure 7 we compute

\[
\kappa_{(0,3,0,2)} = \mathfrak{F}(0,3,0,2) + \mathfrak{F}(2,2,0,1) + \mathfrak{F}(1,3,0,1) + \mathfrak{F}(2,3,0,0).
\]

Implicit in the work of Lascoux and Schützenberger \([15]\) and explicit in that of Assaf and Searles \([2]\)(Corollary 4.9), we have the following analog of Proposition 2.2.6 for Demazure characters.

**Proposition 3.2.4** \((\text{[2]}\)\). For a weak composition $a$, the Demazure character is

\[(3.5)\]

\[
\lim_{m \to \infty} \kappa_0^{m \times a}(x_1, \ldots, x_m, 0, \ldots, 0) = s_{\text{sort}(a)}(X),
\]

where sort($a$) is the partition obtained by sorting $a$ into weakly decreasing order.

Generalizing dual equivalence, a **weak dual equivalence** is a family of involutions $\hat{a}_2, \ldots, \hat{a}_{n-1}$ that give a dual equivalence when weak descent compositions are flattened to descent compositions by removing parts equal to 0 and for which the fundamental slide generating polynomial of each restricted equivalence class under $\hat{a}_i, \ldots, \hat{a}_j$ for $j - i \leq 3$ is a single Demazure character. Parallel to the symmetric case, the main theorem for weak dual equivalence \([6](Theorem\ 3.33)\) states that this local Demazure positivity implies global Demazure positivity.

**Theorem 3.2.5** \((\text{[6]}\)\). The fundamental slide generating polynomial of any weak dual equivalence class is a single Demazure character.

One might now anticipate that Schubert polynomials expand nonnegatively into Demazure characters, parallel to \((3.3)\), and indeed, we have,

\[
\mathfrak{S}_{153264} = \kappa_{(0,3,1,0,1)} + \kappa_{(0,3,2,0,0)}.
\]

Lascoux and Schützenberger \([14]\) give a formula for the key polynomial expansion of a Schubert polynomial as a sum over increasing Young tableau whose row reading word is a reduced word for $w$, where for each such $\rho$ one computes the **left nil key** by considering all reduced words of $w$ that are Coxeter–Knuth equivalent to $\rho$. For details that fill the gaps in \([15]\), see \([18]\)(Theorem 4).
Theorem 3.2.6 ([15, 18]). For \( w \) a permutation, we have
\[
S_w = \sum_{\rho \in R(w)} K_{\text{content}}(K_0^\rho)(\rho),
\]
where \( K_0^\rho \) is the left nil key of \( \rho \).

While theoretically interesting for the nonnegativity, this result does not provide
a direct formula as one is required to compute each Coxeter–Knuth class, and so the
computation is effectively equivalent to computing the fundamental slide expansion.
A simplified proof comes as an immediate application of weak dual equivalence.

Theorem 3.2.7 ([6]). The Coxeter–Knuth involutions \( d_i \) give a weak dual equiv-
alence for \( R_p w q \). In particular, Schubert polynomials are Demazure positive.

That is, the Coxeter–Knuth relations \( d_i \) partition reduced words for a given
permutation into weak dual equivalence classes, each of which has fundamental slide
generating polynomial equal to a single Demazure character. This gives a simplified
proof of Theorem 3.2.6 though the resulting formula is no more tractable.

4. Positive expansions

By Theorem 3.1.4 each Coxeter–Knuth equivalence class corresponds to a term
in the Schur expansion of a Stanley symmetric function. Similarly, by Theo-
rem 3.2.7 each Coxeter–Knuth equivalence class corresponds to a term in the De-
mazure expansion of a Schubert polynomial. To make these positivity results more
compelling, we wish to have canonical representatives from each Coxeter–Knuth
equivalence class from which an exact formula can be easily computed.

Edelman and Greene [10] resolved this for the Schur expansion of Stanley sym-
metric functions, but we wish to give a simple, self-contained proof of their formula
that avoids the subtleties of their insertion algorithms. The end result will be
the same, however, namely that each Coxeter–Knuth equivalence class contains a
unique reduced word whose descent tableau is an increasing Young tableau. Then
the shape of these tableaux determines the Schur expansion.

\begin{figure}
\centering
\begin{tabular}{llll}
  6 & 6 & 6 & 6 \\
  5 & 7 & 5 & 5 \\
  3 & 3 & 3 & 3 \\
  2 & 4 & 4 & 2 \\
\end{tabular}
\caption{The set of increasing Young tableaux whose row reading
words are reduced words for \( w = 13625847 \).}
\end{figure}

For example, we compute the Schur expansion of Stanley symmetric function
\( S_{13625847} \) by constructing the five increasing Young tableaux in Figure 8 giving
\[
S_{13625847} = s_{(3,2,1,1)} + s_{(3,2,2)} + s_{(3,3,1)} + s_{(4,1,1,1)} + s_{(4,2,1)}.
\]

However, in the Schubert case, these are not the correct Coxeter–Knuth equiva-

cence class representatives for giving the Demazure expansion. In light of Proposi-
tion 3.2.3 there are many different candidates for which weak composition should
index each class, even knowing the correct partition. Using the same techniques
For example, Fig. 10 shows the drop alignments for two pairs of increasing words.

![Figure 9. The set of Yamanouchi key tableaux whose row reading words are reduced words for \( w = 13625847 \).](image)

4.1. Increasing Young tableaux. We begin by considering the descent tableaux for reduced words, and, more generally, any tableau with weakly increasing rows for which the reading word is reduced.

**Definition 4.1.1.** Given two increasing words \( \sigma, \tau \) of lengths \( s, t \), respectively, define the drop alignment of \( \sigma \) below \( \tau \) as follows: If \( \tau_j > \sigma_j \) for all \( j \leq \min(s, t) \), then left justify \( \sigma \) with respect to \( \tau \). Otherwise, set \( j_1 \) to be the minimum index such that \( \tau_j \leq \sigma_j \), align \( \sigma_1 \cdots \sigma_{j_1-1} \) directly under \( \tau_1 \cdots \tau_{j_1-1} \) and iterate the process with the drop alignment of \( \sigma_{j_1} \cdots \sigma_s \) below \( \tau_{j_1+1} \cdots \tau_t \).

Visually, begin with \( \sigma \) left justified under \( \tau \), and from left to right, for each not strict column, slide entries of \( \sigma \) from that column onward right by one position. For example, Fig. 10 shows the drop alignments for two pairs of increasing words.

![Figure 10. The drop alignment of \((4,7)\) below \((3,6)\) (left) and of \((2,4)\) below \((3,4,5)\) (right).](image)

If there are \( k \) instances in the drop alignment of \( \sigma \) below \( \tau \) where a cell of \( \tau \) has no cell below it, then we denote these cells of \( \tau \) as \( x_1, \ldots, x_k \) and factor \( \tau = \tau^{(0)}x_1\tau^{(1)} \cdots x_k\tau^{(k)} \) and, correspondingly, \( \sigma = \sigma^{(0)}\sigma^{(1)} \cdots \sigma^{(k)}\sigma^{(k+1)} \) as shown in Fig. 11. When the concatenation \( \tau\sigma \) is a reduced word, then this factorization has the following properties.

**Proposition 4.1.2.** Given two increasing words \( \tau, \sigma \) such that \( \tau\sigma \) is reduced, there is a unique factorization \( \tau = \tau^{(0)}x_1\tau^{(1)} \cdots x_k\tau^{(k)} \) and \( \sigma = \sigma^{(0)}\sigma^{(1)} \cdots \sigma^{(k)}\sigma^{(k+1)} \), with some \( \tau^{(i)} \) or \( \sigma^{(i)} \) possibly empty, such that...
\[ \tau^{(0)} x_1 \tau^{(1)} \ldots x_k \tau^{(k)} \]
\[ \sigma^{(0)} \sigma^{(1)} \ldots \sigma^{(k)} \sigma^{(k+1)} \]

Figure 11. An illustration of the drop alignment of \( \sigma \) below \( \tau \) and corresponding factorizations of \( \tau \) and \( \sigma \).

