The $f$-vector of the descent polytope

Denis Chebikin and Richard Ehrenborg

December 6, 2008

Abstract

For a positive integer $n$ and a subset $S \subseteq [n-1]$, the descent polytope $DP_S$ is the set of points $(x_1, \ldots, x_n)$ in the $n$-dimensional unit cube $[0,1]^n$ such that $x_i \geq x_{i+1}$ if $i \in S$ and $x_i \leq x_{i+1}$ otherwise. We discuss several ways to compute the $f$-vector $(f_0, f_1, \ldots, f_n)$ of this polytope. First, we express the $f$-polynomial $F_S(t) = f_0 + f_1 t + \cdots + f_n t^n$ as a sum over all subsets of $[n-1]$. Second, we use certain factorizations of the associated word over a two-letter alphabet to describe the $f$-polynomial $F_S(t)$. Finally, we give efficient recursions to compute $F_S(t)$. We show that the $f$-vector is maximized when the set $S$ is the alternating set $\{1, 3, 5, \ldots\} \cap [n-1]$. We also derive a generating function for $F_S(t)$, written as a formal power series in two non-commuting variables with coefficients in $\mathbb{Z}[t]$.

1 Introduction

For a set $S \subseteq [n-1] = \{1, 2, \ldots, n-1\}$, define the descent polytope $DP_S$ to be the set of points $(x_1, \ldots, x_n)$ in $\mathbb{R}^n$ such that $0 \leq x_i \leq 1$, and

\[
\begin{cases}
  x_i \geq x_{i+1} & \text{if } i \in S, \\
  x_i \leq x_{i+1} & \text{if } i \notin S.
\end{cases}
\]

Thus $DP_S$ is the order polytope of the ribbon poset $Z_S = \{z_1, z_2, \ldots, z_n\}$ defined by the cover relations $z_i > z_{i+1}$ if $i \in S$ and $z_i < z_{i+1}$ if $i \notin S$; see [9]. Descent polytopes occur as a subdivision of the $n$-dimensional unit cube in the recent work [2] of Ehrenborg, Kitaev, and Perry. It is clear that the set $S$ and its complement $S' = [n-1] - S$ yields the same descent polytope up to an affine transformation.

In this paper our primary goal is to compute the $f$-vector of the descent polytope $DP_S$. Recall that for an $n$-dimensional polytope, the $f$-vector is the integer vector $(f_0, f_1, \ldots, f_n)$, where $f_i$ is the number of $i$-dimensional faces in the polytope. Observe that $f_n = 1$ since the descent polytope is a face of itself. For $S \subseteq [n-1]$, define the $f$-polynomial of the descent polytope $DP_S$ to be

\[
F_S(t) := \sum_{i=0}^{n} f_i \cdot t^i.
\]

To simplify notation, we will often write $F_S$ instead of $F_S(t)$. As we show in Section 2, $F_S$ can be expressed as a sum of polynomials taken over all subsets of $S$.

The volume of the descent polytope $DP_S$ is given by $\beta(S)/n!$, where $\beta(S)$ denotes the number of permutations in the symmetric group $S_n$ with descent set $S$. A classical result in combinatorics is
that $\beta(S)$ is maximized when $S$ is an alternating set, that is, $S = \{1, 3, 5, \ldots\} \cap [n - 1]$, or equivalently $S = \{2, 4, 6, \ldots\} \cap [n - 1]$; see [1, 6, 7, 8, 11]. We show that the same result holds for the $f$-vector of the descent polytope. That is, the number of $i$-dimensional faces of the descent polytope $\text{DP}_S$ is maximized when $S$ is the alternating set.

A different way to encode subsets of the set $[n - 1]$ is by a word in letters $x$ and $y$ of length $n - 1$. Viewing $x$ and $y$ as non-commutative variables and the words as monomials, we consider the non-commutative generating function $\Phi(x, y) = \sum_{\nu} F_{\nu} \cdot \nu$. We determine this generating function, which turns out to be a rational function. Using this rational function we obtain a more concise expression for the $f$-polynomial of the descent polytope $\text{DP}_S$. A second way to encode subsets of $[n - 1]$ is by compositions. In Section 4 we use this encoding to obtain more recurrences to compute the $f$-polynomial $F_S$.

We end the paper with a few open questions and directions for further research.

2 An expression for the $f$-polynomial $F_{\nu}$

Let $x$ and $y$ be two non-commuting variables. For $S \subseteq [m]$, define $\nu_S = \nu_1 \nu_2 \cdots \nu_m$ where

$$
\nu_i = \begin{cases} 
x & \text{if } i \notin S, \\
y & \text{if } i \in S.
\end{cases}
$$

We denote $F_S$ by $F_{\nu_S}$; such notation has an advantage, as $\nu_S$ encodes not only $S \subseteq [n - 1]$ but also the dimension $n$. Since pairs $(n, S \subseteq [n - 1])$ are in bijective correspondence with $xy$-words via $S \mapsto \nu_S$, it is natural to parameterize the $f$-polynomials of descent polytopes by $xy$-words and write $F_{\nu}$, where $\nu = \nu_S$ for some $S \subseteq [n]$, and $[\nu]$ denotes the length of the word $\nu$.

For a $xy$-word $\nu = \nu_1 \nu_2 \cdots \nu_{n-1}$, define the statistic $\kappa(\nu)$ by $\kappa(\nu) = 2 + |\{i : \nu_i \neq \nu_{i+1}\}|$ for $\nu \neq 1$, and $\kappa(1) = 1$. A direct observation is the number of facets of the descent polytope is described by $\kappa$.

Lemma 2.1. The number of $(n - 1)$-dimensional faces of the $n$-dimensional descent polytope $\text{DP}_\nu$ is given by

$$f_{n-1}(\text{DP}_\nu) = n - 1 + \kappa(\nu).$$

Proof. There are $n - 1$ supporting hyperplanes of the form $x_i = x_{i+1}$ that each intersect the polytope in a facet. The hyperplane $x_i = 1$ intersects the polytope in a facet if one of the following three cases holds: $\nu_{i-1} \nu_i = xy$; $i = 1$ and $\nu_1 = y$; or $i = n$ and $\nu_n = x$. A similar statement holds for the hyperplane $x_i = 0$. The lemma follows by adding these three statements.

