Schubert polynomials

Jasper Ty

These are notes based on my study of Schubert polynomials. My main references are [MacdonaldSP] and [KnutsonSP].

Contents

1 Notation and conventions

1.1 Sets

We take $\mathbb{N}$ to be the set of natural numbers *including* zero,

$$\mathbb{N} := \{0, 1, 2, \ldots\}.$$

We take $\mathbb{P}$ to be the set of *positive integers*,

$$\mathbb{P} := \{1, 2, \ldots\}.$$

$\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ are defined as usual.

We denote the set $\{1, \ldots, n\}$ by $[n]$.

1.2 Partitions and compositions

A *weak composition* α of $n \in \mathbb{N}$ is an infinite tuple of nonnegative integers

$$(\alpha_1, \alpha_2, \ldots)$$

such that $\sum_i \alpha_i = n$. We define $|\alpha| = \sum_i \alpha_i$ to have notation for recovering n given α.

A *partition* λ of n is a weak composition whose entries are *weakly decreasing*. That a particular partition λ is a partition of a particular n is denoted $\lambda \vdash n$. We define $|\lambda|$ the exact same way.

I use English notation when drawing diagrams and tableaux, meaning, row index increases *north to south*, and column index increases *west to east*.

1.3 Rings, polynomials, and formal power series

The following notation is (mostly) in accordance with the notation in [GrinbergAC], with a few additions.

All rings considered are commutative and unital. An arbitrary ring will be denoted $\mathbb{K}$.

$\mathbb{K}[[t]]$ will denote the formal power series ring over $\mathbb{K}$ in the indeterminate t.

We will fix notation for the following sets of indeterminates, which we will use when convenient:

(a) $X_N := (x_1, x_2, \ldots x_N)$ for a set of N indeterminates.

(b) $X := (x_1, x_2, \ldots)$ for a set of countably many indeterminates.

(c) Y, Y_N, Z, Z_N, Q, Q_N and so on are defined similarly.

With compositions, partitions, or otherwise any finitely supported tuple of non-negative integers α, we define *multi-index notation* for compactly writing down monomials.

$$x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} \cdots .$$

We will let $[x^\alpha]f$ denote the coefficient of $[x^\alpha]$ in the polynomial or formal power series f.

1.4 Permutations and the symmetric group

S_n will denote the symmetric group on n letters.

I use cycle notation, so e.g the cycle that sends 1 to 7, 7 to 4, and 4 to 1 will be written as (174).

The simple transpositions $(i\ i+1)$ will be denoted s_i.

The identity permutation will be denoted 1.

The length of a permutation w will be denoted $\ell(w)$.

Permutations will act on polynomials or power series by permuting *places*, meaning that if $\sigma \in S_n$ and $f(x_1, \dots, x_n) \in \mathbb{K}[X_n]$, we define

$$\sigma f(x_1, \dots, x_n) := f(x_{\sigma(1)}, \dots x_{\sigma(n)}).$$

2 Permutations

We recall here relevant tidbits about permutations.

Definition 2.0.1. Let $w \in S_n$. An *inversion* of w is a pair $i < j$ such that $w(i) > w(j)$. The *inversion number* of w is the number of inversions of w, and we denote this with $\ell(w)$.

We note that it's particularly easy to see that $\ell(w)$ is well-defined (as the size of a well-defined subset of $[n] \times [n]$).

It gives us, then, an easy way to define the simplest, most famous permutation statistic:

Definition 2.0.2. We define the *sign* of a permutation w to be

$$(-1)^w := (-1)^{\ell(w)}.$$

This coincides with more typical definitions.

Remark 2.0.3. Let $w \in S_n$. The quantity $(-1)^{\ell(w)}$ agrees with the following

(a) $\operatorname{sgn}(w)$, where sgn is the usual *sign homomorphism* $\operatorname{sgn} : S_n \to \{-1, 1\}$.

(b) $(-1)^k$, where k is the length of *any* decomposition of w into a product of transpositions.