1. \( \ell(\tau^{(j)}) = \ell(\sigma^{(j)}) \) for \( j = 1, \ldots, k \);  
2. \( t_i^{(j)} > s_i^{(j)} \) for \( j = 1, \ldots, k \) and \( i = 1, \ldots, \ell(\tau^{(j)}) \);  
3. \( x_j \leq s_i^{(j)} \) with equality only if \( t_1^{(j)} = s_i^{(j)} + 1 \), for \( j = 1, \ldots, k \).

Proof. The algorithm in Definition 4.1.3 is clearly well-defined and conditions (1) and (2) follow immediately by construction. The hypothesis that \( \tau \sigma \) is reduced is needed only for condition (3). Suppose \( j_1 \) is the minimum index such that \( t_{j_1} \leq s_{j_1} \) and that, in fact, \( t_{j_1} = s_{j_1} \). Condition (3) requires \( t_{j_1+1} = s_{j_1} + 1 \). Suppose, for contradiction, this does not hold. Then either \( j_1 = t \) or \( t_{j_1+1} > s_{j_1} + 1 \).

By minimality of \( j_1 \), we have \( t_i > s_i \) for all \( i < j_1 \). In particular, \( s_1, \ldots, s_{j_1-1} < t_{j_1-1} < t_{j_1} \), so the word \( \tau \sigma \) is Coxeter equivalent to the word

\[ \tau_1 \cdots \tau_{j_1-1} s_1 \cdots s_{j_1-1} t_{j_1} \cdots t_{j_1} \sigma_{j_1} \cdots \sigma_s. \]

If \( j_1 = t \), then \( t_{j_1} = s_{j_1} \) are adjacent, contradicting the fact that the word is reduced. If \( j_1 < t \), then \( t_{j_1+1} > s_{j_1} + 1 = t_{j_1} + 1 \), and so the above word is Coxeter equivalent to the word

\[ \tau_1 \cdots \tau_{j_1-1} s_1 \cdots s_{j_1-1} t_{j_1} \sigma_{j_1} \tau_{j_1+1} \cdots t_{j_1} \sigma_{j_1} \cdots \sigma_s. \]

In this case as well, \( t_{j_1} = s_{j_1} \) are adjacent, contradicting the fact that the word is reduced. Thus condition (3) must hold whenever \( \tau \sigma \) is reduced. \( \square \)

We call this factorization the drop alignment because, as we show below, we may drop the unsupported cells \( x_1, \ldots, x_k \) from \( \tau \) down to \( \sigma \) without changing the Coxeter–Knuth equivalence class.

Definition 4.1.3. Given two increasing words \( \tau, \sigma \) such that \( \tau \sigma \) is reduced, define

\[ \text{drop}(\tau \sigma) = \tau^{(0)} \tau^{(1)} \ldots \tau^{(k)} \sigma^{(0)} \sigma^{(1)} \ldots x_k \sigma^{(k)} \sigma^{(k+1)}, \]

where \( \tau = \tau^{(0)} \tau^{(1)} \ldots x_k \tau^{(k)} \) and \( \sigma = \sigma^{(0)} \sigma^{(1)} \ldots \sigma^{(k)} \sigma^{(k+1)} \) is the unique factorization of Proposition 4.1.2 and for \( 1 \leq j \leq k \), we set

\[ \hat{s}_i^{(j)} = \begin{cases} s_i^{(j)} + 1 & \text{for } 1 \leq i \leq b_j \\ s_i^{(j)} & \text{for } b_j + 1 \leq i \leq \ell(\sigma^{(j)}) \end{cases} \]

where \( b_j = \max\{b \mid t_1^{(j)} = s_i^{(j)} + i \forall 1 \leq i \leq b \} \) if \( x_j = t_1^{(j)} - 1 \) and 0 otherwise.

For example, dropping the aligned words in Fig. 10 results in the words in Fig. 12.

The following elementary lemma will be useful in proving that \( \tau \sigma \) is Coxeter–Knuth equivalent to \( \text{drop}(\tau \sigma) \).

Lemma 4.1.4. Given an increasing word \( \sigma = (\sigma_\ell, \ldots, \sigma_1) \) and a letter \( x \) such that \( x < \sigma_\ell \), we have \( \sigma x \sigma_{\ell-1} \cdots \sigma_1 \) is Coxeter–Knuth equivalent to \( \sigma_\ell \sigma_{\ell-1} \cdots \sigma_1 x \).
Figure 12. The drop of (4, 7) below (3, 6) (left) and of (2, 4) below (3, 4, 5) (right).

**Proof.** We claim \( \mathfrak{d}_2 \cdots \mathfrak{d}_{\ell-1} (\sigma_1 \cdots \sigma_{\ell-1} \cdots \sigma_1) = \sigma x \), from which the assertion follows. To see this, notice that from the hypotheses we have \( x < \sigma_k < \sigma_{k-1} \) for \( k = \ell, \ldots, 2 \). Thus we may apply \( \mathfrak{d}_{k-1} \), for \( k = \ell, \ldots, 3 \), with effect

\[
\mathfrak{d}_{k-1}(\sigma_{\ell} \cdots \sigma_k x \sigma_{k-1} \cdots \sigma_1) = \sigma_{\ell} \cdots \sigma_{k-1} x \sigma_{k-2} \cdots \sigma_1.
\]

The culmination of these elementary equivalences gives the desired result. \( \square \)

**Lemma 4.1.5** (Drop Lemma). Given two increasing words \( \tau, \sigma \) such that \( \tau \sigma \) is reduced, drop(\( \tau \sigma \)) is Coxeter–Knuth equivalent to \( \tau \sigma \).

**Proof.** With the factorization as denoted in Proposition 4.1.2(3), consider first the case \( k = 1 \), in which we must show \( \tau(0) x \tau(1) \sigma(0) \sigma(1) (\sigma(2)) \) is Coxeter–Knuth equivalent to \( \tau(0) \tau(1) \sigma(0) x \sigma(1) \sigma(2) \). Let \( \ell_j = \ell(\sigma(j)) \) for \( j = 0, 1, 2 \). Since \( \tau \) is increasing and \( \tau_i(j) > \sigma_i(j) \) for \( j = 0, 1 \), we can apply Lemma 4.1.4 \( \ell_0 + \ell_1 \) times to move letters of \( \sigma(0) \sigma(1) \) left so that \( \tau \sigma \) is Coxeter–Knuth equivalent to

\[
\tau(0) \sigma(0) \cdots \tau(0) \sigma(0) x \tau(1) \sigma_1(1) \cdots \tau(1) \sigma(1) \sigma(2).
\]

If \( \ell_1 = 0 \), then we may apply Lemma 4.1.4 to \( \sigma_i(0) \) for \( i < \ell_0 \) to deduce Eq. (4.1) is Coxeter–Knuth equivalent to

\[
\tau(0) \sigma_1(0) \cdots \tau(0) \sigma_1(0) x \sigma(2),
\]

as desired. If \( \ell_1 > 0 \) and \( \tau_1(1) > \sigma_1(1) + 1 \), then by Proposition 4.1.2(3), \( \sigma(0) x \sigma(1) \sigma(2) \) is not reduced. Applying Lemma 4.1.4 \( \ell_1 \) times to the letters of \( \sigma(1) \), then to \( x \), then \( \ell_0 \) times to the letters of \( \sigma(0) \) shows Eq. (4.1) is Coxeter–Knuth equivalent to \( \tau(0) \tau(1) \sigma(0) x \sigma(1) \sigma(2) \), as desired.

