Words avoiding a reflexive acyclic relation

John Dollhopf

The Hill School, Pottstown, PA 19464

jdollhopf@thehill.org

Ian Goulden

Department of Combinatorics and Optimization, University of Waterloo

Waterloo, Ontario CANADA N2L 3G1

ipgoulde@math.uwaterloo.ca

Curtis Greene

Department of Mathematics, Haverford College

Haverford, PA 19041

cgreene@haverford.edu

Dedicated to Richard Stanley on his sixtieth birthday

Submitted: Dec 24, 2004; Accepted: Jan 30, 2006; Published: Feb 8, 2006

Mathematics Subject Classifications: 05A15, 06A07

Abstract

Let $\mathcal{A} \subseteq [\mathbf{n}] \times [\mathbf{n}]$ be a set of pairs containing the diagonal $\mathcal{D} = \{(i,i) \mid i = 1, \ldots, n\}$, and such that $a \leq b$ for all $(a,b) \in \mathcal{A}$. We study formulae for the generating series $F_{\mathcal{A}}(\mathbf{x}) = \sum_w \mathbf{x}^w$ where the sum is over all words $w \in [\mathbf{n}]^*$ that *avoid* $\mathcal{A}$, i.e., $(w_i, w_{i+1}) \notin \mathcal{A}$ for $i = 1, \ldots, |w| - 1$. This series is a rational function, with denominator of the form $1 - \sum_T \mu_{\mathcal{A}}(T)\mathbf{x}^T$, where the sum is over all nonempty subsets T of $[n]$. Our principal focus is the case where the relation $\mathcal{A}$ is *μ-positive*, i.e., $\mu_{\mathcal{A}}(T) \geq 0$ for all $T \subseteq [\mathbf{n}]$, in which case the form of the generating function suggests a cancellation-free combinatorial encoding of words avoiding $\mathcal{A}$. We supply such an interpretation for several classes of examples, including the interesting class of cycle-free (or crown-free) posets.

1 Introduction

Let $\mathbf{X}$ be a finite set, and let $\mathcal{A} \subseteq \mathbf{X} \times \mathbf{X}$ be a relation on $\mathbf{X}$. We consider the set $\mathcal{L}(\mathcal{A})$ of words whose letters are elements of $\mathbf{X}$, and whose adjacent letters *avoid* the pairs in $\mathcal{A}$. In other words, if $w = w_1 w_2 \cdots w_m$, then $w \in \mathcal{L}(\mathcal{A})$ if and only if $(w_i, w_{i+1}) \notin \mathcal{A}$

for $i = 1, \ldots, m-1$. In this paper we will consider the problem of enumerating words in $\mathcal{L}(\mathcal{A})$ in the special case when $\mathcal{A}$ is a *reflexive* and *acyclic* relation on $\mathbf{X}$. A typical example is the following: let $\mathbf{X} = \{1,2,3\}$ and $\mathcal{A} = \{(1,1),(2,2),(3,3),(1,3)\}$ so that $\mathcal{L}(\mathcal{A})$ is the set of words in $\{1,2,3\}$ avoiding repeated letters and also avoiding the pair $(1,3)$. If $w = w_1 w_2 \cdots w_m$ is a word in $\mathcal{L}(\mathcal{A})$, let $\mathbf{x}^w$ denote the monomial $x_{w_1} x_{w_2} \ldots x_{w_m}$ where the x_i are commuting indeterminates. Then

$$F_{\mathcal{A}}(x_1, x_2, x_3) = \sum_{w \in \mathcal{L}(\mathcal{A})} \mathbf{x}^w = \frac{(1+x_1)(1+x_2)(1+x_3)}{1 - x_1x_2 - x_2x_3 - x_1x_2x_3} \tag{1}$$

and

$$F_{\mathcal{A}}(t,t,t) = \frac{(1+t)^3}{1-2t^2-t^3} \;=\; \frac{(1+t)^2}{1-t-t^2} \tag{2}$$

It follows from (2) that, for this particular relation $\mathcal{A}$, the number $f(k)$ of words in $\mathcal{L}(\mathcal{A})$ of length k is a Fibonacci number.

For any relation $\mathcal{A}$, we will call the series $F_{\mathcal{A}}(x_1, x_2, \ldots)$ defined as in (1) the *pair-avoiding series* for $\mathcal{A}$. Our intent is not to add to the huge literature devoted to techniques for enumerating words avoiding various patterns (e.g., [4],[9], [11], [12]). Rather, we are interested in the combinatorics suggested by the *form* of equations such as (1) and (2). In particular, when the geometric series in (1) is expanded, the resulting series has positive terms. We will identify several large classes of examples for which this phenomenon occurs. Our goal will be to give combinatorial interpretations of all relevant coefficients and bijective correspondences, wherever formulas such as (1) suggest that these exist.

If $\mathcal{A}$ is a reflexive, acyclic relation on $\mathbf{X}$, and $|\mathbf{X}| = n$, we may label the elements of $\mathbf{X}$ with the elements of $[\mathbf{n}] = \{1, 2, \ldots, n\}$ so that $(a,b) \in \mathcal{A}$ implies $a \leq b$. Such a relation $\mathcal{A} \subseteq [\mathbf{n}] \times [\mathbf{n}]$ will be called *monotone*. In Section 2 we prove that for monotone relations $\mathcal{A}$,

$$F_{\mathcal{A}}(\mathbf{x}) \;=\; \frac{\prod_{i=1}^{n}(1+x_i)}{1 - \sum_{\substack{T \subseteq [\mathbf{n}] \\ T \neq \emptyset}} \mu_{\mathcal{A}}(T)\mathbf{x}^T} \tag{3}$$

where $\mu_{\mathcal{A}}(T) \in \mathbf{Z}$, and $x^T = \prod_{w \in T} x_w$. It is natural to consider the case when $\mu(T) \geq 0$ for all $T \subseteq [\mathbf{n}]$, and we will say that a relation $\mathcal{A}$ is *μ-positive* if it has this property. For such relations, we have a positive expansion

$$F_{\mathcal{A}}(\mathbf{x}) \;=\; \sum_{T_0, T_1, T_2, \ldots} x^{T_0} \prod_i \mu_{\mathcal{A}}(T_i)\; x^{T_i}\;. \tag{4}$$

where the sum is over all finite sequences of subsets of $[\mathbf{n}]$ with $T_i \neq \emptyset$ for $i > 1$. The form of (4) suggests that for μ-positive $\mathcal{A}$ the nonnegative integer $\mu_{\mathcal{A}}(T)$ should have a direct combinatorial meaning, and that there should exist an encoding of words in $\mathcal{L}(\mathcal{A})$ by sequences of sets with positive weights determined by the values of $\mu_{\mathcal{A}}$. We will solve this problem for two large classes of μ-positive monotone relations on $[\mathbf{n}]$, namely the following.

V-free relations. These are defined by the condition

$$x\mathcal{A}y \text{ and } x\mathcal{A}z \text{ imply } y\mathcal{A}z \text{ or } z\mathcal{A}y. \tag{5}$$

This class includes all examples whose incidence matrices are *column-convex*, i.e. $x\mathcal{A}z$ implies $y\mathcal{A}z$ for all y such that $x < y < z$. For V-free relations $\mathcal{A}$, $\mu_{\mathcal{A}}(T)$, it turns out that $\mu_{\mathcal{A}}(T)$ can be interpreted as the number of winning positions (minus 1) in a simple combinatorial game associated with $\mathcal{A}$. An analogous theory exists for monotone relations that are Λ-free, where this is defined by a condition dual to (5).

Cycle-free (=crown-free) posets. These are posets P with a natural labeling whose underlying comparability graph contains no chordless cycles of length > 3. Equivalently, P contains no induced subposet order-isomorphic to a crown (see section 4 for precise definitions). We will show that, for *any* poset (cycle-free or not), $\mu_{\mathcal{A}}(T)$ is equal to the value of the Möbius function $\mu_{\mathcal{A}}(\hat{0}, \hat{1})$, computed in the poset obtained by adding a $\hat{0}$ and $\hat{1}$ to T. We show, further, that a poset P is μ-positive if and only if it is cycle-free, and in this case, $\mu_{\mathcal{A}}(T)$ is equal to the number of connected components of T minus 1.

For both of the above examples, we will give complete encoding and decoding algorithms corresponding to the positive expansion formula (4).

This paper is organized as follows. Section 2 reviews some enumerative techniques needed to derive formula (3) for arbitrary monotone relations $\mathcal{A}$. In Section 3 we present a basic paradigm for encoding and decoding of words, in the spirit of (4), and in Section 4 we prove that these algorithms are valid in their simplest form if and only the relation $\mathcal{A}$ is V-free. Section 5 develops some general theory relevant to the poset case, and proves that posets are μ-positive if and only if they are cycle-free. Section 6 shows how to adapt the algorithms in Section 3 to the cycle-free poset case. Section 7 explores several interesting special cases in more detail, and Section 8 discusses analogues of our main results for rearrangements of a multiset.

2 The pair-avoiding series for monotone relations

We begin by developing some general techniques for computing the pair-avoiding series $F_{\mathcal{A}}(\mathbf{x})$. We will rely on a simple determinant formula from [11] (Chapter 4) which is well-suited to our problem. This material is well known, but we have included a complete derivation to make our treatment self-contained.

Theorem 2.1 *Let $\mathcal{A} \subseteq [\mathbf{n}] \times [\mathbf{n}]$ be an arbitrary relation. Let $\overline{\mathcal{A}} = [\mathbf{n}] \times [\mathbf{n}] - \mathcal{A}$ denote the relation complementary to $\mathcal{A}$. Then*

$$F_{\mathcal{A}}(\mathbf{x}) \;=\; \frac{\det(I + XA)}{\det(I - X\bar{A})}, \tag{6}$$

where A and $\bar{A}$ denote the $n \times n$ incidence matrices of $\mathcal{A}$ and $\overline{\mathcal{A}}$, respectively, and $X = \operatorname{diag}(x_1, \ldots, x_n)$.

Proof. Let J denote the $n \times n$ matrix all of whose entries equal 1, so that $\bar{A} = J - A$. The (i, j)-entry of $(X\bar{A})^{k-1}X$ gives the contribution to $F_{\mathcal{A}}(\mathbf{x})$ from words of length k that begin with i and end with j, so

$$\begin{aligned} F_{\mathcal{A}}(\mathbf{x}) &= 1 + \sum_{i,j=1}^{n} \sum_{k \geq 1} [(X\bar{A})^{k-1}X]_{ij} \\ &= 1 + \sum_{i,j=1}^{n} [(I - X\bar{A})^{-1}X]_{ij} \\ &= 1 + \operatorname{trace}(I - X\bar{A})^{-1}XJ \end{aligned} \tag{7}$$

since for all $n \times n$ matrices K,

$$\sum_{i,j=1}^{n} [K]_{ij} = \operatorname{trace} KJ.$$

Now

$$\det(I + M) = 1 + \operatorname{trace} M$$

for any matrix M with $\operatorname{rank}(M) = 1$, so, from (7),

$$\begin{aligned} F_{\mathcal{A}}(\mathbf{x}) &= \det(I + (I - X\bar{A})^{-1}XJ) \\ &= \frac{\det(I + XA)}{\det(I - X\bar{A})} \end{aligned} \tag{8}$$

as claimed.

Theorem 2.1 is a special case of a more general result in [11] (Section 2.8), which uses inclusion-exclusion to give the generating series for words avoiding an arbitrary collection of subwords, not restricted, as in this paper, to words of length 2. Note also that, if the numerator of (6) is replaced by 1, one obtains the generating function appearing in MacMahon's "Master Theorem" ([20]; see also [5]) which enumerates restricted permutations (also called rearrangements) of a multiset. This is not a coincidence; indeed, there is a precise connection between our results and the general theory of multiset rearrangements, and we will explain this connection in Section 8.

As a corollary of Theorem 2.1 we obtain a useful reciprocity theorem for pair-avoiding series, due independently to Gessel [9] and Carlitz, Scoville, and Vaughan [4].