Proof. See section 5.4 in [GrinbergAC]. □

We happen to be interested in a particular kind of decomposition of a permutation w as a product of transpositions:

Definition 2.0.4. Let $w \in S_n$. A *Coxeter word* for w is a sequence of simple transpositions $s_{i_1}, \ldots, s_{i_k}$ such that

$$w = s_{i_1} \cdots s_{i_k}.$$

We call a Coxeter word a *reduced word* if it's of minimal length, that is, there is no shorter Coxeter word for w.

The following theorem is important, and has a detailed proof, as Theorem 5.3.17, in [GrinbergAC].

Theorem 2.0.5. Let $w \in S_n$. Then there exist Coxeter words for w, and their minimal length is $\ell(w)$, i.e reduced words for w have length $\ell(w)$.

Proof (sketch). We kill one and a half birds with one stone by first showing existence of Coxeter words for w with length $\ell(w)$. The remaining half a bird is showing that it is a reduced word.

The key fact is that simples s_i, when multiplied on the right, either increment or decrement the inversion number— if $(i, i+1)$ is an inversion, then s_i *deletes* it, otherwise, s_i *creates* an inversion $(i, i+1)$.

This makes existence amenable to proof by induction on $\ell(w)$.

For the base case, the only permutation w with $\ell(w) = 0$ is the identity permutation, a product of zero simples.

For the induction step, let w be a permutation and suppose $\ell(w) = h > 0$ and assume (induction hypothesis) existence of Coxeter words for all permutations w' where $\ell(w') = h - 1$. Then we hit w with a simple s_k that cancels out one of its inversions.

Then $\ell(ws_k) = \ell(w) - 1 = h - 1$, so there exists a Coxeter word $s_{i_1} \cdots s_{i_{h-1}}$ for ws_k. Then $s_{i_1} \cdots s_{i_{h-1}} s_k$ is a Coxeter word of length $h = \ell(w)$ for w.

The fact that we have a reduced word follows from s_i's at most only incrementing inversion number— you can't get $\ell(w) = h$ with fewer than h simples! □

Then, what do we know about $\ell(w)$?

Definition 2.0.6. Let

$$w_0 := n, n-1, \ldots, 1.$$

Equivalently, it's the permutation that maximizes the number of inversions, which happens to be

$$\ell(w_0) = \frac{n(n-1)}{2}.$$

Theorem 2.0.7. The simple transpositions satisfy the relations

(a) Braid relation

$$s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1} \tag{1}$$

(b) Far commutativity

$$s_i s_j = s_j s_i \qquad \text{whenever} \qquad |i - j| > 1 \tag{2}$$

(c) Contraction

$$s_i^2 = 1 \tag{3}$$

3 Schubert Polynomials

Apparently these are "Schubert cycles in flag varieties".

3.1 Divided difference operators

These strike me as a tool to measure how "unsymmetric" a polynomial is in a local sense, in two variables at a time.

3.1.1 Definition

Definition 3.1.1. Let f be a polynomial. We define the *divided difference operator* ∂_i by

$$\partial_i f := \frac{f - s_i f}{x_i - x_{i+1}} \tag{4}$$

Example 3.1.2. If $f(x_1, x_2, x_3) = x_1 x_2$, then

$$\begin{aligned} \partial_2 f(x_1, x_2, x_3) &= \frac{x_1 x_2 - x_1 x_3}{x_2 - x_3} \\ &= x_1 \left(\frac{x_2 - x_3}{x_2 - x_3} \right) \\ &= x_1 . \end{aligned}$$

3.1.2 Basic facts

We have the following characterization of ∂_i that does not invoke division.