Finally, suppose \( \ell_1 > 0 \) and \( \tau_1(1) = \sigma_i(1) + i \) for \( i < b \) with \( b > 0 \) maximal. By Proposition 4.1.2(3), \( \sigma(0) x \sigma(1) \sigma(2) \) is not reduced. Then notice \( \tau_1(1) = \sigma_i(1) + 1 \) for \( i < b \), since otherwise \( \sigma(1) \sigma(2) \) is not reduced. Therefore \( \sigma(0) x \sigma(1) \sigma(2) \) is not reduced. Putting this together, Eq. (4.1) is Coxeter–Knuth equivalent to

\[
\tau(0) \sigma_1(0) \cdots \tau(0) \sigma_1(0) x \tau(1) (\sigma_1(1) + 1) \cdots \tau(1) (\sigma_i(1) + 1) \sigma_{b+1} \sigma_{b+1} \cdots \tau(1) \sigma(2).
\]

Notice that the letters of \( \sigma(1) \) have now been raised to letters of \( \hat{\sigma}(1) \). Finally, apply Lemma 4.1.4 \( \ell_1 \) times to the letters of \( \hat{\sigma}(1) \), then to \( x \), then \( \ell_0 \) times to the letters of \( \sigma(0) \) to see this is Coxeter–Knuth equivalent to \( \tau(0) \tau(1) \sigma(0) x \sigma(1) \sigma(2) \), as desired.

For \( k > 1 \), the above case shows \( \tau \sigma \) is Coxeter–Knuth equivalent to

\[
\tau(0) x_1 \tau(1) \cdots x_{k-1} \tau(k-1) \tau(k) \sigma(0) \sigma(1) \cdots \sigma(k-1) x_k \hat{\sigma}(k) \sigma(k+1).
\]
However, notice that this new pair factors uniquely by combining
\( \tau^{(k-1)} \) and \( \tau^{(k)} \), as well as \( \sigma^{(k-1)} \), \( x \), and \( \sigma^{(k)}_1 \cdots \sigma^{(k)}_{k-1} \), then also combining \( \sigma^{(k)}_1 \) and \( \sigma^{(k+1)} \). Therefore there are fewer factors in the result, so by induction, we may drop the remaining \( x_1, \ldots, x_{k-1} \) as well, completing the proof. 

Extending Definition 4.1.3 we define the drop of any reduced word based on the rows of its descent tableau.

**Definition 4.1.6.** Let \( \rho \) be a reduced word, and let \( \rho = \rho^{(k)} \cdots | \rho^{(1)} \) be an increasing decomposition of \( \rho \). Define \( \text{drop}_i(\rho) \) by replacing \( \rho^{(k)} \cdots | \rho^{(1)} \) with \( \text{drop}(\rho^{(i+1)}|\rho^{(i)}) \).

Figure 13. All drop sequences taking \( \mathbb{D}(3, 6, 4, 7, 5, 2, 4) \) to the unique increasing tableau in its Coxeter–Knuth equivalence class.

We visualize increasing factorizations as tableaux with strictly increasing rows, and then the drop maps can be visualized as dropping cells in the tableaux, as shown in Fig. 13.

**Proposition 4.1.7.** For \( \rho \) a reduced word with run decomposition \( (\rho^{(k)}|\cdots|\rho^{(1)}) \), \( \mathbb{D}(\rho) \) is an increasing Young tableau if and only if \( \text{drop}_i(\rho) = \rho \) for all \( 1 \leq i < k \).

**Proof.** If \( \mathbb{D}(\rho) \) is increasing, then in Definition 4.1.1 we will always left justify \( \rho^{(j)} \) with respect to \( \rho^{(j+1)} \) resulting in no \( x_i \)’s in the unique factorization of Proposition 4.1.2 and so \( \text{drop}_j(\rho) = \rho \) for all \( j \). Moreover, this is the only case in which there are no \( x_i \)’s, and so the only case when \( \text{drop}_j(\rho) = \rho \) for all \( j \). 

We apply the maps \( \text{drop}_i \) until reaching this terminal state, in which case we have an increasing Young tableau. To see that this is independent of the choice of which rows to drop when, we observe that these maps satisfy the nil-Hecke relations.

**Theorem 4.1.8.** The maps \( \text{drop}_i \) are well-defined on reduced words satisfying

(i) \( \text{drop}_i \circ \text{drop}_i = \text{drop}_i; \)
(ii) \( \text{drop}_i \circ \text{drop}_j = \text{drop}_j \circ \text{drop}_i \) for \( |i - j| > 1; \)
(iii) \( \text{drop}_i \circ \text{drop}_{i+1} \circ \text{drop}_i = \text{drop}_{i+1} \circ \text{drop}_i \circ \text{drop}_{i+1}. \)
Proof. By the Drop Lemma, $\rho$ is Coxeter–Knuth equivalent to $\text{drop}_i(\rho)$. In particular, when $\rho$ is reduced, so is $\text{drop}_i(\rho)$, so we may iterate the maps.

For relation (ii), notice that since $x_i < \sigma_1^{(i)}$ for $i = 1, \ldots, k$, with notation as in Proposition 4.1.2 the columns of $D(\text{drop}(\sigma \tau))$ are strict when the two rows are left justified. Therefore, by uniqueness of the factorization in Proposition 4.1.2 there will be no $x_i$'s for $\text{drop}(\sigma \tau)$. In particular, drop will act trivially.

Relation (iii) follows from the fact that drop $\tau$ considers only rows $k$ and $k + 1$, so for $|i - j| > 1$, the sets of indices $\{i, i + 1\}$ and $\{j, j + 1\}$ are disjoint.

Finally, for relation (iii) it is enough to consider a three term factorization, say $\tau \sigma \rho$ with each of $\tau, \sigma, \rho$ increasing and the concatenation reduced. We must show that $\text{drop}_1 \text{drop}_2 \text{drop}_3(\tau \sigma \rho)$ is equal to $\text{drop}_2 \text{drop}_1 \text{drop}_3(\tau \sigma \rho)$. Align $\rho$ below $\sigma$ with letters of unsupported letters of $\sigma$ denoted $x_1, \ldots, x_k$, and then, maintaining that alignment, align $\sigma$ below $\tau$ with letters of unsupported letters of $\tau$ denoted $y_1, \ldots, y_l$; call this the initial alignment. For a generic example, see Fig. 14. We consider blocks of the initial alignment with respect to the factorization of $\tau$ that put each $y_i$ at the end of a block as indicated.

\[
\begin{align*}
\tau : & \quad x \times x \times y_1 \times \times \times \times y_2 \times y_3 \times y_4 \times \\
\sigma : & \quad x \times x \times x_1 \times x_2 \times x_3 \times \times \times \times \times \\
\rho : & \quad x \times \times \times \times \times \times \times \times \\
\end{align*}
\]

Figure 14. An example of the drop alignment of $\rho$ below $\sigma$ below $\tau$.

Consider first $\text{drop}_1 \text{drop}_2 \text{drop}_3(\tau \sigma \rho)$. Applying $\text{drop}_1(\tau \sigma \rho) = \tau \text{drop}(\sigma \rho)$ results in all $x_i$'s moving down to row 1. When subsequently applying $\text{drop}_2$ to the result, in addition to the $y_i$'s dropping to row 2, for every $x_i$, letting $j_i$ be smallest such that $y_{j_i}$ is right of $x_i$, some entry in row 3 weakly right of $x_i$ and strictly left of $y_{j_i}$, will also drop to row 2. The final application of $\text{drop}_1$ now pulls down one cell from row 2 for each $y_i$, and this will either be $y_i$ or some entry originally from $\tau$ that is weakly right of $y_i$. In particular, for $\text{drop}_1 \text{drop}_2 \text{drop}_3(\tau \sigma \rho)$, we have

- the top row has no $y_i$'s and loses additional cells based on the drop alignment of $\sigma$ with $x_i$'s removed below $\tau$;
- the bottom row has all $x_i$'s and additional cells originally from $\tau$ based on the drop alignment of $\rho$ with $x_i$'s added below $\tau$.