Corollary 2.2 *Let $\mathcal{A} \subseteq [\mathbf{n}] \times [\mathbf{n}]$ be an arbitrary relation. Then*

$$F_{\mathcal{A}}(\mathbf{x}) = \frac{1}{F_{\overline{\mathcal{A}}}(-\mathbf{x})}$$

Theorem 2.1 gives our main formula (3) as an immediate corollary, and also yields several different explicit expressions for $\mu_{\mathcal{A}}(T)$, as we show next. Let $M[T,T]$ denote the principal submatrix obtained by restricting the matrix M to the rows and columns determined by $T \subseteq [\mathbf{n}]$.

Corollary 2.3 *If $\mathcal{A}$ is a monotone relation on $[\mathbf{n}]$ then*

$$F_{\mathcal{A}}(\mathbf{x}) \;=\; \frac{\prod_{i=1}^{n}(1+x_i)}{1-\sum_{\substack{T\subseteq[\mathbf{n}]\\ T\neq\emptyset}} \mu_{\mathcal{A}}(T)\mathbf{x}^T} \tag{9}$$

where $\mu_{\mathcal{A}}(T)\in\mathbf{Z}$ for all $T\subseteq[\mathbf{n}], T\neq\emptyset$. Moreover,

$$\mu_{\mathcal{A}}(T) \;=\; -\det{(A-J)[T,T]} \;=\; \sum_{\substack{S\subseteq T\\ S\in\mathcal{L}(\overline{\mathcal{A}}_0)}} (-1)^{|S|-1} \tag{10}$$

where $\mathcal{A}_0=\mathcal{A}-\mathcal{D}$. Here, $\mathcal{D}$ denotes the diagonal relation on $[\mathbf{n}]$, and we identify a set S with the word obtained by writing its elements in increasing order.

Proof. From (6) we have

$$\begin{aligned}
F_{\mathcal{A}}(\mathbf{x}) &= \frac{\det{(I+XA)}}{\det{(I+X(A-J))}}\\
&= \frac{\det{(I+XA)}}{\sum_{T\subseteq[\mathbf{n}]}\det{(X(A-J)[T,T])}}\\
&= \frac{\det{(I+XA)}}{\sum_{T\subseteq[\mathbf{n}]}\mathbf{x}^T\det{(A-J)[T,T]}}.
\end{aligned}$$

But A is upper unitriangular, so

$$\det{(I+XA)}=\prod_{i=1}^{n}(1+x_i),$$

and expression (9) follows, with the determinantal form of (10).

For the second part of (10), note that the elements of $\mathcal{L}(\mathcal{A}_0)$ can be constructed uniquely by applying the following combinatorial construction to the elements of $\mathcal{L}(\mathcal{A})$: for each word $w\in\mathcal{L}(\mathcal{A})$, replace every symbol a by an arbitrary nonempty string of a's, in all possible ways. This gives us immediately the equation

$$F_{\mathcal{A}_0}(\mathbf{x})=F_{\mathcal{A}}(\frac{x_1}{1-x_1},\ldots,\frac{x_n}{1-x_n}),$$

and so replacement of x_i by $\frac{x_i}{1+x_i}, i=1,\ldots,n$ gives

$$F_{\mathcal{A}}(\mathbf{x})=F_{\mathcal{A}_0}(\frac{x_1}{1+x_1},\ldots,\frac{x_n}{1+x_n}). \tag{11}$$

Now, if $\overline{\mathcal{A}}_0$ denotes the complement of $\mathcal{A}_0$ then all elements of $\mathcal{L}(\overline{\mathcal{A}}_0)$ are strictly increasing words. We will identify such words S with the underlying set of symbols, and write $i\in S$ to indicate that the symbol i appears in the increasing word S. Then directly from the definition of the series F we get

$$F_{\overline{\mathcal{A}}_0}(\mathbf{x}) \;=\; \sum_{S\in\mathcal{L}(\overline{\mathcal{A}}_0)}\prod_{i\in S}x_i.$$

Applying Corollary 2.2 to this last expression, substituting in (11) and multiplying in the numerator and denominator by $\prod_{i=1}^{n}(1+x_i)$ yields

$$F_{\mathcal{A}}(\mathbf{x}) \;=\; \frac{\prod_{i=1}^{n}(1+x_i)}{\sum_{S\in\mathcal{L}(\overline{\mathcal{A}}_0)}\prod_{i\in S}(-x_i)\prod_{i\notin S}(1+x_i)}. \tag{12}$$

The constant term in the denominator is 1, arising from the empty word. This gives a second proof of expression (9), with the alternating sum form for (10).

Example 2.4 Let $n=3$ and $\mathcal{A}=\{(1,1),(1,2),(2,2),(2,3),(3,3)\}$. Then

$$A \;=\; \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

and so

$$F_{\mathcal{A}}(\mathbf{x}) \;=\; \frac{\begin{vmatrix} 1+x_1 & x_1 & 0 \\ 0 & 1+x_2 & x_2 \\ 0 & 0 & 1+x_3 \end{vmatrix}}{\begin{vmatrix} 1 & 0 & -x_1 \\ -x_2 & 1 & 0 \\ -x_3 & -x_3 & 1 \end{vmatrix}} \;=\; \frac{(1+x_1)(1+x_2)(1+x_3)}{1-x_1x_3-x_1x_2x_3}.$$

Example 2.5 Let $n=4$ and let $\mathcal{A}$ be the monotone relation defined by the incidence matrix

$$A \;=\; \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

then

$$F_{\mathcal{A}}(\mathbf{x}) \;=\; \frac{(1+x_1)(1+x_2)(1+x_3)(1+x_4)}{1-x_1x_2-x_3x_4+x_1x_2x_3x_4}, \tag{13}$$

In this example we have $\mu_{\mathcal{A}}(\{1,2,3,4\}) = -1$, showing that, in general, the integer coefficient $\mu_{\mathcal{A}}(T)$ appearing in (9) can assume both positive and negative values.

The presence of negative coefficients in (13) may be expected, since the combinatorial components of (9) and (10) involve cancellation in a variety of ways. It is thus surprising to discover the phenomenon of μ-positivity in several large families of relations. We devote the remainder of this paper to studying the combinatorial significance of μ_A is these cases.

3 A simple decoding model

Suppose that $\mathcal{A}$ is a μ-positive monotone relation. Following (4), we seek a bijection between words $w \in \mathcal{L}(\mathcal{A})$ and certain sequences

$$(T_0, (T_1, x_1), (T_2, x_2), \ldots, (T_k, x_k)) \tag{14}$$

where the T_i are subsets of $[\mathbf{n}]$ (with $T_i \neq \emptyset, i > 0$) and x_i represents a "marking" of T_i for $i = 1, \ldots, k$. Here T_0 denotes an arbitrary (possibly empty) subset. It should be noted that the symbols x_i appearing in (14) are distinct from (and have nothing to do with) the indeterminates x_i appearing in, e.g., (13). This bijection should be *weight-preserving*, in the sense that it preserves the underlying multisets of letters. In order for such a bijection to exist, the number of possible markings of T_i must equal $\mu_{\mathcal{A}}(T_i)$ for each i, by (4). We refer to (14) as a *coding sequence* for w, and the individual terms (T_i, x_i) as *codons*.

If $\mu_{\mathcal{A}}(T) \leq |T|$ for all $T \subseteq [\mathbf{n}]$, then one potential marking scheme for T consists of choosing elements x from a special set $M(T) \subseteq T$ of "markable" elements, where $|M(T)| = \mu_{\mathcal{A}}(T)$ for each T. Assuming that all codons have this form, let us consider the following simple algorithm for decoding sequences $(T_0, (T_1, x_1), (T_2, x_2), \ldots, (T_k, x_k))$ into words $w \in \mathcal{L}(\mathcal{A})$:

Algorithm A (Basic Decoding Algorithm):

(1) *Initially let* $w = [T_0]_\downarrow$. *(Here* $[T]_\downarrow$ *denotes the string obtained from* T *by writing its elements in decreasing order.)*

(2) *If* w *has been defined, adjoin* (T, x) *to* w *by the rule*

$$w(T, x) \longmapsto \begin{cases} wx[T - x]_\downarrow & \text{if the result is in } \mathcal{L}(\mathcal{A}) \\ w[T]_\downarrow & \text{otherwise} \end{cases}$$

(3) *Repeat* (2) *with* $(T_1, x_1), (T_2, x_2), \ldots, (T_k, x_k)$, *until all codons have been adjoined.*

In order for Algorithm A to work (i.e. to be well defined and give a bijection) certain conditions on both the relation $\mathcal{A}$ and the sets $M(T)$ must be satisfied. The precise requirements are contained in the following theorem.

Theorem 3.1 *Suppose that* $\mathcal{A}$ *is a* μ*-positive monotone relation on* $[\mathbf{n}]$*, and that for each nonempty subset* $T \subseteq [\mathbf{n}]$ *a subset* $M(T) \subseteq T$ *has been assigned, such that* $|M(T)| = \mu(T)$ *for all* T*. Suppose further that Algorithm A defines a weight-preserving bijection between words* $w \in \mathcal{L}(\mathcal{A})$ *and coding strings* $(T_0, (T_1, x_1), (T_2, x_2), \ldots, (T_k, x_k))$*. Then*

(1) $\mathcal{A}$ *is* V*-free, and*

(2) *For each* T, $M(T) = W(T) - \{\max(T)\}$*, where* $W(T)$ *is the set of winning positions in the combinatorial game* $\mathbf{NIM}_{\mathcal{A}}(T)$*, whose rules are given below.*

Conversely, if $\mathcal{A}$ *and* $M(T), T \subseteq [\mathbf{n}]$ *satisfy conditions* (1) *and* (2)*, then* $\mathcal{A}$ *is* μ*-positive and Algorithm A is bijective.*

Definition of the game $\mathbf{NIM}_{\mathcal{A}}(T)$: *A chip or stone is placed on an element of T, and two players take turns moving it. Legal moves from element $i \in T$ are to any $j \in T$ for which $i \neq j$ and $i\mathcal{A}j$. A player loses if no legal moves are possible. A winning position is one from which an eventual win can be forced by the player who moves to it.*

Informally, k is winning if there are no legal responses in T, or if every legal response is losing. This rule suffices to define unique sets of winning and losing positions for any finite acyclic directed graph. For example, these sets can be defined iteratively as follows: the sinks are winning, and predecessors of sinks are losing; remove these vertices and all incident edges and repeat, until no vertices remain. This type of game has been studied, for example, in Chapter 14 of [2], where the set of winning positions in a graph G is called the *kernel* of G.

The proof of Theorem 3.1 is contained in the next section, along with some other lemmas and remarks about V-free relations. We conclude this section with several examples to illustrate the algorithm.

Example 3.2 For the Fibonacci example in Section 1, we have $W(\{123\}) = \{23\}$, $W(\{12\}) = \{12\}$, and $W(\{23\}) = \{23\}$, so that $M(\{123\}) = \{2\}$, $M(\{12\}) = \{1\}$, and $M(\{23\}) = \{2\}$. Furthermore, $M(T) = \emptyset$ for all other nonempty $T \subseteq \{123\}$. Hence the valid codons (T, x) are $(321, 2), (21, 1)$, and $(32, 2)$. Here, and in subsequent examples, we are denoting codons (T, x) by $([T]_\downarrow, x)$, i.e., with the elements of T written in decreasing order. As an exercise, the reader may check that, Algorithm A gives the decoding

$$(31)(32,2)(321,2)(21,1)(321,1) \longmapsto 312323121231$$

Example 3.3 Let $\mathcal{A} = \mathcal{D}$, the diagonal relation on $[\mathbf{n}] \times [\mathbf{n}]$. Then $\mathcal{L}(\mathcal{A})$ is the set of all words without repeated letters, sometimes called *Smirnov words* (after [23]; see also [11], p. 68). Then, for any $T \neq \emptyset$, all positions are winning and we have $M(T) = T - \{\max(T)\}$. For each nonempty T there are $|T| - 1$ codons (T, x). The reader may check that Algorithm A gives the decoding

$$(31)(32,2)(321,1)(21,1)(321,1) \longmapsto 312313212132$$

We note that, for this example, Algorithm A can be simplified to the rule: decode $w(T, x)$ as $wx[T - x]_\downarrow$ unless the last letter of w equals x, and in that case decode it as $w[T]_\downarrow$. We also note that, in this example, formula (9) becomes

$$F_{\mathcal{D}}(\mathbf{x}) = \frac{\sum_{i=0}^{n} e_i(\mathbf{x})}{1 - \sum_{i=2}^{n} (i-1) e_i(\mathbf{x})}, \tag{15}$$

where $e_i(\mathbf{x})$ is the i-th elementary symmetric function of $\mathbf{x}$. This generating function appears in [25], where it arises as the coloring polynomial of a path. It is also interesting to note that if the numerator in (15) is replaced by 1, one gets MacMahon's generating function for derangements of a multiset [20, Chapter 3]. See Section 8 for an explanation of the relationship between our formula and MacMahon's.