For a $xy$-word $\nu = \nu_1 \nu_2 \cdots \nu_{n-1}$ and a subset $T$ of $[n - 1]$, define $\nu^T$ to be the subword $\nu^T = \nu_{j_1} \nu_{j_2} \cdots \nu_{j_k}$, where $T = \{j_1 < j_2 < \cdots < j_k\}$. The following theorem provides a way to compute the $f$-polynomial $F_{\nu}$.

Theorem 2.2. Let $\nu$ be a $xy$-word of length $n - 1$. Then the $f$-polynomial of the descent polytope $\text{DP}_\nu$ is given by

$$F_{\nu} = 1 + \sum_{T \subseteq [n-1]} \left(\frac{t+1}{t}\right)^{\kappa(\nu^T)} \cdot t^{|T|+1}.$$
Proof. For a face $\mathcal{F}$ of a polytope, let $\mathcal{F}^I$ denote the relative interior of $\mathcal{F}$. Then the polytope is the disjoint union of $\mathcal{F}^I$ taken over all faces $\mathcal{F}$, including the polytope itself.

Recall that the descent polytope $\text{DP}_v$ consists of all points $(x_1, \ldots, x_n) \in \mathbb{R}^n$ belonging simultaneously to the half spaces $x_i \geq 0, x_i \leq 1$ ($1 \leq i \leq n$), $x_i \leq x_{i+1}$ ($v_i = x$), and $x_i \geq x_{i+1}$ ($v_i = y$). A face $\mathcal{F}$ of $\text{DP}_v$ can be uniquely identified by specifying which of these half spaces contain $\mathcal{F}$ on their boundary hyperplanes, as long as the intersection of the whole polytope and the specified boundary hyperplanes is non-empty. Forming the specification just for the half spaces of the form $x_i \leq x_{i+1}$ or $x_i \geq x_{i+1}$ restricts the location of $\mathcal{F}^I$ in $\mathbb{R}^n$ to the region defined by the relations

$$x_1 = x_2 = \cdots = x_{j_1} \leq x_{j_1+1} = x_{j_1+2} = \cdots = x_{j_2} \leq \cdots \leq x_{j_k+1} = x_{j_k+2} = \cdots = x_n$$

for some $T = \{j_1 < j_2 < \cdots < j_k\} \subseteq [n-1]$, where the symbol $\leq$ denotes strict inequality: $x_{j_i} \neq x_{j_{i+1}}$ if $v_{j_i} = x$, or $x_{j_i} > x_{j_{i+1}}$ if $v_{j_i} = y$. Then $T$ is the set of indexes $j$ for which $\mathcal{F}$ does not lie entirely on the boundary hyperplane $x_j = x_{j+1}$ and thus the relative interior $\mathcal{F}^I$ is contained in the interior of the corresponding half space. Let $\mathcal{R}(T)$ denote the intersection of the region defined by (2.1) and the hypercube $[0,1]^n$. Each point $(x_1, \ldots, x_n)$ of $\text{DP}_v$ belongs to exactly one such region $\mathcal{R}(T)$, namely, the one for $T = \{j \mid x_j \neq x_{j+1}\}$. Thus we have the disjoint union

$$\text{DP}_v = \bigsqcup_{T \subseteq [n-1]} \mathcal{R}(T).$$

Let us show that the term corresponding to $T \neq \emptyset$ in the expression in the statement of the theorem is the contribution to $F_v$ of the faces $\mathcal{F}$ of $\text{DP}_v$ for which $\mathcal{F}^I$ is contained in the region $\mathcal{R}(T)$. In other words, we claim that for $T \neq \emptyset$ we have

$$\sum_{\mathcal{F} : \mathcal{F}^I \subseteq \mathcal{R}(T)} t^{|\mathcal{F}^I|} = \left( \frac{t+1}{t} \right)^{\kappa(v^T)} \cdot t^{|T|+1}.$$  

(2.2)

Fix $\emptyset \neq T \subseteq [n-1]$. To select a particular face $\mathcal{F}$ from the set of all faces with the property $\mathcal{F}^I \subseteq \mathcal{R}(T)$, we need to complete the specification started above, that is, we must specify which of the hyperplanes $x_i = 0, 1$ contain $\mathcal{F}$, and we must make sure that the intersection of the set of the specified hyperplanes and $\mathcal{R}(T)$ is non-empty. In terms of defining relations (2.1), this task is equivalent to setting the common value of some of the “blocks” of coordinates $(x_1, \ldots, x_{j_1})$, $(x_{j_1+1}, \ldots, x_{j_2})$, $\ldots$, $(x_{j_k+1}, \ldots, x_n)$ to 0 or 1. Since the relations must remain satisfiable by at least one point in $[0,1]^n$, only the blocks preceded in (2.1) by $>$ (or nothing) and succeeded by $<$ (or nothing) can be set to 0. Similarly, only the blocks preceded by $<$ (or nothing) and succeeded by $>$ (or nothing) can be set to 1. Thus each block can be set to at most one of 0 and 1. The letters of the $xy$-word $v^T = v_{j_1} \cdots v_{j_k}$ encode the inequality signs in (2.1) (x stands for $<$, and y stands for $>$), so the number of blocks that can be set to 0 or 1 is the total number of occurrences of x followed by y, or y followed by x, in $v^T$, plus 2, as we also need to count the first and the last blocks. In other words, the number of such blocks is $\kappa(v^T)$.

Observe that the dimension of the face of $\text{DP}_v$ obtained by this specification procedure equals the number of blocks that have not been set to 0 or 1: the common values of the coordinates in those blocks form the “degrees of freedom” that constitute the dimension. Let us call such blocks free. The number of faces $\mathcal{F}$ with $\mathcal{F}^I \subseteq \mathcal{R}(T)$ for which the specification procedure results in $m$ free blocks is

$$\left( \frac{\kappa(v^T)}{|T| + 1 - m} \right).$$

3
the number of ways to choose \(|T| + 1 - m\) blocks that are not free out of \(\kappa(v^T)\) possibilities. Hence we have
\[
\sum_{\mathcal{F} \subseteq \mathcal{R}(T)} t^{\dim \mathcal{F}} = \sum_{m=|T|+1-\kappa(v^T)}^{[T]+1} \left( \frac{\kappa(v^T)}{|T| + 1 - m} \right) \cdot t^m \\
= t^{[T]+1-\kappa(v^T)} \sum_{\ell=0}^{\kappa(v^T)} \binom{\kappa(v^T)}{\ell} \cdot t^\ell \\
= t^{[T]+1-\kappa(v^T)} \cdot (t + 1)^{\kappa(v^T)},
\]
proving \((2.2)\).