Consider next $\text{drop}_2 \text{drop}_1 \text{drop}_3(\tau \sigma \rho)$. Applying $\text{drop}_2(\tau \sigma \rho) = \text{drop}(\tau \sigma \rho)$ results in all $y_i$'s moving down to row 2. When subsequently applying $\text{drop}_3$ to the result, every $x_i$ will move down and, in addition, one cell originally from $\tau$ for every $y_i$ and this cell must be weakly right of $y_i$. The final application of $\text{drop}_1$ will pull down one cell from $\tau$ for every $x_i$ for which did not sit directly below some $y_{j_i}$. In particular, the description above holds for the top and bottom rows of $\text{drop}_2 \text{drop}_1 \text{drop}_3(\tau \sigma \rho)$, and so the two must correspond for all three rows. \hfill $\square$

Definition 4.1.9. For a reduced word $\rho$, the drop of $\rho$, denoted by $\text{drop}(\rho)$, is defined as follows: Let $\rho^{(k)}$ denote the $k$th row of $D(\rho)$. Choose any $k$ such that either $\ell(\rho^{(k+1)}) > \ell(\rho^{(k)})$ or $\rho_j^{(k+1)} \preceq \rho_j^{(k)}$ for some $j \leq \ell(\rho^{(k)})$. Replace $\rho^{(k+1)} \rho^{(k)}$ with $\text{drop}(\rho^{(k+1)}) \rho^{(k)}$. Repeat until the result is an increasing Young tableau.

By Theorem 4.1.8 the definition of $\text{drop}(\rho)$ is independent of the order in which rows are consolidated, and so we have the following.
Theorem 4.1.10. For a reduced word \( \rho \), \( \text{drop}(\rho) \) is the unique Coxeter–Knuth equivalent reduced word for which the descent tableau is increasing and of partition shape.

Proof. By Theorem 4.1.8 the end result is independent of the order in which the \( \text{drop}_k \) are applied, and so \( \text{drop}(\rho) \) is well-defined. By Lemma 4.1.5 each application of \( \text{drop}_k \) stays within the Coxeter–Knuth class, and so the theorem follows. \( \square \)

Theorem 4.1.10 shows that each Coxeter–Knuth equivalence class has a unique representative \( \pi \) characterized by the property \( \mathbb{D}(\pi) \) is an increasing Young tableau of partition shape. In particular, combining this with Theorem 3.1.4, this gives a new and simplified proof of the following equivalent formulation of Theorem 3.1.3.

Corollary 4.1.11. For \( w \) a permutation, we have

\[
S_w = \sum_{\rho \in \mathcal{R}(w), \mathbb{D}(\pi) \text{ increasing}} \kappa_{\text{des}(\rho)}.
\]

4.2. Yamanouchi key tableaux. In order to give an explicit, direct formula for the Demazure expansion of a Schubert polynomial, we begin by characterizing the analogs of increasing Young tableaux that will give our canonical representatives for each Coxeter–Knuth equivalence class in the polynomial setting.

Lemma 4.2.1. Given a weak composition \( a \), there exists a unique \( T_a \in \text{SKT}(a) \) for which \( \text{des}(T_a) = a \). Moreover, for \( T \in \text{SKT}(a) \) non-virtual, we have \( \text{des}(T) \geq a \), where we define \( b \geq a \) if and only if \( b_1 + \cdots + b_k \geq a_1 + \cdots + a_k \) for all \( k \).

Proof. We construct \( T_a \) by filling the key diagram for \( a \) with entries \( n, n-1, \ldots, 2, 1 \) left to right along rows beginning with the top. Then rows decrease left to right and columns decrease top to bottom, showing \( T_a \in \text{SKT}(a) \). Definition 3.2.2 will have \( t_i = t_i \) for all \( i \), so \( T_a \) will indeed have \( \text{des}(T_a) = a \). Uniqueness follows since any alternative filling \( T \in \text{SKT}(a) \) necessarily has \( \text{flat}(\text{des}(T)) \neq \text{flat}(\text{des}(T_a)) = \text{flat}(a) \). Finally, the latter condition follows from the upper unitriangularity of Demazure characters with respect to monomials. \( \square \)

Thus it makes sense to consider the set of reduced words that subordinate their Coxeter–Knuth equivalence classes.

Definition 4.2.2. A reduced word \( \rho \) is Yamanouchi if for any non-virtual Coxeter–Knuth equivalent reduced word \( \sigma \), we have \( \text{des}(\sigma) \geq \text{des}(\rho) \). Denote the set of Yamanouchi reduced words for \( w \) by \( \mathcal{YR}(w) \).

For example, there are two Yamanouchi reduced words for the permutation 153264, namely \( (5, 3, 2, 3, 4) \) and \( (3, 5, 2, 3, 4) \). Notice as well that we may reformulate the Demazure expansion for our running example as

\[
S_{153264} = \kappa_{(0,3,1,0,1)} + \kappa_{(0,3,2,0,0)} = \kappa_{\text{des}(5,3,2,3,4)} + \kappa_{\text{des}(3,5,2,3,4)}.
\]

Moreover, this expansion holds in general, giving the following formula.

Theorem 4.2.3. Given a permutation \( w \), we have

\[
S_w = \sum_{\rho \in \mathcal{YR}(w)} \kappa_{\text{des}(\rho)}.
\]
Proposition 4.2.5. Given two increasing words \( \tau, \sigma \) such that \( \tau \sigma \) is reduced, there is a unique factorization \( \tau = \tau^{(0)} \tau^{(1)} \cdots \tau^{(k+1)} \) and \( \sigma = \sigma^{(1)} x_1 \sigma^{(2)} \cdots x_k \sigma^{(k+1)} \), with some \( \tau^{(i)} \) or \( \sigma^{(i)} \) possibly empty, such that

1. \( \ell(\tau^{(j)}) = \ell(\sigma^{(j)}) \) for \( j = 1, \ldots, k+1 \);
2. \( \tau^{(j)}_i < x_j \) for \( j = 1, \ldots, k \);
3. \( \tau^{(j)}_i > \sigma^{(j)}_i \) for \( j = 1, \ldots, k+1 \) with equality only if \( j > 1 \) and \( \tau^{(j)}_1 = x_{j-1} - 1 \).

We call this factorization the lift alignment since we will lift the unblocked cells \( x_1, \ldots, x_k \) from \( \sigma \) up to \( \tau \) without changing the Coxeter-Knuth equivalence class.
Definition 4.2.6. Given two increasing words $\tau, \sigma$ such that $\tau \sigma$ is reduced, define

\[
\text{lift}(\tau \sigma) = \tau^{(0)} \tau^{(1)} \sigma^{(1)} \tau^{(2)} \sigma^{(2)} \cdots x_k \tau^{(k)} \sigma^{(k)} \rho^{(k)} \sigma^{(k+1)} \rho^{(k+1)} ,
\]

where the factorization is the unique one in Proposition 4.2.5 and for $1 \leq j \leq k$

\[
\tau_i^{(j)} = \begin{cases} 
\tau_i^{(j+1)} - 1 & \text{for } 1 \leq i \leq b_j \\
\tau_i^{(j+1)} & \text{for } b_j + 1 \leq i \leq \ell(\tau^{(j+1)})
\end{cases}
\]

for $b_j = \max\{b \mid \tau_i^{(j+1)} = \tau_i^{(j+1)} = x_j + i \forall 1 \leq i \leq b\}$ if $x_j = \tau_i^{(j+1)} - 1$ or else 0.

For example, lifting the aligned words in Fig. 12 results in the words in Fig. 10

Lemma 4.2.7 (Lift Lemma). Given two increasing words $\tau, \sigma$ such that $\tau \sigma$ is reduced, $\text{lift}(\tau \sigma)$ is Coxeter–Knuth equivalent to $\tau \sigma$.