Example 3.4 Let $\mathcal{A} = \mathcal{D} \cup \{(i, i+1) \mid 1 \le i \le n-1\}$. Suppose T consists of k "blocks" of consecutive elements, with adjacent blocks separated by at least 2. For example, the decreasing word 9876321 consists of 2 blocks, namely 9876 and 321. The markable elements of T are the 3rd, 5th,... largest elements in the block containing the largest elements, and the 1st, 3rd, 5th,... largest elements in the other blocks. For example, the markable elements of 9876321 are 7,3,1. Thus if the blocks of T have lengths $b_1, \ldots, b_k$ then

$$|M(T)| \;=\; -1 + \sum_{i=1}^{k} \left\lceil \frac{b_i}{2} \right\rceil,$$

where $\lceil\ \rceil$ is the ceiling function. The reader may verify that Algorithm A gives the decoding

$$(2)(321,1)(5431,3)(6532,3)(7631,3) \longmapsto 2132543136527631$$

Example 3.5 Let $\mathcal{A}$ be the column-convex relation defined by the incidence matrix

$$A \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

The reader may check that Algorithm A gives the decoding

$$(32)(641,4)(431,1)(65432,4) \longmapsto 3264143146532$$

4 Proof of Theorem 3.1

We will first prove the necessity of conditions (1) and (2), assuming that a marking scheme $M(T), T \subseteq [\mathbf{n}]$ has been specified, and Algorithm A is well-defined and bijective. Our objective is to show that $\mathcal{A}$ is V-free, and that, for each nonempty T, $M(T) = W(T) - \{\max(T)\}$. This part of the proof will proceed by a series of short lemmas.

Lemma 4.1 *If $m = \max(T)$, then $m \notin M(T)$, i.e., m cannot be marked.*

Proof. Otherwise, the coding sequence $(m)(T, m)$ has no valid decoding, since neither $mm[T-m]_\downarrow$ nor $m[T]_\downarrow$ are valid words in $\mathcal{L}(\mathcal{A})$.

Lemma 4.2 *If $x\mathcal{A}y$, with $x, y \in T$, then x and y are not both in $M(T)$.*

Proof. Otherwise we have $(x)(T,x) \mapsto x[T]_\downarrow$ and $(x)(T,y) \mapsto x[T]_\downarrow$, and Algorithm A is not injective.

Lemma 4.3 *If $x \neq m$ is an element with no successors in T, then $x \in M(T)$. More generally, if $x \neq m$ is an element whose only successors in T are not markable, then $x \in M(T)$.*

Proof. Otherwise the word $x[T-x]_{\downarrow} \in \mathcal{L}(\mathcal{A})$ cannot be obtained from Algorithm A, and Algorithm A is not surjective.

The preceding lemmas show that $x \neq m$ is markable if and only if none of its successors are markable, which proves that $M(T) = W(T) - \{\max(T)\}$, as desired. The next lemma shows that $\mathcal{A}$ must be V-free.

Lemma 4.4 *If $x\mathcal{A}y, x\mathcal{A}z$, and $y < z$, then $y\mathcal{A}z$.*

Proof. Otherwise the coding sequence $(x)(zy, y)$ has no proper decoding, and Algorithm A is not well defined.

This completes the first part of the proof of Theorem 3.1. To complete the proof, we must show that conditions (1) and (2) imply that $\mathcal{A}$ is μ-positive, and Algorithm A is both injective and surjective. Although it is not difficult to prove each of these last two statements independently, we will prove only that Algorithm A is surjective (which is somewhat easier), and then complete the proof by a counting argument. The following lemma will be helpful.

Lemma 4.5 *If $\mathcal{A}$ is V-free and $z \in T$, then there is a unique element $z^* \in W(T)$ such that $z\mathcal{A}z^*$. Equivalently, if a position z in $\mathbf{NIM}_{\mathcal{A}}(T)$ is not winning, then there is a unique winning response to it.*

Proof. If $z \notin W(T)$, there must be some winning response $y \in W(T)$, with $z\mathcal{A}y$. If there were another such response, say y', then the V-free condition implies either $y\mathcal{A}y'$ or $y'\mathcal{A}y$, implying that either $y \notin W(T)$ or $y' \notin W(T)$. This is a contradiction.

Lemma 4.6 *If conditions (1) and (2) of Theorem 3.1 hold, then Algorithm A is surjective.*

Proof. Suppose that $w \in \mathcal{L}(\mathcal{A})$. If the letters of w are decreasing, then clearly w is the image of (w). Otherwise, there is a unique factorization $w = w_0\,z\,u$ where u is a decreasing subword preceded by the letter z, which is an ascent. Let U denote the set of letters in u, and let $T = U \cup \{z\}$. Let z^* be the unique element in T specified by Lemma 4.5. Then the following rule shows how to obtain w by adjoining a codon to a shorter word in $\mathcal{L}(\mathcal{A})$.

- If $z = z^* \in W(T)$ and $z^* \notin U$, then $w_0(T, z^*) \mapsto w$.
- Otherwise, $w_0 z(U, z^*) \mapsto w$.

A straightforward induction argument now shows that Algorithm A is surjective, and the proof is complete.

The next lemma enables us to compute the number of valid coding sequences of a given length, and thus verify that Algorithm A is injective.

Lemma 4.7 *If $\mathcal{A}$ is V-free, then $\mathcal{A}$ is μ-positive, and*

$$\mu_{\mathcal{A}}(T) \;=\; |\mathcal{W}_{\mathcal{A}}(T)| - 1$$

for all $T \subseteq [\mathbf{n}]$.

Proof. For $T \subseteq [\mathbf{n}]$ and $a \in T$, let

$$c(a,T) = \sum_{\substack{S \subseteq T,\, \min(S)=a \\ S \in \mathcal{L}(\overline{\mathcal{A}}_0)}} (-1)^{|S|-1}. \tag{16}$$

Then by the alternating sum in (10), we have

$$\mu_{\mathcal{A}}(T) = -1 + \sum_{a \in T} c(a,T). \tag{17}$$

Further, it is an easy consequence of (16) that

$$c(a,T) \;=\; 1 \;-\; \sum_{\substack{a\mathcal{A}b \\ b \neq a,\, b \in T}} c(b,T), \tag{18}$$

Clearly, $c(m,T) = 1$ for $m = \max(T)$, and it is easy to show by induction, using (18) and Lemma 4.5, that

$$c(a,T) \;=\; \begin{cases} 1 & \text{if } a \in \mathcal{W}_{\mathcal{A}}(T) \\ 0 & \text{if } a \notin \mathcal{W}_{\mathcal{A}}(T) \end{cases}$$

The lemma now follow immediately.

Corollary 4.8 *For all $L \geq 0$, the number of words $w \in \mathcal{L}(\mathcal{A})$ of length L equals the number of coding sequences $(T_0, (T_1, x_1), \ldots, (T_k, x_k))$ with $\sum_{i=0,k} |T_i| = L$.*

Proof. This follows from Lemma 4.7 and (4).

Combining Lemma 4.6 and Corollary 4.8, we obtain that Algorithm A is bijective, and this completes the proof of Theorem 3.1.

Note that there is another decoding algorithm that is equally simple as Algorithm A – we could form the words in step (2) of the algorithm by adding $[T-x]_{\downarrow}x$ or $[T]_{\downarrow}$ on the *left* of w. In this case, we obtain a result similar to Theorem 3.1, where the V-freeness of $\mathcal{A}$ is replaced by Λ-freeness, and sources play the role of sinks in the definition of $\mathbf{NIM}_{\mathcal{A}}(T)$.

Finally, we note that the proof of Lemma 4.6 contains an encoding algorithm which inverts Algorithm A under the assumption that $\mathcal{A}$ is V-free. We state it separately, as follows:

Algorithm B (Basic Encoding Algorithm): *Assume that $\mathcal{A}$ is V-free and $w \in \mathcal{L}(\mathcal{A})$.*

(1) *If w decreasing, output $([w]_{\downarrow})$ and terminate; otherwise write $w = w_0 z u$ where z is an ascent and u is decreasing. Denote the set of letters in u by U.*

(2) *If $z \notin U$, delete zu from w and output $(U \cup \{z\}, z)$ if this is a valid codon; otherwise delete u from w and output (U, z^*).*

(3) *Repeat the process, scanning w and building a coding sequence from right to left, until w is empty.*

5 Transitive relations (posets)

In this section we consider the case when $\mathcal{A}$ is both monotone and transitive, i.e. it defines a poset, with the integers $\{1, \ldots, n\}$ forming a natural labelling. For a poset P, let $\mu(P)$ denote $\mu_{\hat{P}}(\hat{0}, \hat{1})$, where $\hat{P}$ is the poset obtained from P by adjoining a $\hat{0}$ and $\hat{1}$, and $\mu_{\hat{P}}$ denotes the Möbius function of $\hat{P}$ (see [24] for definitions and basic properties of Möbius functions). We begin by observing that the coefficient $\mu_{\mathcal{A}}(T)$ appearing in (9) can be interpreted as a Möbius function.

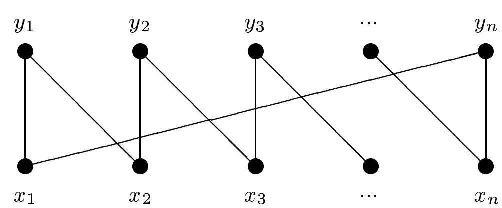
Lemma 5.1 *Suppose that $\mathcal{A} \subseteq [\mathbf{n}] \times [\mathbf{n}]$ is a transitive monotone relation. Then for all $T \subseteq [\mathbf{n}]$,*

$$\mu_{\mathcal{A}}(T) \;=\; \mu(\hat{T}) \tag{19}$$

Here we regard T as a subposet of $[\mathbf{n}]$, obtained by restriction of the relation $\mathcal{A}$.

Proof. This follows directly from the alternating sum in (10) and the well-known formula computing μ as an alternating sum over chains (see, for example, [24], Proposition 3.4.5). Note that each $S \in \mathcal{L}(\overline{\mathcal{A}}_0)$ in the alternating sum in (10) corresponds to a chain in the poset corresponding to T.

In view of (19), we can identify transitive $\mu_{\mathcal{A}}$-positive monotone relations by determining for which posets P it is true that $\mu(Q) \geq 0$ for all nonempty induced subposets $Q \subseteq P$. A poset P with this property will be called *μ-positive.* We will show that such posets can be completely characterized. First a technical definition: call P *cycle-free* if its underlying comparability graph is a *chordal* or *triangulated* graph, i.e., it contains no chordless cycles of length greater than three. Equivalently, a poset is cycle-free if it has no induced subposets of the form



with $n \geq 2$. The excluded posets are called *crowns.* Cycle-free posets have been studied by several authors, e.g., [16], [18], and [26], Chapter 5; see also [10] for a general exposition of triangulated graphs.