Finally, for \(T = \emptyset\), we have \(\mathcal{R}(T) = \{0 \leq x_1 = \cdots = x_n \leq 1\}\), which is just the line segment joining the two vertices \((0, \ldots, 0)\) and \((1, \ldots, 1)\) of \(DP_v\). Thus the contribution of \(\mathcal{R}(T)\) to \(F_v\) is
\[
t + 2 = 1 + \left( \frac{t + 1}{t} \right)^{\kappa(v^\emptyset)} \cdot t.
\]
Adding this equation to the sum of \((2.2)\) taken over the non-empty \(T\) proves the theorem. \(\square\)

Theorem \(2.2\) yields a combinatorial interpretation of the number of vertices of the polytope \(DP_v\). Call a \(xy\)-word \(v = v_1v_2\cdots v_k\ alternating if \(v_i \neq v_{i+1}\) for all \(1 \leq i \leq k - 1\). Then we have the following corollary.

**Corollary 2.3.** For \(v\) a \(xy\)-word of length \(n - 1\), the number of vertices of the descent polytope \(DP_v\) is one greater than the number of subsets \(T \subseteq [n - 1]\) for which the word \(v^T\) is alternating.

**Proof.** The number of vertices of \(DP_v\) is the constant term of \(F_v\). For the summand corresponding to a subset \(T \subseteq [n - 1]\) in the formula of Theorem \(2.2\), the constant term is either 0 or 1, the latter being the case if and only if \(|T| + 1 - \kappa(v^T) = 0\). This condition is equivalent to \(v^T\) being alternating, proving the corollary. \(\square\)

As we mention in the introduction, the descent set statistic \(\beta(S)\) is maximized when \(S\) is the alternating set. The most elegant proof of this fact uses the cd-index of the simplex; see [7]. For a \(xy\)-word \(v\), let \(\nabla\) denote the word obtained from \(v\) by replacing \(x\)'s with \(y\)'s and vice versa. Then the following inequality holds:
\[
\beta(uyxv) > \beta(uyy\nabla), \tag{2.3}
\]
where we use \(xy\)-words to encode the sets. In each of the proofs [11] [10] [5] [11] that the alternating word maximizes the descent set, the arguments rely on proving the inequality \((2.3)\). However, the cd-index proof gives a quick way to verify this inequality. We now state a similar inequality for the \(f\)-vectors of descent polytopes.

**Theorem 2.4.** Let \(u\) and \(v\) be two \(xy\)-words such that \(|u| + |v| = n - 3\). Then the difference
\[
F_{uyxv}(t) - F_{uyy\nabla}(t) \tag{2.4}
\]
has positive coefficients at \(1, t, \ldots, t^{n-1}\). That is, for \(0 \leq i \leq n - 1\) the descent polytope \(DP_{uyxv}\) has more faces of dimension \(i\) than the descent polytope \(DP_{uyy\nabla}\).
Table 1: Calculations for the proof of Theorem 2.4.
Proof. Let \(|u| = m\) and \(|v| = n - m - 3\). For \(T \subseteq [m]\) and \(U \subseteq [n - m - 3]\), define
\[ Q_{T,U}(t) := \sum_{E \subseteq \{1,2\}} \frac{(t+1)}{t} \kappa((u^T(yv))^k v^U) \cdot t^{|T|+|U|+|E|+1}. \] (2.5)
Thus \(Q_{T,U}(t)\) is the sum of four of the terms in the summation formula for \(F_{uvxy}(t)\) given by Theorem 2.2, corresponding to fixed choices of letters drawn from \(u\) and from \(v\). Similarly, let us define
\[ \overline{Q}_{T,U}(t) := \sum_{E \subseteq \{1,2\}} \frac{(t+1)}{t} \kappa((u^T(yv))^k v^U) \cdot t^{|T|+|U|+|E|+1}. \]
Note that \(Q_{T,U}(t)\) depends only on \(|T|, |U|, \kappa(u^T v^U)\), the last letter of \(u^T\), and the first letter of \(v^U\), and not on the remaining letters of \(u^T\) and \(v^U\). Thus to show that the difference \(Q_{T,U}(t) - \overline{Q}_{T,U}(t)\) is a polynomial with non-negative coefficients, it suffices to consider 9 cases corresponding to \(u^T\) (respectively, \(v^U\)) ending (respectively, beginning) with \(x\) or \(y\), or being equal to the empty word 1.

We summarize our calculations in Table 1. We denote \(\kappa(u^T v^U)\) by \(k\), and we divide each polynomial by the common factor \((t + 1)^k \cdot t^{(|T|+|U|+1-k)}\). In the fourth column, which corresponds to \(Q_{T,U}\), the four summands represent the results of inserting 1, \(x\), \(y\), and \(yx\) between \(u^T\) and \(v^U\). For example, if \(u^T\) ends with a \(x\) and \(v^U\) begins with a \(x\), then inserting \(y\) increases the value of the statistic \(k\) by 2, thus contributing a factor of \((t+1)^2 \cdot t = (t+1)^2 \cdot t^{-1}\) to the corresponding term of (2.5). Similarly, the entries in the fifth column consist of a factor resulting from a different value of \(k\) for the word \(u^T v^U\) times the contributions of inserting 1, \(y\) (counted twice), and \(y^2\) between \(u^T\) and \(v^U\).