Lemma 3.1.3. Fix i. Consider some monomial $f = \cdots x_i^a x_{i+1}^b \cdots$. Then

$$\partial_i(\cdots x_i^a x_{i+1}^b \cdots) = \varepsilon_{ba} \sum_{\substack{u,v \geq \min\{a,b\} \\ u+v=a+b-1}} \cdots x_i^u x_{i+1}^v \cdots ,$$

where ε is defined to be

$$\varepsilon_{rs} := \begin{cases} 0 & \text{if } r = s \\ 1 & \text{if } r < s \\ -1 & \text{if } r > s \end{cases} .$$

Proof. The proof is not hard but it's a slog. We compute

$$\begin{aligned} \partial_i(\cdots x_i^a x_{i+1}^b \cdots) &= \frac{(\cdots x_i^a x_{i+1}^b \cdots) - (\cdots x_i^b x_{i+1}^a \cdots)}{x_i - x_{i+1}} \\ &= (\cdots) \frac{x_i^a x_{i+1}^b - x_i^b x_{i+1}^a}{x_i - x_{i+1}} . \end{aligned}$$

Recall that in any commutative ring we have that

$$\frac{x^n - y^n}{x - y} = x^{n-1} y^0 + x^{n-2} y^1 + \cdots + x^1 y^{n-2} + x^0 y^{n-1},$$

which we will modify a little

$$\frac{x^{n+m}y^m - x^m y^{n+m}}{x-y} = x^{m+n-1}y^m + x^{m+n-2}y^{m+1} + \cdots + x^{m+1}y^{m+n-2}x^m y^{m+n-1},$$

and we note that the pairs $(u, v) \in \{(m+n-1, m), \ldots, (m, m+n-1)\}$ are precisely those such that $u, v \geq \min\{a, b\}$ and $u + v = 2m + n - 1$. We then put $a = m + n$ and $b = m$, to get that

$$\frac{x^a y^b - x^a y^b}{x-y} = \sum_{\substack{u,v \geq \min\{a,b\} \\ u+v=a+b-1}} x^u y^v, \qquad \text{given } a \geq b.$$

Then, to forget $a \geq b$, we pick up a ε_{ba} term to keep track of sign. Applying this identity now to our computation, we finish the lemma. □

Then the following properties of the operator ∂_i can be read off

Corollary 3.1.4. Let f be a polynomial.

(a) $\partial_i f$ is a polynomial.

(b) If f is homogeneous of degree d, then $\partial_i f$ is homogeneous of degree $d - 1$.

Proof. Left to reader. □

The following theorem gives us an analogy between the divided difference operators and the simple transpositions. In particular, it tells us that sequences of ∂_i's structurally behave like reduced words when the corresponding sequence of s_i's are reduced words (see Definition 3.1.6 and Theorem 3.1.7), but that the ∂_i's degenerate and collapse to nothing in the case for non-reduced words (see Theorem 3.1.9).

Theorem 3.1.5. The divided difference operators satsify, in analogy to Theorem 2.0.7,

(a) The braid relation

$$\partial_i \partial_{i+1} \partial_i = \partial_{i+1} \partial_i \partial_{i+1} \tag{5}$$

(b) Far commutativity

$$\partial_i \partial_j = \partial_j \partial_i \qquad \text{whenever} \qquad |i - j| > 1 \tag{6}$$

(c) Annihilation (not contraction!)

$$\partial_i^2 = 0 \tag{7}$$

Proof (sketch). For (a), without loss of generality we prove the case

$$\partial_1 \partial_2 \partial_1 = \partial_2 \partial_1 \partial_2.$$

Which we just have to grind out (see Appendix).

It turns out that both sides equal

$$\frac{1 - s_1 - s_2 + s_1 s_2 + s_2 s_1 - s_1 s_2 s_1}{(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)}.$$

The proofs of (b), (c) are straightforward. □

Note that the numerator appearing in the proof happens to be

$$\nabla^- := \sum_{w \in S_3} (-1)^w w,$$

the antisymmetrizer of $\mathbb{Z}[S_3]$.

Given that the definition of ∂_i takes in some s_i as an input, we can naturally come up with a broader definition of ∂ that takes in Coxeter words.