Theorem 5.2 *A poset P is μ-positive if and only if it is cycle-free.*

Proof. The condition is clearly necessary, since adjoining a $\hat{0}$ and $\hat{1}$ to a crown gives a poset with $\mu(\hat{0}, \hat{1}) = -1$. Suppose, on the other hand, that P contains no crowns, and that Q is μ-positive for all proper subposets $Q \subseteq P$. We will show that $\mu_{\hat{P}}(\hat{0}, \hat{1}) \geq 0$.

Consider first the case where P contains a chain of length greater than two. Then there is an element $x \in P$ such that the intervals $[\hat{0}, x]$ and $[x, \hat{1}]$ each have more than two elements. These intervals are both of the form $\hat{Q}$ for nonempty proper subsets $Q \subseteq P$. Hence we have $\mu_{\hat{P}}(\hat{0}, x) \geq 0$ and $\mu_{\hat{P}}(x, \hat{1}) \geq 0$. A simple but useful identity (due originally to Baclawski; see also [24], Lemma 3.14.3) states that for any poset $\hat{P}$ with a $\hat{0}$ and $\hat{1}$, and any $x \in \hat{P} - \{\hat{0}, \hat{1}\}$,

$$\mu_{\hat{P}}(\hat{0}, \hat{1}) \;=\; \mu_{\hat{P}-x}(\hat{0}, \hat{1}) + \mu_{\hat{P}}(\hat{0}, x)\mu_{\hat{P}}(x, \hat{1}) \tag{20}$$

Since $\mu_{\hat{P}-x}(\hat{0}, \hat{1}) \geq 0$ by assumption, the result follows in this case.

Next suppose that every chain in P has length at most two, i.e. the Hasse diagram of P is a bipartite graph. By assumption, this graph is also a forest. If P consists entirely of isolated points, the result is trivial. Otherwise the bipartite graph has an "endpoint" x, which we may assume to be a maximal element. Now $\mu_{\hat{P}}(\hat{0}, x) = 0$, so using (20) again we have

$$\mu_{\hat{P}}(\hat{0}, \hat{1}) \;=\; \mu_{\hat{P}-x}(\hat{0}, \hat{1}) + \mu_{\hat{P}}(\hat{0}, x)\mu_{\hat{P}}(x, \hat{1}) \;=\; \mu_{\hat{P}-x}(\hat{0}, \hat{1}) \geq 0$$

and the proof is complete.

It is worthwhile to have a good constructive characterization of cycle-free posets. For this we need some more machinery.

Definition 5.3 *Let P be a cycle-free poset. Denote by $\mathrm{Top}(P)$ the set of $x \in P$ such that the elements above x in P are linearly ordered. Similarly, let $\mathrm{Bot}(P)$ denote the set of $x \in P$ such that the elements below x are linearly ordered.*

Lemma 5.4 *If P is a cycle-free poset, then*

$$\mathrm{Top}(P) \cup \mathrm{Bot}(P) \;=\; P$$

PROOF. If an element x lies above two incomparable elements and below two others, those four elements form a chordless 4-cycle.

Lemma 5.5 *If P is a cycle-free poset, then*

$$\mathrm{Top}(P) \cap \mathrm{Bot}(P) \;\neq\; \emptyset$$

In other words, there exists an element $x \in P$ such that the set of all y comparable to x forms a chain.

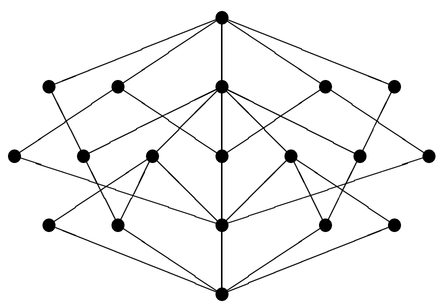
PROOF. An old result, due independently to Dirac (1961) and also Lekkerkerker and Boland (1962), states that any triangulated graph contains a *simplicial vertex*, i.e, a vertex whose neighbors form a clique. (See [10] for a proof.) Our lemma simply restates this result in the language of partially ordered sets.

Corollary 5.6 *Let P be a cycle-free poset. Then P may be constructed by successively adjoining a sequence of elements $x_1, x_2, \ldots, x_{|P|}$ such that for each i, the set of elements comparable to x_i in $\{x_1, x_2, \ldots, x_i\}$ forms a chain.*

Corollary 5.6 translates into a constructive (recursive) algorithm, as follows. To obtain a cycle-free poset of size $n+1$, start with a cycle-free poset P of size n, and then either

- pick an element $y \in \mathrm{Bot}(P)$ and add an element x covering y,
- pick an element $z \in \mathrm{Top}(P)$, and add an element x covered by z,
- pick elements $y \in \mathrm{Bot}(P)$ and $z \in \mathrm{Top}(P)$ with $y < z$, and add an element x covering y and covered by z,
- add a single isolated element x.

The following diagram illustrates a cycle-free poset. The reader may wish to verify that it can be constructed by the algorithm just described.



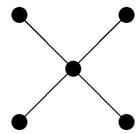
In the algorithm presented above, one can enlarge the class of allowable operations somewhat, for example:

- add an element x which is a new top (or bottom) to P.

This operation preserves the cycle-free property, but the element added is not necessarily simplicial.

The above results show that cycle-free posets are *dismantlable* in the sense of [15]. This means that one can reduce P to the empty poset by removing a succession of doubly irreducible elements. (An element of a poset P is *doubly irreducible* if it covers and is covered by at most one element of P.) However, the existence of a reduction via simplicial elements (as above) is a special property of cycle-free posets: every simplicial element is doubly irreducible, but not conversely. It is easy to construct examples of dismantlable posets that are not cycle-free.

The class of cycle-free posets includes a large number of familiar families whose Hasse diagrams are "cycle-free" in the usual graph-theoretic sense, for example: top-rooted forests, bottom-rooted forests, bipartite forests. However, the two conditions are not equivalent: for example, the following poset has no cycles in its Hasse diagram, but is not cycle-free.



The next lemma gives a combinatorial interpretation for $\mu(Q)$ when $Q \subseteq P$ and P is μ-positive.

Lemma 5.7 *Suppose that P is μ-positive (or, equivalently, cycle-free). Then*

$$\mu(Q) = \text{(number of connected components of } Q\text{)} - 1,$$

for all subsets $Q \subseteq P$.

Proof. We prove this for $Q = P$ by computing $\mu_{\hat{P}}(\hat{0}, \hat{1})$ recursively, from bottom to top, using the rule

$$\mu(\hat{0}, x) \;=\; -\sum_{y<x} \mu(\hat{0}, y)$$

In this computation, if $x \in \mathrm{Bot}(P)$ but x is not a root, i.e., a minimal element of P, then the value of $\mu(\hat{0}, x)$ is zero. Hence such elements may be deleted from P without changing $\mu_{\hat{P}}(\hat{0}, \hat{1})$. Similarly the non-root elements of $\mathrm{Top}(P)$ may be deleted, and hence we may assume that P is bipartite. Since there are no cycles, there is at least one endpoint $x \in P$, i.e., an element x related to exactly one other $y \in P$. Such may also be deleted without changing $\mu_{\hat{P}}(\hat{0}, \hat{1})$. This process may be repeated until P consists entirely of isolated points, and in this case the result is immediate.

6 Encoding/decoding for cycle-free posets

If P is a top-rooted forest, then P is V-free, and hence the methods of Section 3 apply directly, giving an encoding algorithm for this case. For more general cycle-free posets, the situation is more complicated and requires some more machinery.

Definition 6.1 *(Height of an element.) Suppose that a poset P is cycle-free, with a natural labelling $\lambda : P \to [\mathbf{n}]$. If T is a subset of P and $a \in P$, let $T[a]$ denote the path component of a in the comparability graph of $T \cup \{a\}$. Define*

$$h(a, T) = \max\{\lambda(z) \mid z \in T[a]\}.$$

Further, let

$$h(T) = \max\{\lambda(z) \mid z \in T\}.$$

Lemma 5.7 suggests a marking rule in which

$$M(T) = \{x \in T \mid \lambda(x) = h(x, T) < h(T)\}.$$

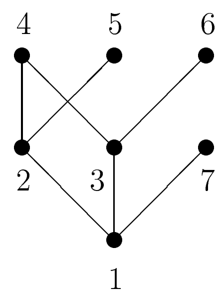
In other words, $M(T)$ consists of the maximum-labeled elements in each of the path components of T in the comparability graph of P, excluding the element whose label equals $h(T)$. This rule gives the correct number of codons and coding sequences. However, Algorithm A is generally neither injective nor surjective in this case, and a rather subtle modification is required to make it work. To explain the details precisely, we need two more definitions.

Definition 6.2 *(Relative height.) If T is a subset of P, define $T/b = \{x \in T \mid x \not\leq_P b\}$. Then $h(a, T/b)$ is called the* height of a relative to (T, b).

Definition 6.3 *(Stability.) If a and b are elements of P, and T is a subset of P, then a is* stable relative to (T, b) *if $h(a, T/b) \leq h(b, T/a)$.*

The last definition is somewhat tricky, so we will give several examples to illustrate it.

Example 6.4 Consider the poset



Then:

- $h(1, \{6,5,4\}/4) = 6$ and $h(4, \{6,5,4\}/1) = 4$, so 1 is unstable relative to $(\{6,5,4\}, 4)$.
- $h(2, \{6,5,4,1\}/4) = 5$ and $h(4, \{6,5,4,1\}/2) = 4$, so 2 is unstable relative to $(\{6,5,4,1\}, 4)$.
- $h(1, \{6,5,4\}/5) = 6$ and $h(5, \{6,5,4\}/1) = 5$, so 1 is unstable relative to $(\{6,5,4\}, 5)$.
- $h(2, \{6,5,4,1\}/5) = 4$ and $h(5, \{6,5,4,1\}/2) = 5$, so 2 is stable relative to $(\{6,5,4,1\}, 5)$.

The modified decoding algorithm can now be described. As before, given a coding sequence

$$(T_0, (T_1, x_1), (T_2, x_2), \ldots, (T_k, x_k)),$$

our object is to produce a word $w \in \mathcal{L}(\mathcal{A})$ having the same weight.

Algorithm C (Decoding for Cycle-Free Posets):

(1) *Initially let $w = [T_0]_\downarrow$.*

(2) *If w has been defined, adjoin codon (T, x) to w as follows:*

- *Preparation step: if w is nonempty, let z denote the last element of w, and suppose that $z <_P x$. If z is unstable relative to (T, x), then insert z in T and delete it from w. Repeat this step until $z \not<_P x$, w is empty, or z is stable.*
- *Then apply the rule*

$$w(T,x) \longmapsto \begin{cases} wx[T-x]_\downarrow & \text{if } z \not<_P x \text{ or } w \text{ is empty} \\ w[T]_\downarrow & \text{otherwise} \end{cases}$$

(3) *Repeat step (2) with $(T_1, x_1), (T_2, x_2), \ldots, (T_k, x_k)$, until all codons have been adjoined.*

Example 6.5 Using the data in Example 6.4 as a guide, the reader may check that for the poset P in that example, Algorithm C gives the following decodings:

$$\begin{array}{rcl} (21)(654,4) & \longmapsto & 46521 \\ (21)(654,5) & \longmapsto & 26541 \end{array}$$

We can now state the principal theorem of this section.

Theorem 6.6 *If P is a cycle-free poset with a natural labeling, then Algorithm C gives a weight-preserving bijection between coding sequences $(T_0, (T_1, x_1), \ldots, (T_k, x_k))$ and words in $\mathcal{L}(\mathcal{A})$.*

Proof. We must prove that the algorithm is well defined, and also that it is bijective. To prove the former, it suffices to prove that step (2) always produces a valid word in $\mathcal{L}(\mathcal{A})$. In case $z \not<_P x$ or w is empty, the algorithm produces $wx[T-x]_\downarrow$, which is a valid word, since if t^+ denotes the maximum-labelled element of T, nothing larger than t^+ is ever added to T. In the alternative case, $z <_P x$, z is stable with respect to (T, x), and the algorithm produces $w[T]_\downarrow$, which is a valid word unless $z <_P t^+$. But $z <_P t^+$ is impossible, since then we would have $t^+ = h(z, T/x) \le h(x, T/z) \le t^+$, which implies that there exists a chordless path[1] from x to t^+ that does not contain z. Combining this path with the path $t^+ \to z \to x$ yields a cycle which either is chordless or contains a chordless cycle of length greater than three, contradicting the cycle-free hypothesis. Hence the algorithm produces valid words in all cases.

