DUMONT'S STATISTIC ON WORDS

MARK SKANDERA

ABSTRACT. We define *Dumont's statistic* on the symmetric group S_n to be the function dmc: $S_n \to \mathbb{N}$ which maps a permutation σ to the number of distinct nonzero letters in code(σ). Dumont showed that this statistic is Eulerian [5]. Naturally extending Dumont's statistic to the rearrangement classes of arbitrary words, we create a generalized statistic which is again Eulerian. As a consequence, we show that for each distributive lattice $J(P)$ which is a product of chains, there is a poset Q such that the f-vector of Q is the h-vector of $J(P)$. This strengthens for products of chains a result of Stanley concerning the flag h-vectors of Cohen-Macaulay complexes [9, Cor. 4.5]. We conjecture that the result holds for all finite distributive lattices.

1. INTRODUCTION

Let S_n be the symmetric group on n letters, and let us write each permutation π in S_n in one line notation: $\pi = \pi_1 \cdots \pi_n$. We call position i a *descent* in π if $\pi_i > \pi_{i+1}$, and an *excedance* in π if $\pi_i > i$. Counting descents and excedances, we define the permutation statistics des and exc by

$$\begin{aligned} \mathrm{des}(\pi) &= \#\{i | \pi_i > \pi_{i+1}\}. \\ \mathrm{exc}(\pi) &= \#\{i | \pi_i > i\}. \end{aligned}$$

It is well known that the number of permutations in S_n with k descents equals the number of permutations in S_n with k excedances. This number is often denoted $A(n, k+1)$ and the generating function

$$A_n(x) = \sum_{k=0}^{n-1} A(n, k+1) x^{k+1} = \sum_{\pi \in S_n} x^{1+\mathrm{des}(\pi)} = \sum_{\pi \in S_n} x^{1+\mathrm{exc}(\pi)}$$

is called the nth *Eulerian polynomial*. Any permutation statistic *stat* satisfying

$$A_n(x) = \sum_{\pi \in S_n} x^{1+\mathrm{stat}(\pi)}$$

is called *Eulerian*.

Mark Skandera, Dept. of Mathematics, Massachusetts Institute of Technology
skan@math.mit.edu.

A third Eulerian statistic, essentially defined by Dumont [5], counts the number of distinct nonzero letters in the code of a permutation. We define code(π) to be the word $c_1 \cdots c_n$, where

$$c_i = \#\{j > i | \pi_j < \pi_i\}.$$

Denoting Dumont's statistic by *dmc*, we have

$$\mathrm{dmc}(\pi) = \#\{\ell \neq 0 | \ell \text{ appears in code}(\pi)\}.$$

Example 1.1.

$$\begin{array}{rcccccccccc} \pi & = & 2 & 8 & 4 & 3 & 6 & 7 & 9 & 5 & 1 \\ \mathrm{code}(\pi) & = & 1 & 6 & 2 & 1 & 2 & 2 & 2 & 1 & 0 \end{array}$$

The distinct nonzero letters in code(π) are $\{1, 2, 6\}$. Thus, dmc(π) = 3.

Dumont showed bijectively that the statistic dmc is Eulerian, i.e. that

$$\#\{\pi \in S_n | \mathrm{dmc}(\pi) = k\} = \#\{\pi \in S_n | \mathrm{des}(\pi) = k\}.$$

While few researchers have found an application for dmc since [5], Foata [7] proved the following equidistribution result involving the statistics INV (inversions) and MAJ (major index). These two statistics belong to the class of *Mahonian* statistics. (See [7] for further information.)

Theorem 1.1. *The Eulerian-Mahonian statistic pairs* (des, INV) *and* (dmc, MAJ) *are equally distributed on S_n, i.e.*

$$\#\{\pi \in S_n | \mathrm{des}(\pi) = k; \mathrm{INV}(\pi) = p\} = \#\{\pi \in S_n | \mathrm{dmc}(\pi) = k; \mathrm{MAJ}(\pi) = p\}.$$

Note that the statistics des, exc, and dmc are all defined on a permutation π to be the cardinalities of certain sets associated to π. To name these sets, we define the *descent set* and *excedance set* of π to be the sets of descents and excedances in π. We denote these by $D(\pi)$ and $E(\pi)$, respectively. Similarly, we define the *letter set* of a word to be the set of its nonzero letters, and will denote the letter set of code(π) by $LC(\pi)$. Thus,

$$\begin{aligned} \mathrm{des}(\pi) &= |D(\pi)|, \\ \mathrm{exc}(\pi) &= |E(\pi)|, \\ \mathrm{dmc}(\pi) &= |LC(\pi)|. \end{aligned}$$

It is easy to see that for every subset T of $[n-1] = \{1, \ldots, n-1\}$, there are permutations π, σ, and ρ in S_n satisfying

$$T = D(\pi) = E(\sigma) = LC(\rho).$$

In fact, Dumont's original bijection [5] shows that for each such subset T we have

$$\#\{\pi \in S_n | E(\pi) = T\} = \#\{\pi \in S_n | LC(\pi) = T\}.$$

However, the analogous statement involving $D(\pi)$ is not true.

Generalizing permutations on n letters are *words* $w = w_1 \cdots w_m$ on n letters, where $m \geq n$. We will assume that each letter in $[n]$ appears at least once in w. Generalizing the symmetric group S_n, we define the *rearrangement class* of w by

$$R(w) = \{w_{\sigma^{-1}(1)} \cdots w_{\sigma^{-1}(m)} | \sigma \in S_m\}.$$

Each element of $R(w)$ is called a *rearrangement* of w.

Many definitions pertaining to S_n generalize immediately to the rearrangement class of any word. In particular, the definitions of descent, descent set, code, letter set of a code, and Dumont's statistic remain the same for words as for permutations. Generalization of excedances requires only a bit of effort.