We conclude that in every case the quotient \(\frac{Q_{T,U}(t) - \overline{Q}_{T,U}(t)}{(t+1)^k \cdot t^{(|T|+|U|+1-k)}}\) is a polynomial of degree 2 with non-negative coefficients. Hence the difference \(Q_{T,U}(t) - \overline{Q}_{T,U}(t)\) is a polynomial with non-negative coefficients. Summing over all possible pairs \((T,U)\) yields that the difference in equation (2.4) has degree \(k(1) + (|T| + |U| + 1 - k) + 2 = |T| + |U| + 2\). This degree can attain any integer value between 2 and \(n - 1\). Thus the leading terms of these differences contribute positively to the coefficients of \(t^2, t^3, \ldots, t^{n-1}\) in the difference (2.4). Furthermore, in the case \(T = U = \emptyset\) we have \(Q_{T,U}(t) - \overline{Q}_{T,U}(t) = (t+1)^2\), which yields a positive contribution to the constant and the linear terms of the overall difference. The proof is now complete.

Let \(z_n\) be the alternating word of length \(n\) starting with the letter \(x\). Then \(\overline{z_n}\) is the alternating word beginning with \(y\). That is, the two alternating words are
\[ z_n = xyy \cdot \overline{y} \quad \text{and} \quad \overline{z_n} = yyy \cdot \overline{y}. \]
We now have the maximization result for the \(f\)-vector of descent polytopes.

**Corollary 2.5.** The \(f\)-vector of the two descent polytopes \(DP_{z_n-1}\) and \(DP_{\overline{z}_n-1}\) is maximal among the \(f\)-vectors of all descent polytopes of dimension \(n\). That is, for each \(0 \leq i \leq n-1\), the polytope \(DP_{z_n-1}\) has more faces of dimension \(i\) than the descent polytope \(DP_\emptyset\) of dimension \(n\) for an non-alternating word \(v\).
3 The power series $\Phi(x, y)$

We now derive a non-commutative generating function $\Phi(x, y)$ for the $f$-polynomial $F_v$, which belongs to the ring $\Phi(x, y) \in \mathbb{Z}[t][\langle x, y \rangle]$. We define the power series $\Phi(x, y)$ by

$$\Phi(x, y) = \sum_v F_v \cdot v,$$

where the sum is over all $xy$-words $v$. Since we have the symmetry $F_v = F_{\bar{v}}$, we obtain that $\Phi(x, y)$ is symmetric with respect to $x$ and $y$, that is,

$$\Phi(x, y) = \Phi(y, x).$$

Let $v$ be a $xy$-word $v_1v_2 \cdots v_{n-1}$. Consider the following polynomials:

$$K_v(t) := \sum_{T \subseteq [n-1] : v_{j_1} = x} \left(\frac{t+1}{t}\right)^{\kappa(T)} \cdot t^{|T|+1},$$

$$L_v(t) := \sum_{T \subseteq [n-1] : v_{j_1} = y} \left(\frac{t+1}{t}\right)^{\kappa(T)} \cdot t^{|T|+1},$$

where $v_{j_1}$ denotes the first letter of the word $v^T = v_{j_1}v_{j_2} \cdots v_{j_k}$, as in the notation of Theorem 2.2. Since $v^T$ begins with either $x$ or $y$ unless $T = \emptyset$, we have

$$F_v = K_v + L_v + t + 2. \quad (3.1)$$

We continue with a lemma that relates the two polynomials $K_v$ and $L_v$.

**Lemma 3.1.** For a $xy$-word $v$ the following four equalities hold:

$$K_{xy} = K_v,$$

$$L_{xy} = L_v,$$

$$K_{xy} = L_{xy} = (t+1) \cdot (K_v + L_v + t + 1).$$

**Proof.** For an integer $i$ and a set $U \subseteq \mathbb{Z}$, let $U + i$ denote the set obtained by adding $i$ to each element of $U$. Also let $v = v_1v_2 \cdots v_{n-1}$, where each $v_i$ is either $x$ or $y$.

Clearly, $(yv)^T$ begins with $x$ if and only if $1 \notin T$ and $v^{T-1}$ begins with $x$, in which case $(yv)^T = v^{T-1}$. Hence $K_{xy} = K_v$.

Now, $(xv)^T$ begins with $x$ if and only if either $1 \in T$, or else $1 \notin T$ and $v^{T-1}$ begins with $x$. In the former case, we have $T = \{1 < j_1 + 1 < j_2 + 1 < \cdots < j_k + 1\}$, and $(xv)^T = xv_{j_1}v_{j_2} \cdots v_{j_k}$. Set $U = (T - \{1\}) - 1 = \{j_1 < \cdots < j_k\}$. Then $\kappa((xv)^T) = \kappa(v^U)$ if $v_{j_1} = x$, and $\kappa((xv)^T) = \kappa(v^U) + 1$ if $v_{j_1} = y$. Hence

$$\sum_{1 \in \mathbb{T} \subseteq [n]} \left(\frac{t+1}{t}\right)^{\kappa((xv)^T)} \cdot t^{|T|+1} = (t+1)^2 + t \cdot \sum_{U : v_{j_1} = x} \left(\frac{t+1}{t}\right)^{\kappa(v^U)} \cdot t^{|U|+1}$$

$$+ (t+1) \cdot \sum_{U : v_{j_1} = y} \left(\frac{t+1}{t}\right)^{\kappa(v^U)} \cdot t^{|U|+1} \quad (3.2)$$

$$= (t+1)^2 + t \cdot K_v + (t+1) \cdot L_v.$$
where the leading term \((t + 1)^2\) corresponds to \(T = \{1\}\) and \(U = \emptyset\). In the case where \(1 \notin T\) and \(v^{T-1}\) begins with \(x\) we have, as before, \((xv)^T = v_{j_1} v_{j_2} \cdots v_{j_h} = v^{T-1}\), and hence

\[
\sum_{T : v_{j_1} = x} \left( \frac{t + 1}{t} \right)^{\kappa((xv)^T)} \cdot t_{|T|+1} = \sum_{v_{j_1} = x} \left( \frac{t + 1}{t} \right)^{\kappa(v^{T-1})} \cdot t_{|T-1|+1} = K_v. \tag{3.3}
\]

Adding (3.2) and (3.3) yields

\[
K_{xv} = (t + 1) \cdot (K_v + L_v + t + 1).
\]

The relations for \(L_{xv}\) and \(L_{yv}\) follow from symmetry that arises from exchanging the variables \(x\) and \(y\). □