Using Lemma 5.7 and formula (4) we can deduce that the number of words of a given weight equals the number of coding sequences of that weight. Hence to complete the proof it suffices to prove that Algorithm C is surjective. The proof begins as in the proof of Lemma 4.6. Let $w \in \mathcal{L}(\mathcal{A})$, and consider the factorization $w = w_0\, z\, u$, where z is the rightmost ascent in w, and u is a maximal decreasing subword. Let U be the set of letters occurring in u, let $T = U \cup \{z\}$. Our goal is to determine a word w' and a codon (T', z^*) so that w is obtained by adjoining (T', z^*) to w'. The first task is to determine an appropriate "markable" $z^* \in T$. The choice of z^* is somewhat tricky, and requires the following lemma.

[1]Note that, although this was not possible in the original codon (T, x), new elements may have been added to T in step 2 of Algorithm C.

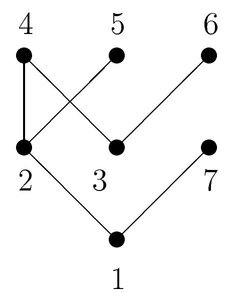
Lemma 6.7 *Given $w \in \mathcal{L}(\mathcal{A})$, let z and T be defined as in the previous paragraph. If z is not a maximal element of T, then there exists a unique maximal element $z^* \in T$ such that $z <_P z^*$ and $h(z, T/z^*) < h(z, T/z)$.*

Proof. Let z^+ be the element of T such that $h(z, T/z) = \lambda(z^+)$, i.e., z^+ has the maximum label among all elements $y \in T$ connected to z by a path that begins with an upward step and does not use any elements $<_P z$. Let

$$z <_P \gamma_1 >_P \gamma_2 <_P \cdots <_P z^+ \tag{21}$$

be a chordless path with this property, and assume further (without loss of generality) that γ_1 is a maximal element of T. Then $z^* = \gamma_1$ satisfies the conclusion of the lemma, since if there were a path from z to z^+ in T/z^*, then combining it with (21) gives a cycle that is either chordless or contains a chordless cycle of length greater than three. Similar reasoning shows that γ_1 is the unique element with the stated property.

Example 6.8 Consider the poset



If $w = 1\,6\,5\,4\,3\,2$, then $z = 1$, $h(z, T/z) = z^+ = 6$, $z^* = 4$, and $h(z, T/z^*) = 5$. If $w = 2\,7\,5\,1$, then $z = 2$, $h(z, T/z) = z^+ = 5$, $z^* = 5$, and $h(z, T/z^*) = 2$. If $w = 2\,7\,6\,5\,4\,3\,2\,1$, then $z = 2$, $h(z, T/z) = z^+ = 6$, $z^* = 4$, and $h(z, T/z^*) = 5$.

Continuing with the proof of Theorem 6.6, if $h(z^*, T) = \lambda(z^*)$, then (T, z^*) is a codon and we can proceed as in Lemma 4.6. However, in general $h(z^*, T) > \lambda(z^*)$, and we must identify and remove a collection of elements from T before the analog of Lemma 4.6 can be implemented. For the details of this procedure, we require two more lemmas.

Lemma 6.9 *Suppose that T and z^* are as defined above, and α and α' are elements of T such that $\alpha, \alpha' \leq_P z^*$. If*

$$h(\alpha, T/z^*) \;=\; h(\alpha', T/z^*),$$

then either $\alpha \leq_P \alpha'$ or $\alpha' \leq_P \alpha$.

Proof. If α and α' are incomparable, let $\beta \in T$ be the element such that $h(\alpha, T/z^*) = h(\alpha', T/z^*) = \lambda(\beta)$. Combining the path from z^* to α to β with the path from β to α' to z^*, we obtain a cycle which is either chordless or contains a chordless cycle of length greater than three.

Lemma 6.10 *Let P be any finite poset, and let $f : P \to \mathbf{Z}$ be any function with the property that, if α and β are incomparable elements of P, then $f(\alpha) \neq f(\beta)$. Then there exists a unique permutation $\alpha_1, \alpha_2, \ldots, \alpha_{|P|}$ of the elements of P such that*

(1) $\alpha_i \not<_P \alpha_{i+1}$ *for all* $i < |P|$, *and*

(2) $f(\alpha_i) > f(\alpha_j)$ *for all* $j > i$ *such that* $\alpha_j \not<_P \alpha_i$.

Proof. As before, if $x \in P$ we let P/x denote the set $\{y \in P \mid y \not\leq_P x\}$. Consider the set

$$A \;=\; \{x \in P \mid f(x) > f(y) \text{ for } y \in P/x\}.$$

Let $x \in P$ be an element for which $f(x)$ is maximum, and such that x is maximal with this property. From our assumption about f, it follows that $x \in A$, and hence A is nonempty. We claim that the elements of A form a totally ordered subset of P. Suppose, otherwise, that $a, b \in A$ with a and b incomparable. Then $a \in P/b$ and $b \in P/a$, which implies that $f(a) > f(b)$ and $f(b) > f(a)$, which is a contradiction. This establishes the claim.

Since A is nonempty and totally ordered, we can define $\alpha_1 = \max(A)$. Next, remove α_1 from P and define α_2 in a similar fashion, continuing with α_3, etc., to obtain a permutation $\alpha_1, \alpha_2, \ldots, \alpha_{|P|}$ of P. Clearly, this permutation satisfies condition (2). To prove condition (1), it suffices to show that $\alpha_1 \not<_P \alpha_2$. Suppose, to the contrary, that $\alpha_1 <_P \alpha_2$. We have

$$f(\alpha_2) > f(y) \text{ for all } y \in (P - \alpha_1)/\alpha_2.$$

But $(P - \alpha_1)/\alpha_2 = P/\alpha_2$, since $\alpha_1 <_P \alpha_2$. Thus $\alpha_2 \in A$, which implies that α_2 was the proper choice instead of α_1 in the initial step of the construction. This is a contradiction, and hence $\alpha_1 \not<_P \alpha_2$.

To prove that the permutation $\alpha_1, \alpha_2, \ldots, \alpha_{|P|}$ is unique, suppose that

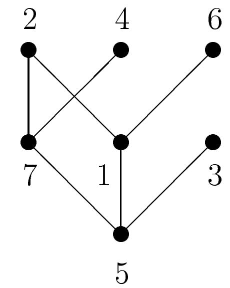
$$\beta_1, \beta_2, \ldots, \beta_k, \alpha_1, \beta_{k+2}, \ldots, \beta_{|P|} \tag{22}$$

is some other permutation satisfying conditions (1) and (2), with $\beta_1 \neq \alpha_1$. Then $\beta_1 \in A$, by (2), and hence $\beta_1 <_P \alpha_1$. Define $A_2, A_3, \ldots, A_k$ to be the sets corresponding to A for the segments of permutation (22) beginning with $\beta_2, \beta_3, \ldots, \beta_k$. It follows from condition (2) that α_1 and β_i are both in A_i, for $i = 2, \ldots, k$, and hence α_1 and β_i are comparable for each i. We claim that, in fact, $\beta_i <_P \alpha_1$ for all i. To see this, suppose first that $\alpha_1 <_P \beta_2$. Then we have

$$\beta_1 <_P \alpha_1 <_P \beta_2$$

and hence $\beta_1 <_P \beta_2$, contradicting condition (1). Hence $\beta_2 <_P \alpha_1$, and a similar argument applies to each successive β_i. In particular, $\beta_k <_P \alpha_1$, which contradicts (1), and so we conclude that $\beta_1 = \alpha_1$. A straightforward induction proves that $\beta_i = \alpha_i$ for all i, which completes the proof of uniqueness and hence of Lemma 6.10.

Example 6.11 Consider the following poset P with function values $f(x)$ as shown.



If we denote each element x by the corresponding $f(x)$, then the unique permutation specified by Lemma 6.10 is

$$7\,6\,4\,3\,2\,1\,5$$

We continue now with the proof of Theorem 6.6. To prove surjectivity of Algorithm C, we must show that, to each word $w \in \mathcal{L}(\mathcal{A})$ we can find a word $w' \in \mathcal{L}(\mathcal{A})$ and a codon (T', x') such that w equals the result of adjoining (T', x') to w'. The process will be presented in the form of an algorithm, essentially constructing an inverse for Algorithm C.

Algorithm D (Encoding for Cycle-Free Posets):

(1) *Given $w \in \mathcal{L}(\mathcal{A})$, let U be the set of letters in the word u, where $w = w_0\, z\, u$ and z is the rightmost ascent in w. Let $T = U \cup \{z\}$.*

(2) *If z is a maximal element of T, let $z^* = z$. Otherwise choose z^* as in Lemma 6.7.*

(3) *Let U_0 be the set of elements $\alpha < z^*$ in U such that $h(\alpha, T/z^*) > \lambda(z^*)$. For each $\alpha \in U_0$, define $f(\alpha) = h(\alpha, T/z^*)$.*[2]

(4) *Let $\alpha_1\alpha_2 \ldots \alpha_k$ be the unique permutation of U_0 specified by Lemma 6.10.*

(5) *Define w' and (T', x') as follows:*

- *If $z = z^*$ and $z \notin U$, then $w' = w_0\, \alpha_1\alpha_2 \ldots \alpha_k$ and $(T', x') = (T - U_0, z^*)$.*
- *Otherwise $w' = w_0\, z\, \alpha_1\alpha_2 \ldots \alpha_k$ and $(T', x') = (U - U_0, z^*)$.*

It remains to show that, if Algorithm D yields w' and (T', x') as output, then w' is a valid word, (T', x') is a valid codon, and $w'\,(T', x') \mapsto w$ as a consequence of Algorithm C. We state these results as a series of lemmas.

Lemma 6.12 *$w' \in \mathcal{L}(\mathcal{A})$.*

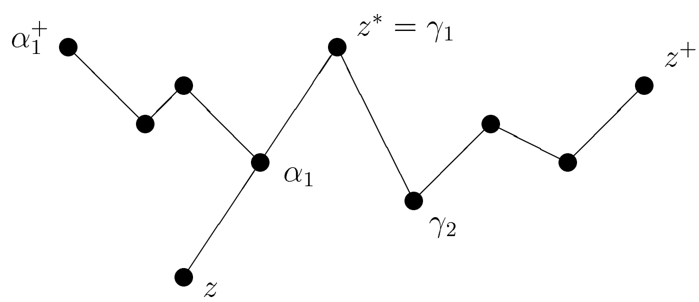
Proof. Suppose first that $z = z^*$ and $z \notin U$, so that $w' = w_0\alpha_1\alpha_2 \cdots \alpha_k$. If $w' \notin \mathcal{L}(\mathcal{A})$, then by Lemma 6.10 (1), we necessarily have $z' \leq_P \alpha_1$, where z' denotes the last letter of w_0. But $\alpha_1 <_P z^* = z$, so that $z' <_P z$, which implies $w \notin \mathcal{L}(\mathcal{A})$, a contradiction.

Suppose next that $z <_P z^*$ or $z \in U$, so that $w' = w_0 z \alpha_1 \cdots \alpha_k$. If $w' \notin \mathcal{L}(\mathcal{A})$, we necessarily have $z \leq_P \alpha_1$. Since $\alpha_1 \in U_0$, $h(\alpha_1, T/z^*) > \lambda(z^*)$, which implies

[2] Note that, by Lemma 6.9, f satisfies the hypotheses of Lemma 6.10.