For any word w, denote by $\bar{w} = \bar{w}_1, \ldots, \bar{w}_m$ the unique *nondecreasing rearrangement* of w. We define position i to be an *excedance* in w if $w_i > \bar{w}_i$. Thus,

$$\mathrm{exc}(w) = \#\{i | w_i > \bar{w}_i\}.$$

If position i is an excedance in word w, we will refer to the letter w_i as the *value* of excedance i. One can see word excedances most easily by associating to the word w the *biword*

$$\begin{pmatrix} \bar{w} \\ w \end{pmatrix} = \begin{pmatrix} \bar{w}_1 \cdots \bar{w}_m \\ w_1 \cdots w_m \end{pmatrix}.$$

Example 1.2. Let $w = 312312311$. Then,

$$\begin{pmatrix} \bar{w} \\ w \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 2 & 2 & 3 & 3 & 3 \\ 3 & 1 & 2 & 3 & 1 & 2 & 3 & 1 & 1 \end{pmatrix}$$

Thus, $E(w) = \{1, 3, 4\}$ and $\mathrm{exc}(w) = 3$. The corresponding excedance values are 3, 2, and 3.

We will use biwords not only to expose excedances, but to define and justify maps in Sections 3 and 4. In particular, if $u = u_1, \ldots, u_m$ and $v = v_1, \ldots, v_m$ are words and y is the biword

$$y = \begin{pmatrix} u \\ v \end{pmatrix},$$

then we will define *biletters* $y_1, \ldots, y_m$ by

$$y_i = \begin{pmatrix} u_i \\ v_i \end{pmatrix},$$

and will define the rearrangement class of y by

$$R(y) = \{y_{\sigma^{-1}(1)} \cdots y_{\sigma^{-1}(m)} | \sigma \in S_m\}.$$

MacMahon, studying the distributions of word statistics on the rearrangement classes of words [8], demonstrated that the word statistics des and exc are equally distributed on the rearrangement class of any word w. That is, for any word w,

$$\#\{y \in R(w) | \mathrm{exc}(y) = k\} = \#\{y \in R(w) | \mathrm{des}(y) = k\}.$$

Analogously to the case of permutation statistics, a word statistic *stat* is called *Eulerian* if it satisfies

$$\#\{y \in R(w) | \mathrm{stat}(y) = k\} = \#\{y \in R(w) | \mathrm{des}(y) = k\}$$

for any word w and any nonnegative integer k.

In Section 2, we state and prove our main result: that dmc is Eulerian as a word statistic. Our bijection is different than that of Dumont [5], which doesn't generalize obviously to the case of arbitrary words. Applying the main theorem to a problem involving f-vectors and h-vectors of partially ordered sets, we state a second theorem in Section 3. This result strengthens a special case of a result of Stanley [9] concerning the flag h-vectors of balanced Cohen-Macaulay complexes. We prove the second theorem in Sections 4 and 5, and finish with some related open questions in Section 6.

2. Main theorem

As implied in Section 1, we define Dumont's statistic on an arbitrary word w to be the number of distinct nonzero letters in $\mathrm{code}(w)$.

$$\mathrm{dmc}(w) = |LC(w)|.$$

This generalized statistic is Eulerian.

Theorem 2.1. *If $R(w)$ is the rearrangement class of an arbitrary word w and k is any nonnegative integer, then*

$$\#\{v \in R(w) | \mathrm{dmc}(v) = k\} = \#\{v \in R(w) | \mathrm{exc}(v) = k\}.$$

Our bijective proof of the theorem depends upon an encoding of a word which we call the *excedance table*.

Definition 2.1. Let $v = v_1 \cdots v_m$ be an arbitrary word and let let $c = c_1 \cdots c_m$ be its code. Define the excedance table of v to be the unique word $\mathrm{etab}(v) = e_1 \cdots e_m$ satisfying

1. If i is an excedance in v, then $e_i = i$.
2. If $c_i = 0$, then $e_i = 0$.
3. Otherwise, e_i is the c_ith excedance of v having value at least v_i.

Note that $\mathrm{etab}(v)$ is well defined for any word v. In particular, if i is not an excedance in v and if $c_i > 0$, then there are at least c_i excedances in v having value at least v_i. To see this, define

$$k = \#\{j \in [m] | v_j < v_i\}.$$

Since c_i of the letters $\bar{v}_1, \ldots, \bar{v}_k$ appear to the right of position i in v, then at least c_i of the letters $\bar{v}_{k+1}, \ldots, \bar{v}_m$ must appear in the first k positions of v. The positions of these letters are necessarily excedances in v.

An important property of the excedance table is that the letter set of $\mathrm{etab}(v)$ is precisely the excedance set of v.

Example 2.2. Let $v = 514514532$, and define $c = \mathrm{code}(v)$. Calculating $e = \mathrm{etab}(v)$, we have

$$\begin{pmatrix} \bar{v} \\ v \\ c \\ e \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 & 3 & 4 & 4 & 5 & 5 & 5 \\ 5 & 1 & 4 & 5 & 1 & 4 & 5 & 3 & 2 \\ 6 & 0 & 3 & 4 & 0 & 2 & 2 & 1 & 0 \\ 1 & 0 & 3 & 4 & 0 & 3 & 4 & 1 & 0 \end{pmatrix}.$$

Calculation of $e_1, \ldots, e_5$ and e_9 is straightforward since the positions $i = 1, \ldots, 5$ and 9 are excedances in v or satisfy $c_i = 0$. We calculate e_6, e_7, and e_8 as follows. Since $c_6 = 2$, and the second excedance in v with value at least $v_6 = 4$ is 3, we set $e_6 = 3$. Since $c_7 = 2$, and the second excedance in v with value at least $v_7 = 5$ is 4, we set $e_7 = 4$. Since $c_8 = 1$, and the first excedance in v with value at least $v_8 = 3$ is 1, we set $e_8 = 1$.

We prove Theorem 2.1 with a bijection $\theta : R(w) \to R(w)$ which satisfies

$$E(v) = LC(\theta(v)), \tag{2.1}$$

and therefore

$$\mathrm{exc}(v) = \mathrm{dmc}(\theta(v)). \tag{2.2}$$

Definition 2.3. Let $w = w_1 \cdots w_m$ be any word. Define the map $\theta : R(w) \to R(w)$ by applying the following procedure to an arbitrary element v of $R(w)$.

1. Define the biword $z = \begin{pmatrix} v \\ \mathrm{etab}(v) \end{pmatrix}$.
2. Let y be the unique rearrangement of z satisfying $y = \begin{pmatrix} u \\ \mathrm{code}(u) \end{pmatrix}$.
3. Set $\theta(v) = u$.