Starting with \(K_1 = L_1 = 0\), one can use Lemma 3.1 to recursively compute \(K_v\) and \(L_v\), and hence \(F_v\), from (3.1). Recall the generating power series

\[
\Phi(x, y) = \sum_v F_v \cdot v,
\]

where the sum is over all \(xy\)-words, including the empty word \(v = v_\emptyset = 1\). Define the two generating power series

\[
K(x, y) := \sum_v K_v \cdot v, \\
\Lambda(x, y) := \sum_v L_v \cdot v.
\]

From the definitions of \(K_v\) and \(L_v\) it follows that \(K_v = L_v\). Hence we have that

\[
K(x, y) = \Lambda(x, y).
\]

Then, by (3.1), we have

\[
\Phi(x, y) = K(x, y) + \Lambda(x, y) + (t + 2) \cdot \sum_v v
\]

\[
= K(x, y) + K(y, x) + (t + 2) \cdot \sum_{r \geq 0} (x + y)^r
\]

\[
= K(x, y) + K(y, x) + (t + 2) \cdot \frac{1}{1 - x - y}.
\]

Using the equations in Lemma 3.1 and recalling that \(K_1 = 0\) we obtain

\[
K(x, y) = \sum_v K_{xv} \cdot xv + \sum_v K_{yv} \cdot yv
\]

\[
= (t + 1) \cdot x \cdot \sum_v (K_v + L_v + t + 1) \cdot v + y \cdot \sum_v K_v \cdot v
\]

\[
= (t + 1) \cdot x \left( K(x, y) + \Lambda(x, y) + (t + 1) \cdot \frac{1}{1 - x - y} \right) + y \cdot K(x, y)
\]

\[
= (t + 1) \cdot x \left( \Phi(x, y) - \frac{1}{1 - x - y} \right) + y \cdot K(x, y),
\]

where the sum is over all \(xy\)-words, including the empty word \(v = v_\emptyset = 1\). Define the two generating power series

\[
K(x, y) := \sum_v K_v \cdot v, \\
\Lambda(x, y) := \sum_v L_v \cdot v.
\]

From the definitions of \(K_v\) and \(L_v\) it follows that \(K_v = L_v\). Hence we have that

\[
K(x, y) = \Lambda(x, y).
\]

Then, by (3.1), we have

\[
\Phi(x, y) = K(x, y) + \Lambda(x, y) + (t + 2) \cdot \sum_v v
\]

\[
= K(x, y) + K(y, x) + (t + 2) \cdot \sum_{r \geq 0} (x + y)^r
\]

\[
= K(x, y) + K(y, x) + (t + 2) \cdot \frac{1}{1 - x - y}.
\]

Using the equations in Lemma 3.1 and recalling that \(K_1 = 0\) we obtain

\[
K(x, y) = \sum_v K_{xv} \cdot xv + \sum_v K_{yv} \cdot yv
\]

\[
= (t + 1) \cdot x \cdot \sum_v (K_v + L_v + t + 1) \cdot v + y \cdot \sum_v K_v \cdot v
\]

\[
= (t + 1) \cdot x \left( K(x, y) + \Lambda(x, y) + (t + 1) \cdot \frac{1}{1 - x - y} \right) + y \cdot K(x, y)
\]

\[
= (t + 1) \cdot x \left( \Phi(x, y) - \frac{1}{1 - x - y} \right) + y \cdot K(x, y),
\]

8
where the last step is by equation (3.4). Rearranging terms we have

\[ K(x, y) = (t + 1) \cdot (1 - y)^{-1} \cdot x \left( \Phi(x, y) - \frac{1}{1 - x - y} \right). \]

Adding this equation and its symmetric version obtained by exchanging \( x \) and \( y \) one has

\[ K(x, y) + K(y, x) = (t + 1) \cdot ((1 - y)^{-1} \cdot x + (1 - x)^{-1} \cdot y) \left( \Phi(x, y) - \frac{1}{1 - x - y} \right), \]

using the symmetry \( \Phi(x, y) = \Phi(y, x) \). Now using equation (3.4) we can solve for \( \Phi(x, y) \) and arrive at the following theorem.

**Theorem 3.2.** The generating power series \( \Phi(x, y) \) is given by

\[ \Phi(x, y) = \left( 1 + \frac{t + 1}{1 - (t + 1) \cdot ((1 - y)^{-1} \cdot x + (1 - x)^{-1} \cdot y)} \right) \cdot \frac{1}{1 - x - y}. \]

**Corollary 3.3.** For a \( xy \)-word \( v \) the \( f \)-vector of the descent polytope \( DP_v \) is given by the sum

\[ F_v(t) = 1 + \sum_{(u_1, \ldots, u_{k-1}, u_k)} (t + 1)^k, \]

where the sum ranges over all factorizations of the word \( v = u_1 \cdots u_{k-1} \cdot u_k \) such that each of factors \( u_1, \ldots, u_{k-1} \) are of the form \( x^i y \) or \( y^i x \), where \( i \geq 0 \), and there is no condition on the last factor \( u_k \).

**Proof.** Rewrite Theorem 3.2 as

\[
\Phi(x, y) = \frac{1}{1 - x - y} + \frac{1}{1 - (t + 1) \cdot ((1 - y)^{-1} \cdot x + (1 - x)^{-1} \cdot y)} \cdot \frac{t + 1}{1 - x - y} \\
= \sum_{v} v + \sum_{j \geq 0} \left( (t + 1) \cdot \sum_{i \geq 0} (y^i x + x^i y) \right)^j \cdot (t + 1) \cdot \sum_{v} v,
\]

where in both sums \( v \) ranges over all \( xy \)-words. The corollary follows by reading the generating function. \( \square \)

**Example 3.4.** Consider the 5-dimensional descent polytope \( DP_{xyyx} \) where \( v = xyyx \). We have the following list of 11 factorizations:

\begin{align*}
\text{v} &= \text{xyyx} & = \text{x \cdot yyx} & = \text{x \cdot y \cdot yx} & = \text{xy \cdot yx} \\
&= \text{x \cdot y \cdot y \cdot x} & = \text{xy \cdot y \cdot x} & = \text{x \cdot y \cdot y \cdot x \cdot 1} & = \text{xy \cdot y \cdot x \cdot 1} \\
&= \text{x \cdot y \cdot yx \cdot 1} & = \text{xy \cdot yx \cdot 1} & = \text{x \cdot yyx \cdot 1}
\end{align*}

Hence the \( f \)-polynomial of the polytope \( DP_{xyyx} \) is given by

\[
F_{xyyx} = 1 + (t + 1) + 2 \cdot (t + 1)^2 + 4 \cdot (t + 1)^3 + 3 \cdot (t + 1)^4 + (t + 1)^5 \\
= 12 + 34 \cdot t + 42 \cdot t^2 + 26 \cdot t^3 + 8 \cdot t^4 + t^5.
\]
For the alternating word $z_{n-1}$ we can say more about the associated descent polytope. The number of vertexes of $DP_{z_{n-1}}$ has Fibonacci number of vertexes; see for instance [10, Chapter 1, Exercise 14e]. More generally, the $f$-vector of $DP_{z_{n-1}}$ is given by the next result.