$h(z, T/z) > \lambda(z^*)$. Following the construction in Lemma 6.7, let z^+ be the element of T with $h(z, T/z) = \lambda(z^+)$, and consider the chordless path (21) from z to z^+, where now $\gamma_1 = z^*$.

We claim that $\alpha_1 \not\geq_P \gamma_2$, since otherwise (to avoid chordless cycles of length 4) the elements $\geq_P \alpha_1$ (including z^*) would have to be linearly ordered. This would force $h(\alpha_1, T/z^*) = \lambda(\alpha_1) < \lambda(z^*)$, contradicting the fact that $\alpha_1 \in U_0$. The following picture suggests how the elements we have been considering might be arranged in P.



Here $\lambda(\alpha_1^+) = h(\alpha_1, T/z^*)$. Note that

$$h(\gamma_2, T/z^*) = \lambda(z^+) > \lambda(\alpha_1^+) = h(\alpha_1, T/z^*) > \lambda(z^*)$$

Thus both α_1 and γ_2 are in U_0. Using the notation of step (3) of Algorithm D, we have $f(\gamma_2) > f(\gamma_1)$, and, since $\alpha_1 \not\geq_P \gamma_2$, this means that γ_2 precedes α_1 in the order specified by Lemma 6.10. This contradicts the assumption that α_1 was the first element of that order, and the proof is complete.

Lemma 6.13 *(T', x') is a valid codon, i.e., $\lambda(x') = h(x', T') < h(T')$.*

PROOF. We have $(T', x') = (T - U_0, z^*)$ or $(U - U_0, z^*)$. Suppose that $h(z^*, T') > \lambda(z^*)$. Then, since z^* is maximal in T, there exists an element $y <_P z^*$ in T' such that $h(z^*, T') = h(y, T'/z^*)$. Clearly, $h(y, T'/z^*) \leq h(y, T/z^*)$. Thus we have

$$\lambda(z^*) < h(z^*, T') = h(y, T'/z^*) \leq h(y, T/z^*)$$

This implies $y \in U_0$, which implies $y \notin T'$, a contradiction. Finally, if m is the maximum-labelled element of T, then $z^* = m$ is impossible, since $z <_P z^*$ and $w \in \mathcal{L}(\mathcal{A})$.

Lemma 6.14 *$w'(T', x') \mapsto w$ as a consequence of Algorithm C.*

PROOF. There are two cases. If

$$w' = w_0\alpha_1\alpha_2 \cdots \alpha_k \text{ and } (T', x') = (T - U_0, z^*),$$

then $z = z^*$ and $z \notin U$. Step (2) of Algorithm C begins by inserting unstable elements into T. From step (3) of Algorithm D, we see that each α_i satisfies $h(\alpha_i, T/z^*) > \lambda(z^*)$, and this implies that each α_i is initially unstable, since trivially $h(\alpha_i, T/z^*) = h(\alpha_i, T'/z^*)$. We need to show that α_i is unstable with respect to the set $T^+ = T' \cup \{\alpha_{i+1}, \ldots, \alpha_k\}$, i.e. it remains unstable after $\alpha_{i+1}, \ldots, \alpha_k$ have been added. But this is guaranteed by the ordering in step (4) of Algorithm D, which insures that $h(\alpha_i, T/z^*) > h(\alpha_j, T/z^*)$ for all α_j with $j > i$ and $\alpha_i \not>_P \alpha_j$. Hence Algorithm C inserts each of the α_i, as desired. If z' is the last element of w_0, z' causes step (2) of Algorithm C to terminate, since $z' \not\leq_P z = z^*$.

In the second case we have

$$w' = w_0 z \alpha_1 \alpha_2 \cdots \alpha_k \text{ and } (T', x') = (U - U_0, z^*).$$

Again, step (2) of Algorithm C inserts each of the elements α_i, and it remains to show that z causes this step to terminate, i.e., z is stable. We have $h(z, T/z) \leq h(z^*, T/z)$ since z^* lies on a path from z to z^+ (see the proof of Lemma (6.12)). Trivially, $h(z, T/z^*) \leq h(z, T/z)$. Combining these last two statements, we get

$$h(z, T/z^*) \leq h(z, T/z) \leq h(z^*, T/z),$$

and hence z is stable. This completes the proof.

7 Specializing to length

If we set $x_1 = x_2 = \cdots = x_n = t$ in the pair-avoiding series $F_{\mathcal{A}}(\mathbf{x})$, we get the ordinary generating function $F_{\mathcal{A}}(t)$ that enumerates words in $\mathcal{L}(\mathcal{A})$ by length. From (12), we obtain immediately

$$F_{\mathcal{A}}(t) \;=\; \frac{(1+t)^n}{\sum_{S \in \mathcal{L}(\overline{\mathcal{A}}_0)} (-t)^{|S|}(1+t)^{n-|S|}},$$

and thus if the longest word in $\mathcal{L}(\overline{\mathcal{A}}_0)$ has length M, we can cancel the factor $(1+t)^{n-M}$ in the numerator and denominator of this expression, to obtain

$$F_{\mathcal{A}}(t) \;=\; \frac{(1+t)^M}{\sum_{S \in \mathcal{L}(\overline{\mathcal{A}}_0)} (-t)^{|S|}(1+t)^{M-|S|}}. \tag{23}$$

Note that, equivalently, M is simply the *height* of the relation $\mathcal{A}$. When $M < n$, we have applied a non-trivial cancellation to obtain (23), so combinatorial explanations do not follow immediately from specializing the algorithms we have considered earlier in this paper. We conclude by providing combinatorial explanations for two cases of (23), which, in turn, raise other questions worth exploring.

7.1 13-avoiding words on $\{1, 2, 3\}$

This example was introduced in the introduction. Let $\mathcal{A} = \{(11), (22), (33), (13)\}$, with $n = 3$. Then

$$F_{\mathcal{A}}(t) \;=\; \frac{(1+t)^3}{1 - 2t^2 - t^3} = \frac{(1+t)^2}{1 - t - t^2},$$

where the first form of the answer comes from (12), and the second from (23). If $f(L)$ denotes the number of words in $\mathcal{L}(\mathcal{A})$, the first form of the answer shows that f satisfies the recurrence

$$f(L) = 2f(L-2) + f(L-3), \quad L \geq 3.$$

Since there are two codons of weight 2 and one of weight 3, Algorithm A gives a simple combinatorial interpretation of this recurrence. Explaining the Fibonacci recurrence suggested by the second form of the answer is less straightforward, but there are fairly simple solutions. Here is one: given a word w in $\mathcal{L}(\mathcal{A})$ of length $L \geq 3$, remove the last two letters if w ends in 23 or 32. Otherwise remove the last letter. The number of words obtained in each case is $f(L-2)$ and $f(L-1)$, respectively, and it is easy to see that this correspondence is bijective.

7.2 Bipartite/Alternating Forests

If a cycle-free poset has height 2, it may be represented as a bipartite graph which is also a forest. When endowed with a natural labelling, such a poset becomes an *alternating forest*, in the sense of [21] and [22]. When the forest is an *alternating tree*, the generating function $F_{\mathcal{A}}(t)$ has an especially simple form.

Theorem 7.1 *Suppose the relation $\mathcal{A}$ is defined by an alternating tree with n vertices. Then*

$$F_{\mathcal{A}}(t) \;=\; \frac{(1+t)^2}{1-(n-2)t}. \tag{24}$$

Consequently, if $f(L)$ denotes the number of words in $\mathcal{L}(\mathcal{A})$ of weight L, then

$$f(L) = (n-1)^2(n-2)^{L-2}, \quad L \geq 2, \tag{25}$$

Both of these results depend only on n and L, and not on the tree itself.

Proof. From (23), we obtain

$$F_{\mathcal{A}}(t) \;=\; \frac{(1+t)^2}{(1+t)^2 - nt(1+t) + (n-1)t^2},$$

which gives (24). Formula (25) for $f(L)$ then follows, since

$$F_{\mathcal{A}}(t) \;=\; 1 + nt + \frac{(n-1)^2t^2}{1-(n-2)t}.$$

We now give a combinatorial explanation for (24) and (25). First, some notation. Since $\mathcal{A}$ is defined by a tree, there is a unique (alternating) path between any pair of distinct vertices. For $a, b \in [n]$, $a \neq b$, let $p_a(b)$ be the vertex immediately preceding b on the unique path from a to b in the tree defining $\mathcal{A}$. Given $w \in \mathcal{L}(\mathcal{A})$ of length $L-1 \geq 2$, where $w = w_0ij$ (so i, j are the last two letters of w), let $k = p_i(j)$. Then consider any $m \in [n]$, with $m \neq j, k$, and from the pair w, m construct w' as follows:

Case 1: If $wm \in \mathcal{L}(\mathcal{A})$, then $w' = wm = w_0ijm$;
Case 2: otherwise, $w' = w_0imj$.

Note that w' is a word in $\mathcal{L}(\mathcal{A})$ of length L: for Case 1 this is immediate; for Case 2 this follows because we have $j\mathcal{A}m$, and $i \neq j$, so $i\mathcal{A}m$ is only possible for $m = k$, which is not an allowable choice of m. We claim also that every word w' in $\mathcal{L}(\mathcal{A})$ of length L is created uniquely in the above construction (referred to below as the *tree construction*), and we leave it to the reader to check this, say, by specifying the inverse mapping. Thus, since there are $n-2$ choices of m for each word w, we have a combinatorial proof that

$$f(L) = (n-2)\, f(L-1), \qquad L \geq 3,$$

for a relation defined by an alternating tree.

The following more general result holds for a relation defined by any bipartite graph.

Theorem 7.2 *If $\mathcal{A}$ is defined by any bipartite graph with n vertices and E edges, then*

$$F_{\mathcal{A}}(t) \;=\; \frac{(1+t)^2}{1-(n-2)t+(E-n+1)t^2}. \tag{26}$$

PROOF. This follows immediately from (23).

We conclude the paper with a combinatorial explanation of (26), which uses the above tree construction, and is given in two stages. First, we consider the case where $\mathcal{A}$ is defined by a connected graph. In this case, find a spanning tree of the graph, and let $\mathcal{T}$ be the relation defined by this spanning tree. Let the $d = E - n + 1$ non-tree edges of $\mathcal{A}$ be given by $\{x_1, y_1\}, \ldots, \{x_d, y_d\}$, where the x_i are minimal elements of $\mathcal{A}$, and the y_i are maximal elements of $\mathcal{A}$. Now note that $\mathcal{L}(\mathcal{A}) \subseteq \mathcal{L}(\mathcal{T})$, so given $w = w_0ij \in \mathcal{L}(\mathcal{A})$, we also have $w \in \mathcal{L}(\mathcal{T})$. Apply the above tree construction to w, m (where $m \neq j, k$, and $k = p_i(j)$ in the tree relation $\mathcal{T}$). The construction again produces a word $w' \in \mathcal{L}(\mathcal{T})$ of length L, but also there are $w' \notin \mathcal{L}(\mathcal{A})$ that are constructed. These extras are:

- for each $t = 1, \ldots, d$, one occurrence of $w' = w_0imj = w_0x_ty_t\, p_{x_t}(y_t)$,
- for each $t = 1, \ldots, d$ and each $i \neq x_t$, one occurrence of $w' = w_0ijm = w_0ix_ty_t$.

Note that these extras are easy to count: if we remove the last two letters from all such w', then we obtain each word w_0i in $\mathcal{L}(\mathcal{A})$ of length $L-2$ exactly $d = E-n+1$ times, i.e. from words obtained by appending, for $t = 1, \ldots, d$, either $y_t\, p_{x_t}(y_t)$ or x_ty_t according to which of the above cases applies. Thus the total number of extras is $(E-n+1)\, f(L-2)$, and we now have completed a combinatorial proof that

$$f(L) = (n-2)\, f(L-1) - (E-n+1)\, f(L-2), \qquad L \geq 3, \tag{27}$$

for a relation defined by a connected bipartite graph with E edges and n vertices. This proof has involved correction for an overcount, since if the connected graph is not simply a tree, then $E - n + 1 > 0$.