Construction of y is quite straightforward. Let $e = e_1 \cdots e_m = \mathrm{etab}(v)$, and linearly order the biletters $z_1, \ldots, z_m$ by setting $z_i < z_j$ if

$$v_i < v_j, \text{ or}$$
$$v_i = v_j \text{ and } e_i > e_j.$$

Break ties arbitrarily. Considering the biletters according to this order, insert each biletter z_i into y to the left of e_i previously inserted biletters.

Example 2.4. Let v and e be as in Example 2.2. To compute $\theta(v)$, we define

$$z = \begin{pmatrix} v \\ e \end{pmatrix} = \begin{pmatrix} 5 & 1 & 4 & 5 & 1 & 4 & 5 & 3 & 2 \\ 1 & 0 & 3 & 4 & 0 & 3 & 4 & 1 & 0 \end{pmatrix}.$$

We consider the biletters of z in the order

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 5 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 1 \end{pmatrix},$$

and insert them individually into y:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 11 \\ 00 \end{pmatrix}, \begin{pmatrix} 112 \\ 000 \end{pmatrix}, \begin{pmatrix} 1132 \\ 0010 \end{pmatrix}, \begin{pmatrix} 14132 \\ 03010 \end{pmatrix}, \ldots$$

Finally we obtain

$$y = \begin{pmatrix} u \\ e \end{pmatrix} = \begin{pmatrix} 1 & 4 & 5 & 5 & 4 & 1 & 3 & 5 & 2 \\ 0 & 3 & 4 & 4 & 3 & 0 & 1 & 1 & 0 \end{pmatrix},$$

and set $\theta(v) = 145541352$.

It is easy to see that any biword z has at most one rearrangement y satisfying Definition 2.3 (2). Such a rearrangement exists if and only if we have

$$e_i \leq \#\{j \in [m] | v_j < v_i\}, \text{ for } i = 1, \ldots, m, \tag{2.3}$$

or equivalently, if and only if

$$\bar{v}_{e_i} < v_i, \text{ for } i = 1, \ldots, m, \tag{2.4}$$

where we define $\bar{v}_0 = 0$ for convenience.

Observation 2.2. *Let $v = v_1 \cdots v_m$ be any word and let $e = \mathrm{etab}(v)$. Then we have*

$$e_i \leq \#\{j \in [m] | v_j < v_i\}, \text{ for } i = 1, \ldots, m.$$

Proof. If i is an excedance in v, then $e_i = i$ and $\bar{v}_1 \leq \cdots \leq \bar{v}_i < v_i$. If $c_i = 0$, then $e_i = 0$. Otherwise, define

$$k = \#\{j \in [m] | v_j < v_i\}.$$

By the discussion following Definition 2.1, at least c_i of the positions $1, \ldots, k$ are excedances in v with values at least v_i. The letter e_i, being one of these excedances, is therefore at most k. □

Thus the map θ is well defined and satisfies (2.1) and (2.2). We invert θ as follows.

Proposition 2.3. *Let y be a biword satisfying*

$$y = \begin{pmatrix} u \\ c \end{pmatrix} = \begin{pmatrix} u \\ \mathrm{code}(u) \end{pmatrix}.$$

The following procedure produces a rearrangement z of y satisfying

$$z = \begin{pmatrix} v \\ e \end{pmatrix} = \begin{pmatrix} v \\ \mathrm{etab}(v) \end{pmatrix}.$$

1. *For each letter ℓ in $L(c)$, find the greatest index i satisfying $c_i = \ell$, and define $z_\ell = y_i$. Let S be the set of such greatest indices, and define $T = [m] \setminus S$.*
2. *For each index $i \in T$, define*
$$d_i = \begin{cases} \#\{j \in S | c_j \leq c_i; u_j \geq u_i\}, & \text{if } c_i > 0, \\ 0, & \text{otherwise.} \end{cases}$$
3. *Let $(y_{\sigma^{-1}(i)})_{i \in T}$ be the unique rearrangement of $(y_i)_{i \in T}$ satisfying*
$$(d_{\sigma^{-1}(i)})_{i \in T} = \mathrm{code}((u_{\sigma^{-1}(i)})_{i \in T}).$$
4. *Insert the biletters $(y_{\sigma^{-1}(i)})_{i \in T}$ in order into the remaining positions of z.*

Proof. First we claim that the procedure is well defined. In particular, we may perform step 3 because we have

$$d_i \leq \#\{j \in T | u_j < u_i\}, \text{ for each } i \in T,$$

as required by (2.3). To see that this is the case, let i be an index in T with $c_i > 0$. In step 1 we have placed d_i biletters y_j with $u_j > u_i$ into positions $1, \ldots, c_i$ of z. Thus, at least d_i biletters y_j with $u_j \leq \bar{u}_{c_i}$ have *not* been placed into these positions. The index j of any such biletter belongs to S only if $c_j > c_i$. However, since $\bar{u}_{c_j} < u_j \leq \bar{u}_{c_i} < u_i$, we have $c_j < c_i$. Thus, j belongs to T.

Next we claim that $e = \mathrm{etab}(v)$. The positions $L(c) = \{c_j | j \in S\}$ are excedances in v, because for each index j in S, we have $v_{c_j} = u_j > \bar{u}_{c_j} = \bar{v}_{c_j}$. As required, we also have $e_{c_j} = c_j$.