Proof. With the factorization as denoted in Proposition 4.2.5, consider first the case $k = 1$, in which we must show $\tau^{(0)} \tau^{(1)} \sigma^{(1)} x \sigma^{(2)}$ is Coxeter–Knuth equivalent to $\tau^{(0)} \tau^{(1)} x \tau^{(2)} \sigma^{(1)} \sigma^{(2)}$. Let $\ell_j = \ell(\tau^{(j)})$ for $j = 0, 1, 2$. Since $\tau$ is increasing and $\tau^{(2)}_i > \sigma^{(1)}_i$, we can apply Lemma 4.1.4 $\ell_1$ times to move letters of $\sigma^{(1)}$ left so that $\tau \sigma$ is Coxeter–Knuth equivalent to

\[
\tau^{(0)} \tau^{(1)} \sigma^{(1)} \cdots \tau^{(1)} \sigma^{(1)} \tau^{(2)} x \sigma^{(2)}.
\]

If $\ell_2 = 0$, then since $\sigma^{(1)}_i < x$ for all $i$, we may apply Lemma 4.1.4 to $\sigma^{(1)}$ for $i \leq \ell_1$ to deduce Eq. (4.4) is Coxeter–Knuth equivalent to $\tau^{(0)} \tau^{(1)} x \sigma^{(1)}$ as desired.

If $\ell_2 > 0$ and $\tau^{(2)}_1 > \sigma^{(2)}_1$, then by Proposition 4.2.5 (3), $x < \sigma^{(2)}_1 < \tau^{(2)}_1$, and so $\partial_2(\tau^{(2)}_1 x \sigma^{(2)}_1) = x \tau^{(2)}_1 \sigma^{(2)}_1$. Applying Lemma 4.1.4 $\ell_2$ times to $x$ and the letters of $\sigma^{(2)}$, then $\ell_1$ times to the letters of $\sigma^{(1)}$ shows Eq. (4.1) is Coxeter–Knuth equivalent to $\tau^{(0)} \tau^{(1)} x \tau^{(2)} \sigma^{(1)} \sigma^{(2)}$, as desired.

Finally, suppose $\ell_2 > 0$ and $\tau^{(2)}_1 = \sigma^{(2)}_1 = x + i$ for $i \leq b$ with $b > 0$ maximal. By Proposition 4.2.5 (3) and the increasing property of $\tau$, we have $\sigma^{(2)}_i \leq \tau^{(2)}_i < \tau^{(2)}_{i+1}$. Thus applying Lemma 4.1.4 shows Eq. (4.4) is Coxeter–Knuth equivalent to

\[
\tau^{(0)} \tau^{(1)} \sigma^{(1)} \cdots \tau^{(1)} \sigma^{(1)} \tau^{(2)} \tau^{(2)} \sigma^{(2)} \cdots \tau^{(2)} \sigma^{(2)}.
\]

Based on the assumptions on $b$, we have the elementary relations

\[
\partial_2(\tau^{(2)}_1 \sigma^{(2)}_1) = \begin{cases} 
\sigma^{(2)}_1 i - 1 & \text{if } i > b, \\
\sigma^{(2)}_1 i - 1 & \text{if } i \leq b
\end{cases}
\]

and $\partial_2(\tau^{(2)}_1 \sigma^{(2)}_1) = x \tau^{(2)}_1 \sigma^{(2)}_1 = x \tau^{(2)}_1 \sigma^{(2)}_1 - 1$. Thus Eq. (4.5) is Coxeter–Knuth equivalent to

\[
\tau^{(0)} \tau^{(1)} \sigma^{(1)} \cdots \tau^{(1)} \sigma^{(1)} \tau^{(2)} \sigma^{(2)} \cdots \tau^{(2)} \sigma^{(2)}.
\]

Notice that the letters of $\sigma^{(2)}$ have now been decremented to letters of $\hat{\sigma}^{(2)}$. Finally, apply Lemma 4.1.4 $\ell_2$ times to the letters of $\hat{\sigma}^{(2)}$, then $\ell_1$ times to the letters of $\sigma^{(1)}$ to see this is Coxeter–Knuth equivalent to $\tau^{(0)} \tau^{(1)} x \tau^{(2)} \sigma^{(1)} \sigma^{(2)}$, as desired.

For $\ell > 1$, the above case shows $\tau \sigma$ is Coxeter–Knuth equivalent to

\[
\tau^{(0)} \tau^{(1)} \tau^{(2)} \cdots \tau^{(k)} \sigma^{(1)} \hat{\sigma}^{(2)} \sigma^{(3)} \cdots \sigma^{(k+1)}.
\]

However, notice that this new pair factors uniquely by combining $\tau^{(0)}$ and $\tau^{(1)}$, combining $\tau^{(2)} \cdots \tau^{(1)} \sigma^{(1)}$, $x$, and $\hat{\sigma}^{(2)}$, and combining $\sigma^{(1)}$ and $\sigma^{(2)}$. Therefore there are fewer factors in the result, so by induction on $k$, we may drop the remaining $x_2, \ldots, x_k$ as well, completing the proof. \[\square\]
Extending Definition 4.2.6, we define $\text{lift}_i$ for a reduced word as follows.

**Definition 4.2.8.** Let $\rho$ be a reduced word, and let $(\rho^{(k)}|\cdots|\rho^{(1)})$ be an increasing decomposition of $\rho$. Define $\text{lift}_i(\rho)$ by replacing $\rho^{(i+1)}\rho^{(i)}$ with $\text{lift}(\rho^{(i+1)}\rho^{(i)})$.

![Figure 16. The lifting algorithm applied to an increasing Young tableau.](image)

Reversing the symmetric situation, we can define an *increasing key tableau* with the corresponding row condition and such that the result is lift-invariant.

**Definition 4.2.9.** A key tableau $T$ is *increasing* if the rows are strictly increasing (left to right) and $\text{lift}_i(T) = T$ for all $i$.

Unlike the case for $\text{drop}_i$, we do not wish to apply the maps $\text{lift}_i$ indiscriminantly until reaching some increasing key tableau. Foremost among the reasons is that these maps do not, in general, satisfy the nil-Hecke relations. For example, Fig. 16 shows all nontrivial $\text{lift}_i$ operators applied the leftmost tableau $P$, which is the unique increasing Young tableau in the Coxeter–Knuth class. Notice

$$\text{lift}_1 \circ \text{lift}_2 \circ \text{lift}_1(P) = \text{lift}_1(P) \neq \text{lift}_1 \circ \text{lift}_2(P) = \text{lift}_2 \circ \text{lift}_1 \circ \text{lift}_2(P).$$

Furthermore, both the third and fourth tableaux (from the left) in the middle row of Fig. 16 are increasing, though only the fourth is Yamanouchi.

Nevertheless, we do have a *canonical lifting path* from the unique increasing Young tableau to the unique Yamanouchi tableau. To define this path, we say that $\text{lift}_i$ acts *faithfully* on a tableau $T$ if $\text{lift}_i(T) \neq T$, and we extend this notion to a sequence of lifting maps in the obvious way.

**Lemma 4.2.10.** Let $P$ be an increasing Young tableau with reduced reading word, and suppose each operator in $\text{lift}_{i_k} \circ \cdots \circ \text{lift}_{i_1}(P)$ acts faithfully. Then

$$\text{Des}(\text{lift}_{i_k} \circ \cdots \circ \text{lift}_{i_1}(P)) = s_{i_k} \cdots s_{i_1} \cdot \text{Des}(P),$$

where $s_i$ acts on compositions by interchanging parts in positions $i$ and $i + 1$. 
Proof. In the notation of Definition \ref{def:EDG} since $\ell(\tau(i)) = \ell(\sigma(i))$ for all $i > 0$, raising the $x_i$’s from $\sigma$ to $\tau$ in lift($\sigma$) precisely exchanges the two parts of Des($\sigma$) provided $\tau(0) = \emptyset$. Considering $P$, since columns strictly increase bottom to top, for any entry $z$ in row $i$, there are at least as many $x < z$ in row $i$ as there are $y < z$ in row $j > i$. In particular, taking $j = i + 1$, ensuring that $\tau(0)$ must be empty when applying lift. Furthermore, when $\tau(0) = \emptyset$, entries that move upward maintain their columns, and only larger entries move left. Therefore this property is maintained throughout lifting, ensuring $\tau(0) = \emptyset$ at each step. \hfill $\Box$

For $i \leq j$, define the lifting sequence lift$_{[i,j]}$ by
\begin{equation}
\text{lift}_{[i,j]} = \text{lift}_{j} \circ \text{lift}_{j-1} \circ \cdots \circ \text{lift}_{i}.
\end{equation}

We say that a lifting sequence lift$_{[i,j]}$ acts faithfully on a tableau $T$ if lift$_i$ acts faithfully on $T$ and lift$_k$ acts faithfully on lift$_{[i,k-1]}(T)$ for all $i < k \leq j$.