We now turn to a combinatorial proof of (27) in the case of a relation $\mathcal{A}'$ defined by a graph which is not connected. Suppose that $\mathcal{A}'$ is defined by a graph with $c+1$ components, $c \geq 1$. Add c edges $\{u_1, v_1\}, \ldots, \{u_c, v_c\}$ to the graph, where the u_i are minimal elements of $\mathcal{A}'$, and the v_i are maximal elements of $\mathcal{A}'$, so that the resulting graph is connected (for an isolated vertex, arbitrarily assign it to the set of minimal or maximal vertices). Let $\mathcal{A}$ be the relation defined by this connected graph, and let $\mathcal{T}$ be the relation defined by a spanning tree. Suppose that $w = w_0 ij \in \mathcal{L}(\mathcal{A}')$ is a word of length $L-1$. We use $\mathcal{T}$ to define $p_i(j)$, and then for each $m \neq j, p_i(j)$ apply the tree construction to w, m, obtaining a word w' of length L. It is straightforward to check that this map is injective, and that, as in the previous argument, the resulting words are all in $\mathcal{L}(\mathcal{A}')$ with $d = (E+c) - n + 1$ exceptions. Moreover, there are $w' \in \mathcal{L}(\mathcal{A}')$ that are *not* constructed. The omitted cases are:

- for each $s = 1, \ldots, c$, one occurrence of $w' = w_0 u_s v_s u_s$,
- for each $s = 1, \ldots, c$ and each $i \neq u_s$, one occurrence of $w' = w_0 i u_s v_s$.

These cases are also easy to count: if we remove the last two letters from all such w', then we obtain each word $w_0 i \in \mathcal{L}(\mathcal{A}')$ of length $L-2$ exactly c times, i.e., from words obtained by appending, for $s = 1, \ldots, c$, either $u_s v_s$ or $v_s u_s$, according to which of the above cases applies. Thus the total number of omitted cases is is $c\, f(L-2)$, and we now have completed a combinatorial proof that

$$(n-2)\, f(L-1) = f(L) + d\, f(L-2) - c\, f(L-2), \quad L \geq 3,$$

for a relation defined by a bipartite graph with E edges and n vertices, where d is the number of edges of $\mathcal{A}$ not in a spanning tree of $\mathcal{A}$. But $\mathcal{A}$ has $E+c$ edges, and a spanning tree of $\mathcal{A}$ has $n-1$ edges, so $d - c = E - n + 1$, and (27) follows.

In the disconnected case, we can have $E-n+1 < 0$. As an example of this, consider the 13-avoiding words on $\{1,2,3\}$, and additional edge $\{u_1, v_1\} = \{2,3\}$. The combinatorial construction above has the simplified statement that was given in Section 7.1.

In both the connected and disconnected case, we can have $E - n + 1 > 0$, in which case the combinatorial construction above involves correction for overcounting. As of this writing, we do not know how to introduce such a correction into our algorithms for arbitrary relations, or even posets (which would be required when we don't have μ-positivity).

8 Rearrangements of a multiset

In this section we explain the relationship between our results and the theory of rearrangements of a multiset, as developed in [5] (see also [17] for an excellent exposition of this topic). [3]

[3] We thank the referee for asking questions that prompted us to add this section.

Suppose $\mathbf{X}$ is a set of n elements, which we now assume to have a linear ordering. If M is a multiset formed from the elements of $\mathbf{X}$, a *rearrangement* σ of M is a two-line array $\binom{M}{\sigma(M)}$, where the first row consists of the elements of M in weakly increasing order, and the second row is a permutation of those elements. Alternately, we will represent σ by (and consider equivalent to σ) any array obtained by permuting the columns of σ, as long as columns with identical first row entries appear in the same order.

If $\mathcal{A}$ is an arbitrary relation on $\mathbf{X}$, let $\mathcal{K}(\mathcal{A})$ denote the set of rearrangements σ of $\mathbf{X}$ containing no column $\binom{x}{y}$ with $(x,y) \in \mathcal{A}$, i.e., $\mathcal{K}(\mathcal{A})$ is the set of rearrangements that "avoid" $\mathcal{A}$ in columns. Let $G_{\mathcal{A}}(\mathbf{x})$ be the generating series for $\mathcal{K}(\mathcal{A})$, where σ is marked by $\mathbf{x}^{\sigma(M)}$ as before. Then, by the MacMahon Master Theorem [20],

$$G_{\mathcal{A}}(\mathbf{x}) = \frac{1}{\det(I - \mathbf{X}\bar{A})}, \tag{28}$$

where $\bar{A}$ is the $n \times n$ matrix defined in Theorem 2.1. Comparing this with Theorem 2.1 and proceeding as in Corollary 2.3 we obtain

$$G_{\mathcal{A}}(\mathbf{x}) = \frac{1}{1 - \sum_T \mu_{\mathcal{A}}(T)\mathbf{x}^T}, \tag{29}$$

where $\mu_{\mathcal{A}}$ is defined as in Corollary 2.3.[4] Examining the proof of Corollary 2.3, we see that (28) and (29) hold for arbitrary relations $\mathcal{A}$, not necessarily assumed to be monotone.

When $\mathcal{A}$ is monotone, formulae (9) and (29) differ only by a multiplicative factor of the form $\prod_{i=1}^{n}(1+x_i)$. The next lemma gives a simple bijection linking these two formulae.

Lemma 8.1 *Suppose that $\mathcal{A}$ is a monotone relation. Then there is a bijection*

$$w \longleftrightarrow (S, \sigma)$$

between words $w \in \mathcal{L}(\mathcal{A})$ and pairs (S, σ), where $S \subseteq \mathbf{X}$ and $\sigma \in \mathcal{K}(\mathcal{A})$.

Proof. Suppose that $w \in \mathcal{L}(\mathcal{A})$. There exists a unique factorization $w = w^{(1)}w^{(2)} \cdots w^{(m)}$ into subwords $w^{(i)}$ such that, if b_i denotes the first letter of $w^{(i)}, i = 1, \ldots, m$, then b_i is the unique smallest letter in $w^{(i)}$ and $b_1 \geq b_2 \geq \cdots b_m$. The b_i are the weak left-to-right minima of w, i.e., the letters b such that $a \geq b$ for all a preceding b in w.

Since w avoids $\mathcal{A}$, each $x \in \mathbf{X}$ can appear at most once as a singleton word among the $w^{(i)}$. Moreover, such a word must be the rightmost $w^{(i)}$ for which $b_i = x$. Delete all such singleton words, and define S to be the set of elements contained in them. Next, denoting the remaining subwords in left-to-right order by $w^{(1)}, w^{(2)}, \ldots, w^{(k)}$, associate to each $w^{(i)}$ a "generalized cycle" $\gamma^{(i)}$, constructed as a two-line array by writing $w^{(i)}$ in the first row, and then cyclically shifting it one step to the left to form the second row. Finally, construct a rearrangement σ by concatenating the two-line arrays of $\gamma^{(k)}, \gamma^{(k-1)}, \ldots, \gamma^{(1)}$ (in that order, i.e., reversed) and sorting the first row of the resulting array in weakly

[4] Despite a superficial resemblance, our μ is not the same as the Möbius function appearing in [5], Chapters 2 and 5.

increasing order. (The operation of concatenation and sorting is one interpretation of the "intercalation product" defined in [5]). It is easy to check that the map from w to (S, σ) is a bijection with the desired properties, and the proof is complete.

We call σ the *canonical rearrangement* constructed from $w^{(1)}, \ldots, w^{(k)}$. By a *generalized cycle*, or sometimes just a *cycle*, we will always mean a two-line array in which the minimum element b occurs uniquely, as the first element of the first row and the last element of the second row. Our factorization of w into subwords $w^{(i)}$ is a variation on one that appears in [5], called the *descending decomposition* of a word. Although our construction is slightly different, we will also use this phrase to describe it. See also [17], especially Theorem B of Section 5.1.2, and Problem 5.1.2 #5, which accounts precisely for the absence of cycles of length one in rearrangements $\sigma \in \mathcal{K}(\mathcal{A})$.

Example 8.2 *Let $\mathcal{A}$ be the diagonal relation on $\mathbf{X} = \{1, \ldots, 5\}$, so that $\mathcal{L}(\mathcal{A})$ consists of Smirnov words (see Example 3.3). Let*

$$w = 3\,4\,5\,3\,2\,4\,1\,3\,2\,1\,3\,1\,.$$

Then the descending decomposition of w is

$$3\,4\,5\,|\,3\,|\,2\,4|\,1\,3\,2\,|\,1\,3|\,1,$$

so that $S = \{1, 3\}$. Removing S and concatenating the remaining cycles in reverse order give the two-line array

$$\begin{pmatrix} 1 & 3 & 1 & 3 & 2 & 2 & 4 & 3 & 4 & 5 \\ 3 & 1 & 3 & 2 & 1 & 4 & 2 & 4 & 5 & 3 \end{pmatrix}.$$

After sorting the last array, we obtain

$$(S, \sigma) \;=\; \left(\{1, 3\}\,, \begin{pmatrix} 1 & 1 & 2 & 2 & 3 & 3 & 3 & 4 & 4 & 5 \\ 3 & 3 & 1 & 4 & 1 & 2 & 4 & 2 & 5 & 3 \end{pmatrix} \right).$$

By analogy with our earlier results, it is also natural to ask whether (29) has a direct combinatorial explanation when $\mathcal{A}$ is μ-positive, i.e., an encoding of rearrangements $\sigma \in \mathcal{K}(\mathcal{A})$ by sequences of codons. We will not answer this question for arbitrary $\mathcal{A}$, or even monotone relations in general, but will solve the problem completely for V-free relations by giving an encoding/decoding algorithm for rearrangements analogous to that given for words in Sections 3 and 4. When applied to the diagonal relation (Smirnov words), this algorithm gives a combinatorial interpretation of MacMahon's generating function for derangements of a multiset (formula (15) with the numerator replaced by 1), as discussed in Example 3.3.

Suppose that $\mathcal{A}$ is a V-free monotone relation. By Lemma 4.7, $\mathcal{A}$ is μ-positive, and we seek a bijection between rearrangements $\sigma \in \mathcal{K}(\mathcal{A})$ and coding sequences

$$\Sigma = ((T_1, x_1), (T_2, x_2), \ldots, (T_j, x_j)) \tag{30}$$

where each (T_i, x_i) is a codon (based on $\mathcal{A}$) as defined in Section 3. Note that coding sequences (30) differ from those of (14) only in the absence of an initial subset T_0. We will give a decoding algorithm that utilizes Algorithm A of Section 3. The first step is to break the coding sequence (30) into subsequences which are decoded into generalized cycles and then combined to give a rearrangement σ.