**Corollary 3.5.** The $f$-polynomial of the $n$-dimensional descent polytope $DP_{z_{n-1}}$ is described by

$$F_{z_{n-1}} = 1 + \sum_{(c_1, c_2, \ldots, c_k)} (t + 1)^k,$$

where the sum is over all compositions of $n$ such that all but the last part is less than or equal to 2, that is, $c_1, \ldots, c_{k-1} \in \{1, 2\}$.

**Proof.** The only factors of the alternating word $z_{n-1}$ of the form $x^i y$ or $y^i x$ has $i = 0, 1$. Hence it is enough to record the length of each factor $u_i$, that is, $d_i = |u_i|$. Thus we are summing over vectors of non-negative integers $(d_1, \ldots, d_k)$ such that the sum of the entries is $n - 1$ and $d_1, \ldots, d_{k-1} \in \{1, 2\}$ and $d_k \geq 0$. By adding one to the last entry $d_k$ we have a composition of $n$. \hfill \square

This corollary yields the generating function

$$\sum_{n \geq 1} F_{z_{n-1}} x^n = \frac{x}{1 - x} + \frac{1}{1 - (t + 1) \cdot (x + x^2)} \cdot (t + 1) \cdot \frac{x}{1 - x}. \quad (3.5)$$

Setting $t = 0$ in this generating function and adding constant 1 yields $(1 + x)/(1 - x - x^2)$, the generating function for the Fibonacci numbers as expected.

## 4 More recurrence relations

In this section we derive a different set of recurrences determining $F_S(t)$ than the ones we used to obtain Theorem 3.2. Here it will be more convenient to associate integer sets with compositions. Let $Comp'(m)$ denote the set of integer compositions $(\gamma_1, \gamma_2, \ldots)$ of $m$ with $\gamma_1 \geq 0$ and $\gamma_2, \gamma_3, \ldots > 0$. For $\gamma = (\gamma_1, \gamma_2, \ldots) \in Comp'(m)$, define $v_\gamma = x^{\gamma_1} y^{\gamma_2} x^{\gamma_3} y^{\gamma_4} \cdots$. This a bijection between compositions in $Comp'(m)$ and $xy$-words of length $m$. By composing this bijection with the correspondence between words and subsets we have a bijection between subsets of $[m]$ and $Comp'(m)$. For example, if we have $S = \{1, 3, 4\} \subseteq [6]$, then $v_S = yxyxyx = x^0 y^1 x^1 y^2 x^2$ and thus $c(S) = (0, 1, 1, 2, 2)$.

Write $F_S(t) = F_c(S)(t)$. In the notation of Theorem 3.2, define

$$G_\gamma(t) := t + 1 + \sum_{T : i_1 > \gamma_1} \left( \frac{t + 1}{t} \right)^{\kappa(v_\gamma^T)} \cdot t^{|T|+1},$$

$$H_\gamma(t) := \sum_{T : i_1 \leq \gamma_1} \left( \frac{t + 1}{t} \right)^{\kappa(v_\gamma^T)} \cdot t^{|T|+1},$$

where $T = \{i_1 < i_2 < \cdots \} \subseteq [n-1]$. Thus $F_\gamma = 1 + G_\gamma + H_\gamma$. The extra $t + 1$ in $G_\gamma$ corresponds to the term for $T = \emptyset$. Note that $G_\gamma$ does not depend on the value of $\gamma_1$, and that $\gamma_1 = 0$ implies $H_\gamma = 0$. For $\gamma = (\gamma_1, \gamma_2, \ldots)$, write $\gamma^{(r)} = (\gamma_{\ell+1}, \gamma_{\ell+2}, \ldots)$. Also, for a nonnegative integer $r$, define

$$p_r = p_r(t) := \frac{(t + 1)^r - 1}{t} = [r]_q = q+1,$$
where \([r]_q = \frac{q^r - 1}{q - 1}\) is the classical \(q\)-analog of \(r\). Breaking up the summation formula for \(F_{\gamma}\) allows to obtain the following recurrence relations.

**Lemma 4.1.** For \(r \geq 0\), we have \(G_{(r)} = t + 1\) and \(H_{(r)} = pr \cdot (t + 1)^2\). If \(\gamma = (\gamma_1, \gamma_2, \ldots)\) has at least two parts, then

\[
G_{\gamma} = G_{\gamma(1)} + H_{\gamma(1)},
\]

\[
H_{\gamma} = p_{\gamma_1} \cdot \left( (t + 1)^2 + t \cdot (H_{\gamma(1)} + H_{\gamma(2)} + \cdots) + H_{\gamma(1)} + H_{\gamma(3)} + H_{\gamma(5)} + \cdots \right).
\]