The positions $L(c)$ are in fact the only excedances in v. For each index j in T, denote by $\phi(j)$ the position of z into which we have placed y_j. Supposing that some indices $\{\phi(j) | j \in T\}$ are excedances in v, choose $i \in T$ so that $\phi(i)$ is the leftmost of these excedances and define

$$k = \#\{j \in [m] | u_j < u_i\}.$$

Then,

$$k > \#\{j \in S | c_j \leq k\} + \#\{j \in T | \sigma(j) < \sigma(i)\}. \tag{2.5}$$

Since $c_i \leq k$ by (2.3), we have

$$\#\{j \in S | c_j \leq k\} = \#\{j \in S | c_j \leq c_i\} + \#\{j \in S | c_i < c_j \leq k\},$$

and by the definition of σ, we have

$$\begin{aligned}\#\{j \in T | \sigma(j) < \sigma(i)\} &= \#\{j \in T | u_j < u_i\} - d_i \\ &= \#\{j \in T | u_j < u_i\} - \#\{j \in S | c_j \leq c_i; u_j \geq u_i\}.\end{aligned}$$

Using these identities to simplify (2.5), we obtain

$$\#\{j \in S | u_j < u_i; c_j > c_i\} > \#\{j \in S | c_i < c_j \leq k\}. \tag{2.6}$$

If j belongs to the set on the left hand side of (2.6) and satisfies $c_j > k$, then we have

$$u_j > \bar{u}_{c_j} \geq \bar{u}_k = u_i - 1,$$

a contradiction. On the other hand, if each index j in this set satisfies $c_j \leq k$, then we have the inclusion

$$\{j \in S | u_j < u_i; c_j > c_i\} \subset \{j \in S | c_i < c_j \leq k\},$$

which contradicts (2.6). We conlcude that the set $\{\phi(j) | j \in T\}$ is precisely the set of non-excedances in v.

To complete the proof, let c' be the code of v. We must show that for each index $i \in T$, we have

$$e_{\phi(i)} = c_i = \begin{cases} \text{the } c'_{\phi(i)}\text{th excedance in } v \text{ having value at least } u_i, & \text{if } c'_{\phi(i)} > 0, \\ 0, & \text{otherwise.} \end{cases}$$

By our definition of the sequence $(d_i)_{i \in T}$, it suffices to show that $c'_{\phi(i)} = d_i$ for each index i. The subword $v_{\phi(i)+1} \cdots v_m$ of v includes d_i letters $v_{\phi(j)}$ with $j \in T$ and $v_{\phi(j)} < v_{\phi(i)}$. On the other hand, any excedance in v to the right of $\phi(i)$ has value *greater* than $v_{\phi(i)}$. We conclude that $c'_{\phi(i)} = d_i$. □

The above procedure inverts θ because the biword z it produces is the *unique* rearrangement of y having the desired properties.

Proposition 2.4. *Let $v = v_1 \cdots v_m$ be an arbitrary word, and define*

$$z = \begin{pmatrix} v \\ e \end{pmatrix} = \begin{pmatrix} v \\ \mathrm{etab}(v) \end{pmatrix}.$$

If there is any rearrangement z' of z satisfying

$$z' = \begin{pmatrix} v' \\ e' \end{pmatrix} = \begin{pmatrix} v' \\ \mathrm{etab}(v') \end{pmatrix},$$

then $z' = z$.

Proof. Let L be the letter set of e. By Definition 2.1, we must have $E(v) = E(v') = L$. Let i be an excedance of v and v'. By condition (1) of Definition 2.1 we must have $e_i = e'_i = i$, and by condition (3) the upper letters v_i and v'_i must be as large as possible. Thus, $(z_i)_{i\in L} = (z'_i)_{i\in L}$.

Let $T = [m] \setminus L$ be the set of non-excedance positions of v and v', and consider the corresponding subsequences of biletters $(z_i)_{i\in T}$ and $(z'_i)_{i\in T}$. By condition (3) of Definition 2.1, the codes of $(v_i)_{i\in T}$ and $(v'_i)_{i\in T}$ are determined by the excedances and excedance values in v and v'. Thus, the two codes must be identical. Applying the argument following Example 2.4, we conclude that $(z_i)_{i\in T} = (z'_i)_{i\in T}$. □

3. An application of Dumont's statistic

As an application of Dumont's (generalized) statistic, we will strengthen a special case of a result of Stanley [9, Cor. 4.5] concerning f-vectors and h-vectors of simplicial complexes.

Given a $(d-1)$ dimensional simplicial complex Σ, we define its f-vector to be

$$f_\Sigma = (f_{-1}, f_0, f_1, \ldots, f_{d-1}),$$

where f_i counts the number of i-dimensional faces of Σ. By convention, $f_{-1} = 1$. Similarly, we may define the f-vector of a poset P by identifying P with its *order complex* $\Delta(P)$. (See [10, p. 120].) That is, we define

$$f_P = f_{\Delta(P)} = (f_{-1}, f_0, f_1, \ldots, f_{d-1}),$$

where f_i counts the number of $(i+1)$-element chains of P. Again, $f_{-1} = 1$ by convention.

In abundant research papers, authors have considered the f-vectors of various classes of complexes and posets, and have conjectured or obtained significant information about the coefficients. (See [1], [2], [11, Ch. 2,3].) Such information includes linear relationships between coefficients and properties such as symmetry, log concavity and unimodality.

Related to the f-vector f_Σ is the h-vector $h_\Sigma = (h_0, h_1, \ldots, h_d)$, which we define by

$$\sum_{i=0}^{d} f_{i-1}(x-1)^{d-i} = \sum_{i=0}^{d} h_i x^{d-i}.$$

From this definition, it is clear that knowing the h-vector of a complex is equivalent to knowing the f-vector. For some conditions on a simplicial complex, one can show that its h-vector is the f-vector of another complex. Specifically, we have the following result, due to Stanley [9, Cor. 4.5].

Theorem 3.1. *If Σ is a balanced Cohen-Macaulay complex, then its h-vector is the f-vector of some simplicial complex Γ.*

We define a simplicial complex to be *Cohen-Macaulay* if it satisfies a certain topological condition ([11, p. 61]), and *balanced* if we can color the vertices with d colors such that no face contains two vertices of the same color ([11, p. 95]). The class of balanced Cohen-Macaulay complexes is quite important because it includes the order complexes of all distributive lattices. The distributive lattices, in turn, contain information about all posets. (See [10, Ch. 3].)

By placing an additional restriction on the complex Σ, one arrives at a special case of the theorem which has an elegant bijective proof. Let us require that Σ be the order complex of a distributive lattice $J(P)$. In this case, $h_\Sigma = h_{J(P)}$ counts the number of linear extensions of P by descents. (See [4].) That is, h_k is the number of linear extensions of P with k descents. Therefore, Theorem 3.1 asserts that for any poset P, there is a bijective correspondence between linear extensions of P with k descents and $(k-1)$-faces of some simplicial complex Γ.

$$\{\pi | \pi \text{ a linear extension of } P;\, \mathrm{des}(\pi) = k\} \stackrel{1-1}{\longleftrightarrow} \{\sigma | \sigma \text{ a } (k-1)\text{-face of } \Gamma\}.$$

Using [3, Remark 6.6] and [6, Cor. 2.2], one can construct a family $\{\Xi_n\}_{n>0}$ of simplical complexes such that for any poset P on n elements, the complex Γ corresponding to $\Sigma = \Delta(J(P))$ is a subcomplex of Ξ_n.