**Definition 4.2.11.** For $P$ an increasing Young tableau whose row reading word is reduced, define the lift of $P$, denoted by lift($P$), to be the tableau of key shape constructed as follows. Set $T_0 = P$, and for $k > 0$,
\begin{enumerate}
\item if lift$_i(T_{k-1}) = T_{k-1}$ for all $i$, then lift($P$) = $T_{k-1}$;
\item otherwise, set $T_k = \text{lift}_{[i_k,j_k]}(T_{k-1})$ where
\begin{enumerate}
\item $j_k$ is the maximum $j$ for which there exists $i \leq j$ such that lift$_{[i,j]}$ acts faithfully on $T_{k-1}$, and
\item $i_k$ is the minimum $i \leq j_k$ for which lift$_{[i,k]}$ acts faithfully on $T_{k-1}$.
\end{enumerate}
\end{enumerate}

For example, with $P$ the leftmost tableau in Fig. 10, we have
$$\text{lift}(P) = \text{lift}_{[1,1]} \circ \text{lift}_{[2,3]}(P) = \text{lift}_1 \circ \text{lift}_3 \circ \text{lift}_2(P)$$
which is the rightmost tableau in Fig. 10 and is Yamanouchi.

**Theorem 4.2.12.** For $\rho$ a reduced word with $D(\rho)$ an increasing Young tableau, the word lift($\rho$) is Yamanouchi.

**Proof.** By Theorem \ref{thm:EDG}, there is a des-preserving isomorphism, say $\theta$, from the Coxeter–Knuth class of $\rho$ to SKT($a$) for some weak composition $a$. Moreover, by Theorem \ref{thm:SKT} and Proposition \ref{prop:sort}, we must have sort($a$) = Des($\rho$).

Given any $T$ in SKT($a$) for which des($T$) rearranges the parts of $a$, if $S$ in SKT($a$) has des($S$) = des($T$), then $S = T$, which is to say that for each weak composition $b$ that rearranges $a$, there is at most one element of SKT($a$) with weak descent composition $b$. In particular, $\theta(\rho)$ is determined as is $\theta(\sigma)$ for any $\sigma$ obtained from $\rho$ by a sequence of lifts by Lemma \ref{lem:lift}. Moreover, by Lemma \ref{lem:lift} we may extend the maps lift$_i$ to those SKT($a$) whose weak descent compositions rearrange $a$ so that they intertwine with the isomorphism $\theta$. Thus it suffices to show that Definition \ref{def:EDG} applied via $\theta$ to SKT($a$) gives the Yamanouchi standard key tableau whose row reading word is the reverse of the identity.

Let $P = \theta(\rho) \in$ SKT($a$), and set $\lambda =$ Des($P$) say with length $\ell$. Let $\alpha$ be the composition obtained by removing zero parts of $\alpha$. After the first pass of Definition \ref{def:EDG} applied to $P$, we have $T_1 = $ lift$_{[i_1,j_1]}(P) \in$ SKT($a$) where $j_1$ is the largest index $j$ for which $\lambda_j \neq \alpha_j$, but $\lambda_k = \alpha_k$ for all $k > j$, and $i_1$ is the largest index $i < j_1$ for which $\lambda_i = \alpha_{j_1}$. Therefore Des($T_1$) = $s_{j_1} \cdots s_{i_1} \cdot \lambda$ agrees with $\alpha$ in all positions $j \geq j_1$. Furthermore, reading the rows of $T_1$ left to right from the top down to $j_1$, we precisely have the reverse of the identity. Thus we may proceed by induction on the first $j_1 - 1 < \ell$ rows of $T_1$. \hfill $\Box$
For an example of the maps \( \text{lift}_i \) induced on elements of \( \text{SKT}(a) \) whose weak descent compositions rearrange \( a \), see Fig. 17. Compare this with Fig. 16.

![Figure 17](image)

For example, lifting the increasing Young tableaux in Figure 8 we arrive at the Yamanouchi key tableaux in Figure 9, as promised.

**Corollary 4.2.13.** For \( w \) a permutation, we have

\[
S_w = \sum_{\rho \in R(w)} k_{\text{des}}(\text{lift}(\rho)) \cdot \text{D(\rho) increasing}
\]

For example, lifting the increasing Young tableaux in Figure 8 we arrive at the Yamanouchi key tableaux in Figure 9 as promised.

5. **Insertion Algorithms**

Edelman and Greene [10] define an insertion algorithm mapping reduced words to pairs of Young tableaux where the left is increasing and the right is standard. In this context, the left tableau gives the canonical Coxeter–Knuth equivalence class representative for obtaining the Schur expansion of a Stanley symmetric function, and the right tableau gives an explicit bijection between elements of the Coxeter–Knuth equivalence class and standard Young tableaux of fixed shape. We recall their definitions and main results for the purpose of generalizing them to the polynomial setting. In the generalization, the left tableau will be a Yamanouchi key tableau, and the right tableau will be a standard key tableau. Thus the left tableau gives the canonical Coxeter–Knuth equivalence class representative for obtaining the Demazure expansion of a Schubert polynomial, and the right tableau gives an explicit bijection between elements of the Coxeter–Knuth equivalence class and standard key tableaux of fixed shape.
5.1. **Edelman–Greene insertion.** Edelman and Greene [10] (Definition 6.20) defined the following procedure for inserting a letter into an increasing tableau.

**Definition 5.1.1 ([10]).** Let \( P \) be an increasing Young tableau, and let \( x \) be a positive integer. Let \( P_i \) be the \( i \)th lowest row of \( P \). Define the *Edelman–Greene insertion* of \( x \) into \( P \), denoted by \( P \leftarrow x \), as follows. Set \( x_0 = x \) and for \( i \geq 0 \), insert \( x_i \) into \( P_{i+1} \) as follows: if \( x_i > z \) for all \( z \in P_{i+1} \), place \( x_i \) at the end of \( P_{i+1} \) and stop; otherwise, let \( x_{i+1} \) denote the smallest element of \( P_{i+1} \) such that \( x_{i+1} > x_i \) (we say that \( x_i \) *bumps* \( x_{i+1} \) in row \( i+1 \)), replace \( x_{i+1} \) by \( x_i \) in \( P_{i+1} \) only if \( x_{i+1} \neq x_i + 1 \) or \( x_i \) is not already in \( P_{i+1} \), and continue.

This algorithm generalizes the insertion algorithm of Schensted [20], building on work of Robinson [19], later generalized by Knuth [12]. Robinson–Schensted insertion becomes a bijective correspondence between permutations and pairs of standard Young tableaux by constructing a second tableau to track the order in which new cells are added. The pair is typically denoted by \((P, Q)\), where \( P \) is called the *insertion tableau*, and \( Q \) is called the *recording tableau*.