**Proof.** Comparing with Theorem 2.2 observe that \(G_{\gamma} = F_{\gamma(1)} - 1 = G_{\gamma(1)} + H_{\gamma(1)}\).

Now consider the terms in the definition of \(H_{\gamma}\). These terms correspond to \(T \subseteq [n - 1]\) such that \(T \cap [\gamma_1] \neq \emptyset\). Write \(T = \{i_1 < \cdots < i_\ell < j_1 < \cdots < j_{k-\ell}\}\), where \(i_\ell \leq \gamma_1\) and \(j_1 > \gamma_1\). Let \(I = \{i_1 < \cdots < i_\ell\}\) and \(J = \{j_1 < \cdots < j_{k-\ell}\}\), so that \(v_J^T = v_J^{I^J}\). Since \(v_J^{I^J}\) is a positive power of \(x\), we have \(\kappa(v_J^{I^J}) = \kappa(v_J^{I^J})\) if \(v_J^{I^J}\) begins with a \(x\), and \(\kappa(v_J^{I^J}) = \kappa(v_J^{I^J}) + 1\) if \(v_J^{I^J}\) begins with a \(y\) or if \(J = \emptyset\). Hence we have

\[
H_{\gamma} = \left( \sum_{\emptyset \neq I \subseteq [\gamma_1]} t^{\lvert I\rvert} \right) \cdot \left( \sum_{J \subseteq [n-1] \setminus [\gamma_1]} \left( \frac{t + 1}{t} \right)^{\kappa(v_J^{I^J})} \cdot t^{\lvert J\rvert + 1} \right),
\]

where \(c_J^{I^J}\) is 0 if \(v_J^{I^J}\) begins with a \(x\), and 1 otherwise. The sum on the left is \((t + 1)^{\gamma_1} - 1 = t \cdot p_{\gamma_1}\). Therefore

\[
H_{\gamma} = t \cdot p_{\gamma_1} \cdot \left( \sum_{J} \left( \frac{t + 1}{t} \right)^{\kappa(v_J^{I^J})} \cdot t^{\lvert J\rvert + 1} \right)
+ \frac{1}{t} \cdot \sum_{J : c_J^{I^J} = 1} \left( \frac{t + 1}{t} \right)^{\kappa(v_J^{I^J})} \cdot t^{\lvert J\rvert + 1}
= t \cdot p_{\gamma_1} \cdot \left( H_{\gamma(1)} + H_{\gamma(2)} + \cdots + t + 1 \right)
+ \frac{1}{t} \cdot \left( H_{\gamma(1)} + H_{\gamma(3)} + H_{\gamma(5)} + \cdots + t + 1 \right),
\]

because \(H_{\gamma(i)}\) is the sum of \(((t + 1)/t)^{\kappa(v_J^{I^J})} \cdot t^{\lvert J\rvert + 1}\) taken over \(J\) with \(\gamma_1 + \cdots + \gamma_i < j_1 \leq \gamma_1 + \cdots + \gamma_{i+1}\), and \(c_J^{I^J} = 1\) if and only if this condition on \(y_1\) holds for odd \(i\) (or if \(J = \emptyset\), which is accounted for by the two \(t + 1\) terms). \(\square\)

The next lemma provides a more concise recurrence relation for \(H_{\gamma}\).

**Lemma 4.2.** For a composition \(\gamma = (\gamma_1, \gamma_2, \ldots)\) with at least two parts, the following equality holds:

\[
H_{\gamma} + H_{(\gamma_1, \gamma_2, \gamma_3, \gamma_4, \ldots)} = p_{\gamma_1} \cdot \left( t \cdot G_{\gamma(1)} + (t + 1) \cdot (1 + G_{\gamma}) \right).
\]


**Proof.** Observe that, by Lemma 4.1,

\[ G_{\gamma} = G_{\gamma(1)} + H_{\gamma(1)} = G_{\gamma(2)} + H_{\gamma(2)} = \cdots = t + 1 + \left( H_{\gamma(1)} + H_{\gamma(2)} + \cdots \right) \]

since \(G(0) = t + 1\). Apply the relation (4.1) to \(H_\gamma\) and \(H_{(\gamma_1, \gamma_3, \gamma_4, \ldots)}\), and then add the two resulting equations:

\[
H_\gamma + H_{(\gamma_1, \gamma_3, \gamma_4, \ldots)} = t \cdot p_{\gamma_1} \cdot \left( G_\gamma + \frac{1}{t} \cdot \left( H_{\gamma(1)} + H_{\gamma(3)} + H_{\gamma(5)} + \cdots + t + 1 \right) \right)
\]

\[
= t \cdot p_{\gamma_1} \cdot \left( G_{\gamma(1)} + \frac{1}{t} \cdot \left( H_{\gamma(2)} + H_{\gamma(4)} + H_{\gamma(6)} + \cdots + t + 1 \right) \right)
\]

\[
= t \cdot p_{\gamma_1} \cdot \left( G_{\gamma} + G_{\gamma(1)} + \frac{1}{t} \cdot (G_\gamma + t + 1) \right).
\]

The lemma follows. 

A useful consequence of Lemma 4.1 is that in working with \(G_\gamma\) and \(H_\gamma\) we can concentrate on just the compositions with the first part equal to 1. Specifically, we have the following corollary.

**Corollary 4.3.** The polynomials \(G\) and \(H\) satisfy

\[
G_{(\gamma_1, \gamma_2, \ldots)} = G_{(1, \gamma_2, \gamma_3, \ldots)},
\]

\[
H_{(\gamma_1, \gamma_2, \ldots)} = p_{\gamma_1} \cdot H_{(1, \gamma_2, \gamma_3, \ldots)}.
\]

**Proof.** The first identity follows from an earlier observation that \(G_\gamma\) is independent of the first part of \(\gamma\), and the second one follows from Lemma 4.1 since \(p_1 = 1\).

Thus for a composition \(\gamma\) with \(k\) parts, we can compute \(F_\gamma\) by applying the recurrence relations of Lemmas 4.1 and 4.2 \(k\) times. For instance, to compute \(F_{(2,4,3)}\) we proceed as follows:

\[
G_{(1)} = G_{(3)} = t + 1,
\]

\[
H_{(1)} = (t + 1)^2,
\]

\[
G_{(1,3)} = G_{(4,3)} = G_{(3)} + H_{(3)} = G_{(3)} + p_3 \cdot H_{(1)},
\]

\[
H_{(1,3)} = -H_{(1)} + (t \cdot G_{(3)} + (t + 1) \cdot (1 + G_{(1,3)})),
\]

\[
G_{(1,4,3)} = G_{(2,4,3)} = G_{(4,3)} + H_{(4,3)} = G_{(1,3)} + p_4 \cdot H_{(1,3)},
\]

\[
H_{(1,4,3)} = -H_{(1,3)} + (t \cdot G_{(4,3)} + (t + 1) \cdot (1 + G_{(1,4,3)})),
\]

and finally

\[
F_{(2,4,3)} = 1 + G_{(2,4,3)} + H_{(2,4,3)} = 1 + G_{(1,4,3)} + p_2 \cdot H_{(1,4,3)}.
\]