On the other hand, any additional restriction placed on the complex Σ in Theorem 3.1 should allow us to prove more than a special case of the theorem. It should allow us to strengthen the special case by asserting specific properties of the complex Γ in the conclusion of the theorem. In particular, let us require that Σ be the order complex of a distributive lattice $J(P)$ which is a product of chains. (See [10, Ch. 3] for definitions.) We will prove the following result.

Theorem 3.2. *Let the distributive lattice $J(P)$ be a product of chains. Then there is a poset Q such that the h-vector of $J(P)$ is the f-vector of Q.*

Let us reconsider this theorem in terms of rearrangements of words. If $J(P)$ is a product of chains having cardinalities $p_1 + 1, \ldots, p_n + 1$, then P is the disjoint *sum* of chains $(\mathbf{p_1} + \cdots + \mathbf{p_n})$. It is not difficult to see that linear extensions of P are in bijective correspondence with rearrangements of the word $w = 1^{p_1} \cdots n^{p_n}$. Combining this observation with Theorem 2.1, we restate Theorem 3.2 in terms of Dumont's statistic.

Proposition 3.3. *Let w be any word and define the vector $h = (h_0, \ldots, h_d)$ by*

$$h_i = \#\{u \in R(w) | \mathrm{dmc}(u) = i\},$$

where d is the maximum cardinality of $LC(u)$ over all rearrangements u of w. Then, there is a poset Q whose f-vector is h.

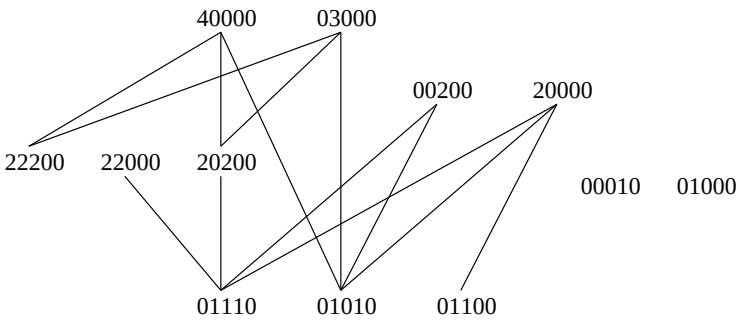

FIGURE 3.1

Let $C(w)$ be the set of codes of all rearrangements of w.

$$C(w) = \{\text{code}(u) | u \in R(w)\}.$$

Proposition 3.3 asserts that for any word w, there is a bijection between k-letter words in $C(w)$ and k-element chains in some poset Q.

$$\{c \in C(w) | c \text{ a } k\text{-letter code}\} \overset{1-1}{\longleftrightarrow} \{(v_1 <_Q \cdots <_Q v_k) \in \Delta(Q)\}.$$

To prove the proposition, and therefore Theorem 3.2, we will construct such a poset $Q = Q(w)$ as in Definition 3.1. In Sections 4 and 5 we will give an explicit bijection $\Psi : C(w) \to \Delta(Q)$, taking k-letter codes in $C(w)$ to k-element chains in Q.

Definition 3.1. For any word w, let Q be the subset of one-letter codes in $C(w)$, and let c and c' be codes in Q whose letters are ℓ and ℓ', respectively. Define $c <_Q c'$ if

1. $\ell < \ell'$.
2. The multiplicity of ℓ in c is strictly *greater* than that of ℓ' in c'.
3. For each position i such that $c'_i = \ell'$, we have $c_{i+\ell'-\ell} = \ell$.

Figure 3.1 shows the poset Q corresponding to the word 11223.

4. The chain map Ψ

Fix a nondecreasing word $w = w_1 \cdots w_m$ on n letters, and define the poset Q as in Definition 3.1. We will define a *chain map* $\Psi : C(w) \to \Delta(Q)$ which will identify a code c with a chain

$$\Psi(c) = v_1 <_Q \cdots <_Q v_k,$$

of elements in Q. If c is a code on the k letters $\ell_1 < \cdots < \ell_k$, then each poset element v_i will be a code whose unique nonzero letter is ℓ_i. Specifically, we will determine v_i by applying a *vertex map* $\psi_{\ell_i} : C(w) \to Q$ to c.

$$v_i = \psi_{\ell_i}(c).$$

After proving that $\psi_{\ell_i}(c) <_Q \psi_{\ell_j}(c)$ whenever $\ell_i < \ell_j$, we will define the chain map to be a product of vertex maps,

$$\Psi(c) = \psi_{\ell_1}(c) <_Q \cdots <_Q \psi_{\ell_k}(c).$$

We begin by observing that several simple operations on codes in $C(w)$ yield other codes in $C(w)$.

Observation 4.1. *Let u be a rearrangement of w and let $c = \mathrm{code}(u)$.*

1. *If $c_i > c_{i+1}$, then the word*
$$c_1 \cdots c_{i-1} \cdot c_{i+1} \cdot (c_i - 1) \cdot c_{i+2} \cdots c_m$$
belongs to $C(w)$.
2. *If for some $r > i$, c_i is strictly less than $c_{i+1}, \ldots, c_r$, and $c_i > c_{r+1}$, then the word*
$$c_1 \cdots c_{i-1} \cdot c_{r+1} \cdot c_{i+1} \cdots c_r \cdot (c_i - 1) \cdot c_{r+2} \cdots c_m$$
belongs to $C(w)$.
3. *If $c_i < c_{i+1}$, or if $c_i = c_{i+1}$ and $u_i < u_{i+1}$, then the word*
$$c_i \cdots c_{i-1} \cdot (c_{i+1} + 1) \cdot c_i \cdot c_{i+2} \cdots c_m$$
belongs to $C(w)$.

Proof. (1) This is the code of the word obtained by switching the letters u_i and u_{i+1}, in the case that $u_i > u_{i+1}$.