Definition 3.1.6. Let $w \in S_n$, and let $a = (a_1, \ldots a_k)$ be a Coxeter word for w, i.e $k = \ell(w)$ and $s_{a_1} \ldots s_{a_k} = w$. Then define

$$\partial_a := \partial_{a_1} \ldots \partial_{a_k}.$$

We'll actually use this to bootstrap another definition– divided difference operators parametrized by permutations. That doesn't quite come for free, so we need to first prove the following fact:

Theorem 3.1.7. Let $w \in S_n$. If $a = (a_1, \ldots, a_k)$ and $b = (b_1, \ldots, b_k)$ are reduced words for w, then $\partial_a = \partial_b$.

Proof. This follows from the fact any two reduced words for a permutation w are equivalent modulo far commutativity and the braid relation. Then recall Theorem 3.1.5— Equations 5 and 6 tell us exactly that the divided difference operators also satisfy those relations. □

So we can now properly define the following:

Theorem 3.1.8. Let $w \in S_n$, and let $a = (a_1, \ldots, a_k)$ be some reduced word for w. Then define

$$\partial_w := \partial_a = \partial_{a_1} \ldots \partial_{a_k}.$$

In the case for sequences that *do not* correspond to reduced words, we have the following reason to not really care about them:

Theorem 3.1.9. Let $a = (a_1, \ldots, a_k)$ be a sequence that is not a reduced word for any $w \in S_n$. Then

$$\partial_a = 0.$$

Proof. Because a is not a reduced word, it is possible to do a sequence of moves on the Coxeter word which contains a contraction. Mapping these moves over to the divided difference word results in an application of Equation 7, killing the whole term. □

3.2 The definition of a Schubert polynomial

Definition 3.2.1. The *Schubert polynomials* $\mathfrak{S}_w$ are defined by the rules

$$\begin{cases} \mathfrak{S}_{w_0} := x_1^{n-1} x_2^{n-2} \cdots x_{n-1}^1, \\ \partial_i \mathfrak{S}_w := \mathfrak{S}_{ws_i} \end{cases}$$

Actually, this definition is a theorem if we start with the "Representatives of cohomology classes of Schubert cycles in flag varieties" definition, but I don't understand that unfortunately.

4 The ring of coinvariants of

Theorem 4.0.1. The Schuberts form a basis for the coinvariant ring

4.1 Definition

5 Appendix

Detailed proof of Theorem 3.1.5. Define $[ij] := x_i - x_j$. We have the following relations:

$$\begin{aligned} s_1[12] &= -[12] & s_2[12] &= [13] \\ s_1[13] &= [23] & s_2[13] &= [12] \\ s_1[23] &= [13] & s_2[23] &= -[23] \end{aligned}$$

First, we expand the left hand side, which is

$$\partial_1 \partial_2 \partial_1 = \left(\frac{1 - s_1}{[12]}\right)\left(\frac{1 - s_2}{[23]}\right)\left(\frac{1 - s_1}{[12]}\right).$$

We do the first application, which is the ∂_2 hitting the ∂_1,

$$\begin{aligned} \partial_2 \partial_1 &= \left(\frac{1 - s_2}{[23]}\right)\left(\frac{1 - s_1}{[12]}\right) \\ &= \left(\frac{\left(\frac{1-s_1}{[12]}\right) - s_2\left(\frac{1-s_1}{[12]}\right)}{[23]}\right), \end{aligned}$$

then we apply the s_2,

$$\begin{aligned} &= \left(\frac{\left(\frac{1-s_1}{[12]}\right) - s_2\left(\frac{1-s_1}{[12]}\right)}{[23]}\right) \\ &= \left(\frac{\frac{1-s_1}{[12]} - \frac{s_2 - s_2 s_1}{s_2[12]}}{[23]}\right) \\ &= \left(\frac{\frac{1-s_1}{[12]} - \frac{s_2 - s_2 s_1}{[13]}}{[23]}\right) \\ &= \left(\frac{1 - s_1}{[12][23]} - \frac{s_2 - s_2 s_1}{[13][23]}\right). \end{aligned}$$