The special case of the alternating word \(z_{n-1}\) corresponds to the composition \(\gamma = 1^{n-1} = (1, 1, \ldots, 1) \mid n - 1\). The generating function for \(F_{1^{n-1}}(t) = F_{z_{n-1}}(t)\) in equation (3.5) can be obtained from the recurrences of this section. From Lemmas 4.1 and 4.2 we get the relations

\[
G_{1^{n-1}} = G_{1^{n-2}} + H_{1^{n-2}},
\]

\[
H_{1^{n-1}} + H_{1^{n-2}} = t \cdot G_{1^{n-2}} + (t + 1) \cdot (1 + G_{1^{n-1}}).
\]
for \( n \geq 2 \), where we set \( G_1 = 0 \) and \( H_1 = t + 1 \) for convenience (it can be easily seen that the relations are valid for \( n = 2 \)). Then we multiply the above equations by \( x^n \) and sum over all \( n \geq 2 \) to obtain the system of equations

\[
\begin{align*}
G & = x \cdot (G + H), \\
H + x \cdot H & = t \cdot x \cdot G + (t + 1) \cdot (x^2 \cdot (1 - x)^{-1} + G) + (t + 1) \cdot x,
\end{align*}
\]

where \( G = G(t, x) := \sum_{n \geq 1} G_{1n-1}(t) x^n \) and \( H = H(t, x) := \sum_{n \geq 1} H_{1n-1}(t) x^n \). Solving this system for \( G \) and \( H \), we get the generating function in equation (3.5).

5 Concluding remarks

A more general invariant to study of the descent polytopes is the flag \( f \)-vector. The flag \( f \)-vector is efficiently encoded by the \textit{cd}-index. Is there a way to describe the \textit{cd}-index of the descent polytope in terms of the \textit{xy}-word \( u \)?

Setting \( t = 1 \) in the polynomial \( F_v(t) \) we obtain the number of faces of the descent polytope \( DP_v \). Especially, for the alternating word \( z_n \) we obtain the sequence \( \{F_{z_n-1}(1)\}_{n \geq 1} = 3, 7, 19, 51, … \). This sequence has a different combinatorial interpretation, as it matches the sequence A052948 in the Online Encyclopedia of Integer Sequences [5] defined as the number of paths from \((0, 0)\) to \((n + 1, 0)\) with allowed steps \((1, 1)\), \((1, 0)\) and \((1, -1)\) contained within the region \(-2 \leq y \leq 2\). The generating function

\[
\frac{1 - 2x^2}{1 - 3x + 2x^3}
\]

given in [5] indeed results if \( t = 1 \) is substituted into (3.5) and the constant 1 is added. Is there a bijective proof? A first step to find such a bijective proof would be to find a statistic on these lattice paths with the same distribution as the dimensions of the faces of the descent polytope \( DP_{z_n-1} \).

For a \textit{xy}-word \( v = v_1 v_2 \cdots v_n \) let \( v^* \) denote the reverse of the word, that is, \( v^* = v_n \cdots v_2 v_1 \). Note that the two descent polytopes \( DP_v \) and \( DP_{v^*} \) only differ by a linear transformation and hence their \( f \)-polynomials agree, that is, \( F_v = F_{v^*} \). However the expressions for the \( f \)-polynomials for \( F_v \) and \( F_{v^*} \) in Corollary 3.3 differ. Is there a bijection between the factorizations of \( v \) and \( v^* \)? The number of factorizations of \( v \) is also equal the number of alternating subwords of \( v \); see Corollary 2.3. This fact also asks for a bijective proof.

More inequalities for the descent statistic has been proved in [3, 4]. Can these inequalities be extended to the \( f \)-polynomial \( F_v \)? For instance, Ira Gessel asked the following question: where does the maximum of the descent set statistic occur when restricting to words \( v \) of length \( n - 1 \) having exactly \( k \) runs of \( x \)'s and \( y \)'s. He conjectured and it was proved in [4] that the maximum occurs at the composition \((r, r+1, \ldots, r+1, r, \ldots, r)\) where \( r = \lfloor (n - 1)/k \rfloor \) and \( a = (n - 1) - r \cdot k \). Would the \( f \)-polynomial be maximized at the same composition?

References

[1] N. G. de Bruijn, Permutations with given ups and downs, \textit{Nieuw Arch. Wiskd.} 18 (1970), 61–65.
[2] R. Ehrenborg, S. Kitaev and P. Perry, A spectral approach to pattern avoiding permutations, preprint 2007.

[3] R. Ehrenborg, M. Levin and M. Readdy, A probabilistic approach to the descent statistic, J. Combin. Theory Ser. A 98 (2002), 150–162.

[4] R. Ehrenborg and S. Mahajan, Maximizing the descent statistic, Ann. Comb. 2 (1998), 111–129.

[5] The Online Encyclopedia of Integer Sequences

[6] I. Niven, A combinatorial problem on finite sequences, Nieuw Arch. Wiskd. 16 (1968), 116–123.

[7] M. A. Readdy, Extremal problems for the Möbius function in the face lattice of the n-octahedron, Discrete Math., Special issue on Algebraic Combinatorics 139 (1995), 361–380.

[8] B. E. Sagan, Y.-N. Yeh and G. Ziegler, Maximizing Möbius functions on subsets of Boolean algebras, Discrete Math. 126 (1994), 293–311.

[9] R. P. Stanley, Two poset polytopes, Disc. Comput. Geom. 1 (1986), 9–23.

[10] R. P. Stanley, “Enumerative Combinatorics, Vol. I,” Wadsworth and Brooks/Cole, Pacific Grove, 1986.

[11] G. Viennot, Permutations ayant une forme donnée, Discrete Math 26 (1979), 279–284.

D. Chebikin, Department of Mathematics, MIT, Cambridge, MA 02139, chebikin@gmail.com.
R. Ehrenborg, Department of Mathematics, University of Kentucky, Lexington, KY 40506, jrge@ms.uky.edu.