(2) This is the code of the word obtained by switching the letters u_i and u_{r+1}, in the case that u_i is less than $u_{i+1}, \ldots, u_r$, c_i is less than $c_{i+1}, \ldots, c_r$, and $u_i > u_{r+1}$.

(3) This is the code of the word obtained by switching the letters u_i and u_{i+1}, in the case that $u_i < u_{i+1}$. □

Using this observation we will define two families of maps from $C(w)$ to itself, $\lambda_1, \ldots, \lambda_{m-1}$ and $\mu_1, \ldots, \mu_{m-1}$. Then, composing maps from these two families, we will define the family of vertex maps $\psi_1, \ldots, \psi_{m-1}$.

The map $\lambda_{\ell_i} : C(w) \to C(w)$ removes from a code c all letters ℓ_j which are greater than ℓ_i. It essentially changes each such letter ℓ_j to ℓ_i and moves it $\ell_j - \ell_i$ places to the right in c. If we identify c with the k-element chain $v_1 <_Q \cdots <_Q v_k$, then we will identify $\lambda_{\ell_i}(c)$ with the i-element subchain $v_1 <_Q \cdots <_Q v_i$.

Definition 4.1. Let ℓ be a nonzero letter. Define the map $\lambda_\ell : C(w) \to C(w)$ by performing the following procedure on a code c.

For $i = m, m-1, \ldots, 1$, if $c_i > \ell$, then

1. Set $\delta = c_i - \ell$.
2. Redefine $c = c_1 \cdots c_{i-1} \cdot c_i \cdots c_{i+\delta} \cdot \ell \cdot c_{i+\delta+1} \cdots c_m$.

Analogous to λ_{ℓ_i}, the map $\mu_{\ell_i} : C(w) \to C(w)$ removes all letters which are *less* than ℓ_i. It simply changes each lower letter to 0. Thus, if we identify c with the k-element chain $v_1 <_Q \cdots <_Q v_k$, then we will identify $\mu_{\ell_i}(c)$ with the $(k-i+1)$-element subchain $v_i <_Q \cdots <_Q v_k$.

Definition 4.2. Let ℓ be a nonzero letter. Define the map $\mu_\ell : C(w) \to C(w)$ by $\mu_\ell(c) = a_1 \cdots a_m$, where

$$a_i = \begin{cases} 0, & \text{if } c_i \leq \ell, \\ c_i, & \text{otherwise.} \end{cases}$$

The maps $\lambda_1, \ldots, \lambda_{m-1}$, and $\mu_1, \ldots, \mu_{m-1}$ are well defined, for their definitions are merely repeated applications of Observation 4.1 (1) and (2). Note that the composition $\mu_\ell \lambda_\ell$ produces a code on the single letter ℓ. This code is an element of Q, and a *vertex* of $\Delta(Q)$.

Definition 4.3. Let ℓ be a nonzero letter. Define the *vertex map* $\psi_\ell : C(w) \to Q$ by

$$\psi_\ell = \mu_\ell \lambda_\ell.$$

It is easy to see that $\lambda_\ell^2 = \lambda_\ell$, and therefore that $\psi_\ell \lambda_\ell = \psi_\ell$. These and the following relations will be essential in establishing a bijection between $C(w)$ and $\Delta(Q)$.

Proposition 4.2. *Let ℓ and ℓ' be letters, $1 \leq \ell < \ell' \leq n$. The maps $\lambda_\ell, \lambda_{\ell'}, \psi_\ell$, and $\psi_{\ell'}$ satisfy the relations*

1. $\lambda_{\ell'} \lambda_\ell = \lambda_\ell \lambda_{\ell'} = \lambda_\ell$.
2. $\psi_\ell \lambda_{\ell'} = \psi_\ell$.
3. $\psi_\ell(c) <_Q \psi_{\ell'}(c)$, *if c contains both letters.*

Proof. (1) Let $c = \mathrm{code}(u)$ be an element of $C(w)$. By the comments following Definition 4.2, we may interpret $\lambda_\ell(c)$ as follows. Define $b = b_1 \cdots b_m$ by

$$b_i = \begin{cases} \ell, & \text{if } c_i > \ell, \\ c_i, & \text{otherwise,} \end{cases}$$

and rearrange the biword

$$\begin{pmatrix} u \\ b \end{pmatrix}$$

as

$$\begin{pmatrix} u' \\ b' \end{pmatrix}$$

so that $b' = \mathrm{code}(u')$. Then, $b' = \lambda_\ell(c)$.

It is not hard to see that there is a unique such rearrangement. Using this interpretation, it is easy to see that $\lambda_{\ell'}\lambda_\ell$, $\lambda_\ell\lambda_{\ell'}$, and λ_ℓ describe the same procedure.

(2) Using (1), we have $\psi_\ell\lambda_{\ell'} = \mu_\ell\lambda_\ell\lambda_{\ell'} = \mu_\ell\lambda_\ell = \psi_\ell$.

(3) We may assume that ℓ' is the greatest letter in c. (Otherwise, we define $d = \lambda_{\ell'}(c)$ and note that $\psi_\ell(c) = \psi_\ell(d)$ and $\psi_{\ell'}(c) = \psi_{\ell'}(d)$.) Let $e = \psi_\ell(c)$ and $e' = \psi_{\ell'}(c)$. Clearly, the multiplicity of ℓ in e is strictly greater than that of ℓ' in e', for

$$\#\{i|e_i = \ell\} = \#\{i|c_i \geq \ell\} > \#\{i|c_i \geq \ell'\} = \#\{i|e'_i = \ell'\}.$$

Next, we show that for any position i of e' satisfying $e'_i = \ell$, we must have $e_{i+\ell'-\ell} = \ell$. Since by assumption, ℓ' is the greatest letter in c, we have $e'_i = \ell'$ if and only if $c_i = \ell'$. To find e, we first calculate $\lambda_\ell(c)$ by the procedure of Definition 4.1. At each iteration i such that $c_i = \ell'$, we place the letter ℓ into position $i + \ell' - \ell$ of $\lambda_\ell(c)$. This position will not be altered by iterations $i-1, \ldots, 1$, since all letters of c are no greater than ℓ'. Finally, since $e = \mu_\ell\lambda_\ell(c)$, and μ_ℓ changes only those letters less than ℓ, we see that $e_{i+\ell'-\ell} = \ell$ for every position i such that $e_i = \ell'$. □

Now we may define the map Ψ.