Similarly, we construct the Edelman–Greene correspondence of a reduced word \( \rho = (\rho_k, \ldots, \rho_1) \) by successively inserting the letters of \( \rho \) from \( k \) to \( 1 \) into the empty tableau to create the *Edelman–Greene insertion tableau* of \( \rho \), denoted by \( P(\rho) \). For example, Figure 18 shows the Edelman-Greene insertion tableau for the reduced word \( \rho = (3, 6, 4, 7, 5, 2, 4) \).

```
\emptyset \leftarrow 3 \leftarrow 3 \ 6 \leftarrow 4 \leftarrow 6 \leftarrow 7 \leftarrow 3 \leftarrow 3 \leftarrow 3 \leftarrow 4 \leftarrow 6 \leftarrow 7
```

**Figure 18.** The Edelman-Greene insertion tableau for the reduced word \( \rho = (3, 6, 4, 7, 5, 2, 4) \).

Since Edelman–Greene insertion adds a single cell to an existing Young diagram, when inserting \( \rho_i \), we create the *Edelman–Greene recording tableau* of \( \rho \), denoted by \( Q(\rho) \), by adding a cell with entry \( k - i + 1 \) into the position of \( P(\rho_k, \ldots, \rho_i) \backslash P(\rho_k, \ldots, \rho_{i+1}) \). For example, Figure 19 shows the Edelman-Greene recording tableau for the reduced word \( \rho = (3, 6, 4, 7, 5, 2, 4) \).

```
\emptyset \ 1 \ 1 \ 1 \ 1 \ 1 \ 2 \ 1 \ 2 \ 1 \ 2 \ 4 \ 3 \ 5 \ 3 \ 5 \ 6 \ 6 \ 7
```

**Figure 19.** The Edelman-Greene recording tableau for the reduced word \( \rho = (3, 6, 4, 7, 5, 2, 4) \).

Edelman and Greene derived many properties of this generalized insertion algorithm, including that the insertion tableau \( P(\rho) \) defined by inserting \( \rho_k, \ldots, \rho_1 \) into the empty tableau is a well-defined increasing tableau whose row reading word is a reduced word for \( w \). They proved the following [10] (Theorem 6.24) relating their insertion to Coxeter–Knuth equivalence.
Theorem 5.1.2 ([10]). For reduced words \( \sigma, \tau \), we have \( P(\sigma) = P(\tau) \) if and only if \( \sigma \) and \( \tau \) are Coxeter-Knuth equivalent.

Theorem 5.1.2 gives canonical Coxeter–Knuth equivalence class representatives as the reduced words occurring as reading words of increasing Young tableaux. Thus the result follows as well from Theorem 4.1.10.

Further, Edelman and Greene characterize how the recording tableaux differ for two reduced words that differ by an elementary Coxeter–Knuth equivalence. Refining [10] (Definition 6.14), we have the following definition from [4].

Definition 5.1.3. The elementary dual equivalence involutions, denoted by \( d_i \), act on standard Young tableaux by

\[
d_i(T) = \begin{cases} 
    s_{i-1} \cdot T & \text{if } i+1 \text{ lies between } i \text{ and } i-1 \text{ in } \text{word}(T), \\
    s_i \cdot T & \text{if } i-1 \text{ lies between } i \text{ and } i+1 \text{ in } \text{word}(T), \\
    T & \text{if } i \text{ lies between } i-1 \text{ and } i+1 \text{ in } \text{word}(T),
\end{cases}
\]

where \( s_i \) acts by interchanging \( i \) and \( i+1 \), and \( \text{word}(T) \) is the row reading word.

For example, the elementary dual equivalence involutions on standard Young tableaux of shape \( (3, 2) \) are shown in Fig. 20.

![Figure 20. The elementary dual equivalence involutions on SYT(3, 2).](image)

Edelman and Greene [10] (Corollary 6.15) relate Coxeter–Knuth equivalence with dual equivalence through the recording tableaux as follows.

Theorem 5.1.4 ([10]). For reduced words \( \sigma, \tau \), we have \( Q(\sigma) = d_i(Q(\tau)) \) if and only if \( \sigma = d_{n-i+1}(\tau) \).

Theorem 5.1.4 follows from Theorem 3.1.4, proving that Edelman–Greene insertion establishes a Des-preserving bijection between elements of a Coxeter–Knuth equivalence class and standard Young tableaux of fixed shape.

Edelman and Greene use their insertion and recording tableaux to establish the following bijective correspondence [10] (Theorem 6.25).

Corollary 5.1.5 ([10]). The Edelman–Greene correspondence \( \rho \rightarrow (P(\rho), Q(\rho)) \) establishes a bijection

\[
R(w) \xrightarrow{\sim} \bigcup_{\lambda} (\text{Inc}_\lambda(w) \times \text{SYT}(\lambda)),
\]

where \( \text{Inc}_\lambda(w) \) is the set of increasing reduced words \( \sigma \) for \( w \) such that \( \text{Des}(\sigma) = \lambda \). Moreover, under this correspondence, \( \text{Des}(\rho) = \text{Des}(Q(\rho)) \).

Taking fundamental quasisymmetric generating functions gives Theorem 3.1.3.
5.2. Weak insertion. We generalize Edelman-Greene insertion to an algorithm on reduced words that outputs a pair tableau of key shape such that the insertion tableau is a Yamanouchi key tableau (in particular, it is increasing) and the recording tableau is a standard key tableau. Leveraging Definition 5.1.1 along with Definitions 4.1.9 and 4.2.11, we have the following.

Definition 5.2.1. For \( P \) a Yamanouchi key tableau and \( x \) a positive integer, define the weak insertion of \( x \) into \( P \), denoted by \( P \triangleleft x \), to be lift \( p \) drop \( p \) \( P \triangleleft q \). Construct the weak correspondence of a reduced word \( \rho = (\rho_k, \ldots, \rho_1) \) by successively inserting the letters of \( \rho \) from \( k \) to 1 into the empty tableau to create the weak insertion tableau of \( \rho \), denoted by \( rPp_{\rho}q \). For example, Figure 21 shows the weak insertion tableau for the reduced word \( \rho = (3, 6, 4, 7, 5, 2, 4) \).

\[
\begin{array}{cccccccc}
\emptyset & 3 & 6 & 4 & 7 & 5 & 2 & 4 \\
3 & 6 & 4 & 7 & 3 & 4 & 7 & 3 & 4 & 5 & 2 & 4
\end{array}
\]

Figure 21. The weak insertion tableau for the reduced word \( \rho = (3, 6, 4, 7, 5, 2, 4) \).

Parallel to Theorem 5.1.2, we have the following.

Theorem 5.2.2. For reduced words \( \sigma, \tau \), we have \( \hat{P}(\sigma) = \hat{P}(\tau) \) if and only if \( \sigma \) and \( \tau \) are Coxeter-Knuth equivalent.

Proof. By Theorem 4.1.10, dropping a word so that the result is an increasing Young tableau maintains the Coxeter–Knuth equivalence class. By Theorem 4.2.12, the Yamanouchi words are constructed by lifting the increasing Young tableaux, and by Theorem 4.2.3 they are the canonical representatives for each weak dual equivalence class. By Theorem 5.2.7, Coxeter–Knuth equivalence classes are weak dual equivalence classes, and so the result follows.

In order to define a weak recording tableau, we must show that the successive shapes when insertion a word are nested. To this end, we have the following.

Lemma 5.2.3. Let \( \rho \) be a Yamanouchi reduced word and \( x \) a letter such that the concatenation \( \rho x \) is reduced, and set \( \sigma = \text{lift}(\text{drop}(\rho x)) \). Then \( \text{des}(\rho)_i \leq \text{des}(\sigma)_i \) for all \( i \), and if \( j \) is the unique index such that \( \text{des}(\sigma)_j = \text{des}(\rho)_j + 1 \), then \( \text{des}(\sigma)_j \neq \text{des}(\rho)_i \) for all \( i < j \).