Now we apply ∂_1 to our just computed $\partial_2\partial_1$,

$$
\begin{aligned}
\partial_1(\partial_2\partial_1) &= \left(\frac{1-s_1}{[12]}\right)\left(\frac{1-s_1}{[12][23]} - \frac{s_2-s_2s_1}{[13][23]}\right)\\
&= \left(\frac{\left(\frac{1-s_1}{[12][23]} - \frac{s_2-s_2s_1}{[13][23]}\right) - s_1\left(\frac{1-s_1}{[12][23]} - \frac{s_2-s_2s_1}{[13][23]}\right)}{[12]}\right)\\
&= \left(\frac{\frac{1-s_1}{[12][23]} - \frac{s_2-s_2s_1}{[13][23]} - \frac{s_1-s_1s_1}{s_1[12]s_1[23]} + \frac{s_1s_2-s_1s_2s_1}{s_1[13]s_1[23]}}{[12]}\right)\\
&= \left(\frac{\frac{1-s_1}{[12][23]} - \frac{s_2-s_2s_1}{[13][23]} - \frac{s_1-1}{(-[12])[13]} + \frac{s_1s_2-s_1s_2s_1}{[23][13]}}{[12]}\right)\\
&= \left(\frac{\frac{1-s_1}{[12][23]} - \frac{s_2-s_2s_1}{[13][23]} - \frac{1-s_1}{[12][13]} + \frac{s_1s_2-s_1s_2s_1}{[23][13]}}{[12]}\right)\\
&= \left(\frac{\frac{1-s_1}{[12][23]} - \frac{1-s_1}{[12][13]}}{[12]}\right) + \left(\frac{-\frac{s_2-s_2s_1}{[13][23]} + \frac{s_1s_2-s_1s_2s_1}{[23][13]}}{[12]}\right)\\
&= \left(\frac{1-s_1}{[12]^2[23]} - \frac{1-s_1}{[12]^2[13]}\right) + \left(\frac{-(s_2-s_2s_1)+s_1s_2-s_1s_2s_1}{[12][13][23]}\right)\\
&= \left(\frac{1-s_1}{[12]^2}\left(\frac{1}{[23]} - \frac{1}{[13]}\right)\right) + \left(\frac{-s_2+s_2s_1+s_1s_2-s_1s_2s_1}{[12][13][23]}\right)\\
&= \left(\frac{1-s_1}{[12]^2}\left(\frac{[13]-[23]}{[23][13]}\right)\right) + \left(\frac{-s_2+s_2s_1+s_1s_2-s_1s_2s_1}{[12][13][23]}\right)\\
&= \left(\frac{1-s_1}{[12]^2}\left(\frac{[12]}{[23][13]}\right)\right) + \left(\frac{-s_2+s_2s_1+s_1s_2-s_1s_2s_1}{[12][13][23]}\right)\\
&= \left(\frac{1-s_1}{[12][13][23]}\right) + \left(\frac{-s_2+s_2s_1+s_1s_2-s_1s_2s_1}{[12][13][23]}\right)\\
&= \frac{1-s_1-s_2+s_2s_1-s_1s_2s_1}{[12][13][23]}
\end{aligned}
$$

□

References

[StanleyEC2] Richard P. Stanley, *Enumerative Combinatorics. Volume 2*, Cambridge University Press 2023.

[GrinbergAC] Darij Grinberg, *An Introduction to Algebraic Combinatorics*, http://www.cip.ifi.lmu.de/~grinberg/t/21s/lecs.pdf

[KnutsonSP] Allen Knutson, *Schubert Polynomials and Symmetric Functions*, https://pi.math.cornell.edu/~allenk/schubnotes.pdf

[MacdonaldSP] Ian Macdonald, *Notes on Schubert Polynomials*