Definition 4.4. Define the *chain* map $\Psi : C(w) \to \Delta(Q)$ by

$$\Psi(c) = \psi_{\ell_1}(c) <_Q \cdots <_Q \psi_{\ell_k}(c),$$

where $\ell_1 < \cdots < \ell_k$ are the distinct nonzero letters in c.

5. Inverting Ψ

We will define a map $\Phi : \Delta(Q) \to C(w)$ which takes a k-element chain in Q to a k-letter code in $C(w)$. By demonstrating that Φ inverts Ψ, we will complete the proof of Proposition 3.3.

We begin by defining an operation $\vee : C(w) \times Q \to C(w)$ which joins a new letter to a code.

Definition 5.1. Let $d \in Q$ be a code whose unique nonzero letter is ℓ', and let $c \in C(w)$ be a code whose greatest letter is ℓ. Assume that $\psi_\ell(c) <_Q d$. Define the code $e = c \vee d$ by the following procedure.

1. For each i such that $d_i = \ell'$, set $e_i = \ell'$ and cross out the ℓ in position $i + \delta$ of c.
2. Fill the remaining positions of e with the remaining components of c, in order.

Note that $L(e) = L(c) \cup \{\ell\}$. Therefore, we may map a chain of k one-letter codes to a single k-letter code by iterating the join operation.

Definition 5.2. Let $v_1 <_Q \cdots <_Q v_k$ be a chain of one-letter codes on the letters $\ell_1 < \cdots < \ell_k$, respectively. Define the map $\Phi : \Delta(Q) \to C(w)$ by

$$\Phi(v_1 <_Q \cdots <_Q v_k) = (\cdots((v_1 \vee v_2) \vee v_3)\cdots) \vee v_k.$$

The following proposition shows that the join operation is well defined. It follows that Φ is well defined also.

Proposition 5.1. *If c and d are codes in $C(w)$ satisfying the hypotheses of Definition 5.1, then $c \vee d$ also belongs to $C(w)$.*

Proof. Let u and y be words in $R(w)$ whose codes $c = \mathrm{code}(u)$ and $d = \mathrm{code}(y)$ satisfy the conditions of Definition 5.1. Consider the leftmost position i in c such that $c_i = \ell$ and $d_{i-\delta} = \ell'$. By assumption, $c_{i-1} \leq c_i$. If $c_{i-1} < c_i$, or if $c_{i-1} = c_i$ and $u_{i-1} < u_i$, then we may apply Observation 4.1 (3) $\ell' - \ell$ times to obtain the word

$$c_1 \cdots c_{i-\delta-1} \cdot \ell' \cdot c_{i-\delta+1} \cdots \hat{c_i} \cdots c_m,$$

which belongs to $C(w)$. (Here, $\hat{c_i}$ means that the letter c_i is omitted.) Repeating this process for each such position i, we redefine the join operation. Therefore it suffices to show that for every position i satisfying $d_{i-\delta-1} = 0$, $d_{i-\delta} = \ell'$, and $c_{i-1} = c_i = \ell$, we have $u_{i-1} < u_i$.

Let i be such a position and suppose that $u_{i-1} = u_i$. Since $d_{i-\delta} = \ell'$ and $d_{i-\delta-1} = 0$, there are exactly $i - \delta - 1 + \ell' = i + \ell - 1$ letters in y which are strictly less than y_i. In particular, we have

$$w_{i+\ell-1} < w_{i+\ell}. \tag{5.1}$$

Let k be the number of positions preceeding i such that $u_{i-k} = u_{i-k+1} = \cdots = u_i$ and $c_{i-k} = c_{i-k+1} = \cdots = \ell$. Then there are exactly $i - k - 1 + \ell$ letters in u which are strictly less than u_i $(= u_{i-1} = \cdots = u_{i-k})$. In particular, we have

$$w_{i-k-1+\ell} < w_{i-k+\ell} = w_{i-k+1+\ell} = \cdots = w_{i+\ell},$$

which contradicts (5.1). We conclude that $u_{i-1} < u_i$, and therefore that $c \vee d$ belongs to $C(w)$. □

To begin demonstrating that Φ inverts Ψ we note the following relations satisfied by ψ, λ and $\vee$.

Proposition 5.2. *The pair of maps (ψ, λ) inverts the operation $\vee$ in the following sense.*

1. *Let $c \in C(w)$ be a code with greatest letter ℓ, and let $d \in Q$ be a code with letter $\ell' > \ell$ and satisfying $\psi_\ell(c) <_Q d$. Then we have*

$$\psi_{\ell'}(c \vee d) = d,$$
$$\lambda_\ell(c \vee d) = c.$$

2. *Let $c \in C(w)$ be a code whose greatest two letters are $\ell < \ell'$. Then we have*

$$\lambda_\ell(c) \vee \psi_{\ell'}(c) = c.$$

Proof. (1) Let S be the set of positions of d containing the letter ℓ', and let $\delta = \ell' - \ell$.

Define the words $e = c \vee d$, $d' = \psi_{\ell'}(c \vee d)$, and $c' = \lambda_\ell(c \vee d)$. Calculating e, we have

$$(e_i)_{i \in S} = \ell' \cdots \ell',$$
$$(e_i)_{i \notin S} = (c_i)_{i-\delta \notin S}.$$

Since e contains no letters greater than ℓ', we have $d' = \psi_{\ell'}(e) = \mu_{\ell'}(e)$. Thus, $d' = d$:

$$(d'_i)_{i \in S} = \ell' \cdots \ell',$$
$$(d'_i)_{i \notin S} = 0 \cdots 0.$$

Calculating $c' = \lambda_\ell(e)$, we change each occurrence of ℓ' in e to ℓ, and move it δ positions to the right. Since $\psi_\ell(c) <_Q d$, we see that $c' = c$:

$$(c'_i)_{i-\delta \in S} = \ell \cdots \ell = (c_i)_{i-\delta \in S},$$
$$(c'_i)_{i-\delta \notin S} = (c_i)_{i-\delta \notin S}.$$

(2) Similar. □

Completing the proof of Proposition 3.3, the following proposition shows that Ψ is bijective.