Proof. By Lemma 4.2.10, the locations of the nonempty rows of \( \hat{P}(\rho) \) are determined by the entries in the first column of \( P(\rho) \), and these are the nonzero entries of \( \text{des}(\text{lift}(\text{drop}(\rho))) \). Thus inserting an additional letter will increase the length of one (possibly empty) row. Therefore \( \text{des}(\rho)_i \leq \text{des}(\sigma)_i \) for all \( i \).

Let \( j \) denote the unique index for which \( \text{des}(\sigma)_j = \text{des}(\rho)_j + 1 \). Suppose, for contradiction, \( \text{des}(\sigma)_j = \text{des}(\rho)_j + 1 \) and \( \text{des}(\rho)_i = 1 \) for some \( i < j \). Note the entries in \( \hat{P}(\sigma) \) in rows \( i, j \) must be \( i, j \), respectively. Furthermore, in \( P(\sigma) \), both \( i \) and \( j \) are singleton cells, so the Edelman–Greene insertion of \( x \) into \( \rho \) must have caused...
i to bump j. In particular, j must have been a singleton cell of P(ρ), but this
contradicts that des(σ)_j = des(ρ)_j + 1. Similarly, if des(σ)_j = des(ρ)_i = c > 1 for
some i < j, then the entry in row j column c must be strictly larger than the entry
in row i column c, and the same bumping argument applies, presenting the same
contradiction. Thus we must have des(σ)_j ≠ des(ρ)_i for all i < j as desired. □

We may now define the weak recording tableau of ρ, denoted by Q(ρ), by adding
a cell with entry i into the position of Q(ρ), . . . , ρ, . . . , ρ, . . . , ρ, . . ., ρ, . . .
for some i, ĵ as desired.

We can characterize how the recording tableaux differ for two reduced words
that differ by an elementary Coxeter–Knuth equivalence using elementary weak
dual equivalences [6] (Definition 3.21).

Definition 5.2.5 [6]. The elementary weak dual equivalence involutions, denoted
by d_i act on standard key tableaux as follows. Let u, v, w be the cells with entries
i − 1, i, i + 1 taken in column reading order. Then

\[ d_i(T) = \begin{cases} 
  b_j(T) & \text{if } u, w \text{ are in the same row and } v \text{ is not,} \\
  s_{i−1}(T) & \text{if } v \text{ has entry } i + 1, \\
  s_i(T) & \text{if } v \text{ has entry } i − 1, \\
  T & \text{otherwise,}
\end{cases} \]

where b_j cycles j − 1, j, j + 1 so that j shares a row with j ± 1.

The elementary weak dual equivalence involutions on standard key tableaux of
shape (0, 3, 0, 2) are shown in Fig. 23.

We relate Coxeter–Knuth equivalence with weak dual equivalence through the
weak recording tableaux as follows.

Theorem 5.2.6. For reduced words σ, τ, we have Q(σ) = d_i(Q(τ)) if and only if
σ = d_i(τ).
Figure 23. The elementary weak dual equivalence involutions on SKT(0, 3, 0, 2).

Proof. By [6] (Theorem 3.24), the bijection Φ : SKT(a) → SYT(sort(a)) that drops entries in a key tableau to partition shape, replaces i with n − i + 1, and sorts columns to increase bottom to top intertwines the elementary weak dual equivalence involutions on standard key tableaux with the elementary dual equivalence involutions on standard Young tableaux by Φ(\tilde{d}_i(T)) = d_{n-i+1}(\Phi(T)). The result now follows from Theorem 5.1.4.

The weak insertion and recording tableaux establish the following bijective correspondence, parallel to Corollary 5.1.5.

Corollary 5.2.7. The weak correspondence ρ → (\tilde{P}(\rho), \tilde{Q}(\rho)) establishes a bijection

\[(5.4) \quad R(w) \overset{\sim}{\longrightarrow} \bigsqcup_{a} (YR_{a}(w) \times \text{SKT}(a)),\]

where YR_{a}(w) is the set of Yamanouchi reduced words σ for w such that des(σ) = a. Moreover, under this correspondence, des(ρ) = des(\tilde{Q}(\rho)).

Taking fundamental slide generating polynomials gives Corollary 4.2.13.

References

[1] Sami Assaf and Dominic Searles. Schubert polynomials, slide polynomials, Stanley symmetric functions and quasi-Yamanouchi pipe dreams. Adv. in Math., 306:89–122, 2017.
[2] Sami Assaf and Dominic Searles. Kohnert tableaux and a lifting of quasi-Schur functions. J. Combin. Theory Ser. A, 156:85–118, 2018.
[3] Sami H. Assaf. Dual equivalence graphs, ribbon tableaux and Macdonald polynomials. PhD thesis, University of California Berkeley, 2007.
[4] Sami H. Assaf. Dual equivalence graphs I: A new paradigm for Schur positivity. Forum Math. Sigma, 3:e12, 33, 2015.
[5] Sami H. Assaf. Multiplication of a Schubert polynomial by a Stanley symmetric polynomial. arXiv:1702.00132, 2017.
[6] Sami H. Assaf. Weak dual equivalence for polynomials. arXiv:1702.04051, 2017.
[7] Sara C. Billey, William Jockusch, and Richard P. Stanley. Some combinatorial properties of Schubert polynomials. J. Algebraic Combin., 2(4):345–374, 1993.
[8] Michel Demazure. Désingularisation des variétés de Schubert généralisées. Ann. Sci. École Norm. Sup. (4), 7:53–88, 1974. Collection of articles dedicated to Henri Cartan on the occasion of his 70th birthday, I.
[9] Michel Demazure. Une nouvelle formule des caractères. Bull. Sci. Math. (2), 98(3):163–172, 1974.
[10] Paul Edelman and Curtis Greene. Balanced tableaux. Adv. in Math., 63(1):42–99, 1987.
[11] Ira M. Gessel. Multipartite P-partitions and inner products of skew Schur functions. In Combinatorics and algebra (Boulder, Colo., 1983), volume 34 of Contemp. Math., pages 289–317. Amer. Math. Soc., Providence, RI, 1984.
[12] Donald E. Knuth. Permutations, matrices, and generalized Young tableaux. Pacific J. Math., 34:709–727, 1970.
[13] Axel Kohnert. Weintrauben, Polynome, Tableaux. Bayreuth. Math. Schr., (38):1–97, 1991. Dissertation, Universität Bayreuth, Bayreuth, 1990.
[14] Alain Lascoux and Marcel-Paul Schützenberger. Polynomes de Schubert. C. R. Acad. Sci. Paris Sér. I Math., 294(13):447–450, 1982.
[15] Alain Lascoux and Marcel-Paul Schützenberger. Keys & standard bases. In Invariant theory and tableaux (Minneapolis, MN, 1988), volume 19 of IMA Vol. Math. Appl., pages 125–144. Springer, New York, 1990.
[16] I. G. Macdonald. Notes on Schubert polynomials. LACIM, Univ. Quebec a Montreal, Montreal, PQ, 1991.
[17] I. G. Macdonald. Symmetric functions and Hall polynomials. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, second edition, 1995. With contributions by A. Zelevinsky, Oxford Science Publications.
[18] Victor Reiner and Mark Shimozono. Key polynomials and a flagged Littlewood-Richardson rule. J. Combin. Theory Ser. A, 70(1):107–143, 1995.
[19] G. de B. Robinson. On the Representations of the Symmetric Group. Amer. J. Math., 60(3):745–760, 1938.
[20] C. Schensted. Longest increasing and decreasing subsequences. Canad. J. Math., 13:179–191, 1961.
[21] Richard P. Stanley. On the number of reduced decompositions of elements of Coxeter groups. European J. Combin., 5(4):359–372, 1984.
[22] Jacques Tits. Le problème des mots dans les groupes de Coxeter. In Symposia Mathematica (INDAM, Rome, 1967/68), Vol. 1, pages 175–185. Academic Press, London, 1969.

Department of Mathematics, University of Southern California, 3620 South Vermont Avenue, Los Angeles, CA 90089-2532, U.S.A.

E-mail address: shassaf@usc.edu