Proposition 5.3. *Let $c \in C(w)$ be a code on the letters $\ell_1 < \cdots < \ell_k$, and let $v_1 <_Q \cdots <_Q v_k$ be a k-element chain in Q, where the letter of v_i is ℓ_i for each i. The maps Ψ and Φ satisfy*

1. $\Psi\Phi(v_1 <_Q \cdots <_Q v_k) = v_1 <_Q \cdots <_Q v_k$.
2. $\Phi\Psi(c) = c$.

Proof. (1) By Definition 5.2, we have

$$\Phi(v_1 <_Q \cdots <_Q v_k) = (\cdots((v_1 \vee v_2) \vee v_3) \cdots) \vee v_k.$$

Applying $\Psi = \psi_{\ell_1} \times \cdots \times \psi_{\ell_k}$ to this code, we calculate $\psi_{\ell_i}((\cdots(v_1 \vee v_2) \vee \cdots) \vee v_k)$, for $i = 1, \ldots, k$. By Proposition 5.2 (1), we have

$$\begin{aligned}\psi_{\ell_i}((\cdots(v_1 \vee v_2) \vee \cdots) \vee v_k) &= \psi_{\ell_i}\lambda_{\ell_i}\lambda_{\ell_{i+1}} \cdots \lambda_{\ell_k}((\cdots(v_1 \vee v_2) \vee \cdots) \vee v_k) \\ &= \psi_{\ell_i}((\cdots(v_1 \vee v_2) \vee \cdots) \vee v_i) \\ &= v_i,\end{aligned}$$

as desired.

(2) By Definition 4.4, we have

$$\Psi(c) = \psi_{\ell_1}(c) <_Q \cdots <_Q \psi_{\ell_k}(c).$$

Applying Φ to this chain, we join vertices one at a time. Noting that $\psi_{\ell_1}(c) = \lambda_{\ell_1}(c)$, we use Proposition 5.2 (2) to calculate

$$\begin{aligned}\lambda_{\ell_i}(c) \vee \psi_{\ell_{i+1}}(c) &= \lambda_{\ell_i}\lambda_{\ell_{i+1}}(c) \vee \psi_{\ell_{i+1}}\lambda_{\ell_{i+1}}(c) \\ &= \lambda_{\ell_i}(\lambda_{\ell_{i+1}}(c)) \vee \psi_{\ell_{i+1}}(\lambda_{\ell_{i+1}}(c)) \\ &= \lambda_{\ell_{i+1}}(c).\end{aligned}$$

Thus, after $k-1$ join iterations, we recover c. □

6. Open questions

Since the class of balanced Cohen-Macaulay complexes contains so many widely studied classes of complexes, there are many possibilities to refine Theorem 3.1. In Theorem 3.2, we have required that Σ be an order complex of the form $\Delta(J(P))$, where P is a disjoint sum of chains. One could also ask if the theorem holds for more general classes of posets. (See [10], [11] for definitions in the questions that follow.) For instance, the following questions are open.

Question 6.1. If P is a forest, then is there another poset Q such the h-vector of $J(P)$ is the f-vector of Q?

Question 6.2. If P is a series-parallel poset, then is there another poset Q such the h-vector of $J(P)$ is the f-vector of Q?

We conjecture that the answers to both questions are affirmative. In fact, we conjecture that the answer remains affirmative for *any* choice of a poset P.

Conjecture 6.1. *Let $J(P)$ be any distributive lattice. Then there is another poset Q such that the h-vector of $J(P)$ is the f-vector of Q.*

This conjecture has been tested by computer for all distributive lattices $J(P)$ arising from posets P having up to seven elements.

Other open questions which define refinements of Theorem 3.1 allow the complexes Σ and Γ to be more general than in the conjecture above.

Question 6.3. If Σ is the order complex of a modular (or semimodular) lattice, then is there a *flag* complex Γ such that the h-vector of Σ is the f-vector of Γ?

Alternatively, one could place requirements on Γ instead of on Σ.

Question 6.4. For which balanced Cohen-Macaulay complexes Σ is h_Σ the f-vector of a *graded* poset (or $(\mathbf{3}+\mathbf{1})$-free poset, or flag complex)?

References

[1] M. M. Bayer and L. J. Billera, *Counting faces and chains in polytopes and posets*, Contemporary Mathematics, 34 (1984), pp. 207–252.

[2] L. J. Billera and A. Björner, *Face numbers of polyopes and complexes*, in Handbook of Discrete and Computational Geometry, J. E. Goodman and J. O'Rourke, eds., CRC Press, Boca Raton/New York, 1997, pp. 291–310.

[3] A. Björner, P. Frankl, and R. Stanley, *The number of faces of balanced Cohen-Macaulay complexes and a generalized Macaulay theorem*, Combinatorica, 7 (1987), pp. 23–34.

[4] A. Björner, A. Garsia, and R. Stanley, *An introduction to the theory of Cohen-Macaulay posets*, in Ordered Sets, I. Rival, ed., Reidel, Dordrecht/Boston/London, 1982, pp. 583–615.

[5] D. Dumont, *Interprétations combinatoires des nombres de Genocchi*, Duke Math. J., 41 (1974), pp. 305–318.

[6] P. Edelman and V. Reiner, *H-shellings and h-complexes*, Adv. Math., 106 (1994), pp. 36–64.

[7] D. Foata, *Distributions Eulériennes et Mahoniennes sur le groupe des permutations*, in Higher Combinatorics, M. Aigner, ed., vol. 19, D. Reidel Publishing Combany, Dordrecht-Holland, 1977, pp. 27–49.

[8] P. MacMahon, *Combinatory Analysis*, vol. 1 and 2, Cambridge Univ. Press, Cambridge, 1915.

[9] R. Stanley, *Balanced Cohen-Macaulay complexes*, Trans. of the AMS, 241 (1979), pp. 139–157.

[10] ———, *Enumerative Combinatorics*, vol. 1, Wadsworth & Brooks/Cole, Belmont, CA, 1986.

[11] ———, *Combinatorics and Commutative Algebra*, Birkhäuser, Boston, MA, 1996.