Algorithm A′ (Decoding Algorithm for Rearrangements) *The input is a coding sequence Σ of the form (30) with codons defined by $\mathcal{A}$, and the output is a rearrangement $\sigma \in \mathcal{K}(\mathcal{A})$ of the underlying multiset. The relation $\mathcal{A}$ is assumed to be V-free.*

(1) *Read Σ from left to right, keeping track of the minimum element b_i in each T_i. If $b_i \le b_j$ for all $j < i$, i.e., b_i is a weak left-right minimum, break the string and form a new coding sequence beginning with (T_i, x_i). The result is an ordered list of coding sequences $\Sigma^{(1)}, \Sigma^{(2)}, \ldots, \Sigma^{(k)}$, each of the form (30). In each $\Sigma^{(i)}$, the minimum element b_i occurs uniquely, and appears in the leftmost codon.*

(2) *Decode each $\Sigma^{(i)}$ using Algorithm A of Section 3 (in all cases, T_0 is the empty set), obtaining an ordered list of words $w^{(1)}, w^{(2)}, \ldots, w^{(k)}$ in $\mathcal{L}(\mathcal{A})$.*

(3) *If $w^{(i)}$ fails to avoid $\mathcal{A}$ as a circular word, i.e., if $w^{(i)}$ factors as xuy with $(y, x) \in \mathcal{A}$, replace the initial segment $x[T - x]_\downarrow$ in $w^{(i)}$ by $[T]_\downarrow$. Here (T, x) denotes the initial codon of $\Sigma^{(i)}$. Shift $w^{(i)}$ cyclically so that b_i appears in the first position, to obtain words $\widehat{w}^{(i)}$, $i = 1, \ldots, k$, each representing a cycle $\widehat{\gamma}^{(i)} \in \mathcal{K}(\mathcal{A})$.*

(4) *Let σ be the canonical rearrangement constructed from $\widehat{w}^{(1)}, \ldots, \widehat{w}^{(k)}$ by concatenating the $\widehat{\gamma}^{(i)}$ in reverse order, as in the proof of Lemma 8.1.*

Example 8.3 *Let $\mathcal{A}$ be the relation defined in Example 3.4, i.e., the diagonal relation plus adjacent pairs $(i, i+1)$. Let Σ be the coding sequence*

$$(7653, 5)(63, 3)(865, 6)(86, 6)(321, 1)(64, 4)(4321, 2)(621, 2)(542, 2).$$

Step (1) of Algorithm A′ *breaks Σ into subsequences $\Sigma^{(i)}$, $i = 1, \ldots, 5$, as follows:*

$$(7653, 5) \mid (63, 3)(865, 6)(86, 6) \mid (321, 1)(64, 4) \mid (4321, 2) \mid (621, 2)(542, 2).$$

Steps (2) and (3) decode each $\Sigma^{(i)}$ as a cyclically shifted word $\widehat{w}^{(i)}$, as follows:

$$3\,5\,7\,6 \mid 3\,6\,8\,6\,5\,8\,6 \mid 1\,3\,2\,4\,6 \mid 1\,4\,3\,2 \mid 1\,5\,4\,2\,6\,2.$$

Forming two-line arrays, concatenating in reverse order, and sorting, we obtain the canonical rearrangement $\sigma =$

$$\begin{pmatrix} 1 & 1 & 1 & 2 & 2 & 2 & 2 & 3 & 3 & 3 & 3 & 4 & 4 & 4 & 5 & 5 & 5 & 6 & 6 & 6 & 6 & 6 & 6 & 7 & 8 & 8 \\ 5 & 4 & 3 & 6 & 1 & 1 & 4 & 2 & 2 & 6 & 5 & 2 & 3 & 6 & 4 & 8 & 7 & 2 & 1 & 8 & 5 & 3 & 3 & 6 & 6 & 6 \end{pmatrix}.$$

Theorem 8.4 *If $\mathcal{A}$ is a V-free monotone relation, then Algorithm* A′ *gives a bijection between the set of all coding sequences of the form (30), with codons defined by $\mathcal{A}$, and the set of all rearrangements $\sigma \in \mathcal{K}(\mathcal{A})$.*

Proof. Using the fact that $\mathcal{A}$ is V-free, it is readily verified that step (3) of this algorithm always yields a word $\widehat{w}^{(i)} \in \mathcal{L}(\mathcal{A})$ that avoids $\mathcal{A}$ circularly. This in turn implies that $\widehat{\gamma}^{(i)} \in \mathcal{K}(\mathcal{A})$ for all i, and hence $\sigma \in \mathcal{K}(\mathcal{A})$.

Algorithm A$'$ constructs σ from cycles $\widehat{\gamma}^{(i)}$ such that the smallest element b_i of $\widehat{\gamma}^{(i)}$ occurs uniquely, and such that $b_1 \geq b_2 \geq \cdots \geq b_k$. This decomposition into cycles can be uniquely reconstructed from σ. Hence, to prove Theorem 8.4, it suffices to prove that Algorithm A$'$ gives a bijection between coding sequences and cycles, when restricted to coding sequences of type (30) with the property that the minimum element occurs uniquely, in the first codon. We will say that such a coding sequence is of *cyclic type*. If Σ is of cyclic type, let $\tilde{D}(\Sigma)$ denote the cycle obtained by applying Algorithm A$'$ of this section. We will construct an encoding $\tilde{E}(\gamma)$ of cycles $\gamma \in \mathcal{K}(\mathcal{A})$, and show that $\tilde{D}$ and $\tilde{E}$ are inverses.

If w is a word in $\mathcal{L}(\mathcal{A})$, denote by $E(w)$ the encoding of w by a coding sequence of the form (14), obtained by applying Algorithm B in Section 4. Note that $E(w)$ is not necessarily of the form (30), i.e., it may contain a nonempty T_0.

Suppose that $\gamma \in \mathcal{K}(\mathcal{A})$ is a generalized cycle with minimum element b and first line w'. Define w^* to be the unique cyclic shift of w' with the property that either (a) w^* is a decreasing word, or (b) $w^* = auv$, where a is an ascent, u is a decreasing subword, and b appears in u. It is immediate that $w^* \in \mathcal{L}(\mathcal{A})$. The encoding $\tilde{E}$ can now be described precisely:

Algorithm B$'$ (Encoding Algorithm for Cycles) *The input is a cycle $\gamma \in \mathcal{K}(\mathcal{A})$, and the output is a coding sequence $\tilde{E}(\gamma)$ of cyclic type.*

(1) *Define w^* as in the previous paragraph.*

(2a) *If w^* is decreasing, output $\tilde{E}(\gamma) = ((W, b^*))$ and terminate. Here W denotes the set of letters in w, and b^* denotes the unique (winning) successor of b in W, as defined in Lemma 4.5.*

(2b) *If $w^* = auv$ is not decreasing, write*

$$E(w^*) = E(auv) = (T_0, (T_1, x_1), \ldots, (T_n, x_n)),$$

the coding sequence of type (14) obtained by applying Algorithm B to w^*.

(3) *If $T_0 = \emptyset$, define*

$$\tilde{E}(\gamma) \;=\; \begin{cases} ((T_1, x_1), (T_2, x_2), \ldots, (T_n, x_n)) & \textit{if } b \in T_1 \\ ((T_2, x_2), \ldots, (T_n, x_n), (T_1, x_1)) & \textit{if } b \notin T_1. \end{cases}$$

(4) *If $T_0 = \{a\}$, first write*

$$E(uva) = (T'_0, (T'_1, x'_1), \ldots, (T'_m, x'_m)).$$

Then define

$$\tilde{E}(\gamma) \;=\; \begin{cases} ((T'_0, a^*), (T'_1, x'_1), \ldots (T'_m, x'_m)) & \textit{if } b \in T'_0 \\ ((T'_1, x'_1), \ldots (T'_m, x'_m), (T'_0, a^*)) & \textit{if } b \notin T'_0. \end{cases}$$

It is easy to verify that, in the coding sequence constructed in Step (2), T_0 is always either the empty set or the singleton $\{a\}$. Furthermore, one can check that Steps (3) and (4) always result in a coding sequence of cyclic type.

We will outline the case-by-case analysis needed to prove that $\tilde{E}(\tilde{D}(\Sigma)) = \Sigma$, where Σ is of cyclic type. We omit the details, which are straightforward but somewhat tedious. Given $\Sigma = ((T_1, x_1), (T_2, x_2), \ldots, (T_j, x_j))$, let $w = w^{(i)}$ be the word constructed initially in Step (3) of Algorithm A′, i.e., before performing the cyclic shift to get $\widehat{w}^{(i)}$. There are seven possibilities for w, with $b \in T_1$ in each case:

$$
\begin{array}{lll}
\text{(i)} & w = x_1[T_1 - x_1]_\downarrow \cdots y & \text{with } j \geq 2, x_1 \neq b,\ y \not\mathcal{A}\, x_1\,, \\
\text{(i}'\text{)} & w = x_1[T_1 - x_1]_\downarrow & \text{with } j = 1, x_1 \neq b, b \not\mathcal{A}\, x_1\,, \\
\text{(ii)} & w = [T_1]_\downarrow \cdots s & \text{with } j \geq 2, s\mathcal{A}x_1\,, \\
\text{(ii}'\text{)} & w = [T_1]_\downarrow & \text{with } j = 1, b\mathcal{A}x_1\,, \\
\text{(iii)} & w = b[T_1 - b]_\downarrow \cdots x_j[T_j - x_j]_\downarrow & \text{with } j \geq 2, \\
\text{(iii}'\text{)} & w = b[T_1 - b]_\downarrow & \text{with } j = 1, \\
\text{(iv)} & w = b[T_1 - b]_\downarrow \cdots t[T_j]_\downarrow & \text{with } j \geq 2, x_j = t^*\,.
\end{array}
$$

Step (4) of Algorithm A′ then constructs $\gamma = \tilde{D}(\Sigma) = \binom{w'}{w''}$, where w' and w'' are appropriate cyclic shifts of w. When $\tilde{E}$ is applied to γ using Algorithm B′, the first step is to compute w^*, which is also a cyclic shift of w. In cases (ii′) and (iii′) above, w^* is decreasing (in both cases $w^* = [T_1]_\downarrow$), and we immediately obtain $\tilde{E}(\gamma) = \Sigma$ in these cases. For the remaining five cases, we write $w^* = auv$, as follows:

$$
\begin{array}{ll}
\text{(i)} & w^* = x_1[T_1 - x_1]_\downarrow \cdots y \\
\text{(i}'\text{)} & w^* = x_1[T_1 - x_1]_\downarrow \\
\text{(ii)} & w^* = s[T_1]_\downarrow \cdots \\
\text{(iii)} & w^* = x_j[T_j - x_j]_\downarrow\, b[T_1 - b]_\downarrow \cdots \\
\text{(iv)} & w^* = t[T_k]_\downarrow\, b[T_1 - b]_\downarrow \cdots
\end{array}
$$

Step (2b) of Algorithm B′ computes $E(w^*) = (T_0, (T_1, x_1), \ldots, (T_n, x_n))$. In cases (i), (i′), and (iii) we get $T_0 = \emptyset$, and one can check that Step (3) gives $\tilde{E}(\gamma) = \Sigma$, as desired. In cases (ii) and (iv) we get $T_0 = \{s\}$ and $T_0 = \{t\}$, respectively. Similar reasoning gives $\tilde{E}(\gamma) = \Sigma$ in each of these cases as well, completing the argument that $\tilde{D}$ is injective. Surjectivity of $\tilde{D}$ now follows by a cardinality argument similar to that of Corollary 4.8, and the proof of Theorem 8.4 is finished.

9 Final remarks

It has occurred to us that formulae such as (3), (23), (24), (26), and (29) may have interpretations in a more algebraic setting, e.g., as Poincaré series of certain rings or modules. In that case, combinatorial constructions such as Algorithms A–D, A', B', and the mappings in Section 7 might also have algebraic significance.

Although we have not established precise connections of this sort, several colleagues (including Victor Reiner and Takayuki Hibi) have drawn our attention to papers such as [7], where crown-free posets appear, [13], where chordal graphs play a central role, and [19], where formula (24) appears explicitly as a Poincaré series. It seems that Koszul algebras feature prominently on the algebraic side (see, e.g., [1], [3], [14], [19]). We suggest that this might be an interesting area for further investigation.

Acknowledgements

This work was supported by research grants from NSERC and NSF.

References

1. L. L. Avramov, D. Eisenbud, Regularity of modules over a Koszul algebra, J. Algebra **153** (1992), 85-90.

2. C. Berge, *Graphs and Hypergraphs*, North-Holland, Amsterdam, 1973.

3. W. Bruns, J. Herzog, *Cohen-Macaulay rings*, Cambridge Studies in Appl. Math., Cambridge University Press, 1998.

4. L. Carlitz, R. Scoville, T. Vaughan, Enumeration of pairs of sequences by rises, falls, and levels, Manuscripta Math. **19** (1976), 211–243.

5. P. Cartier, D. Foata, Problėms combinatoire de commutation et réarrangements, Lecture Notes in Math **85** (1969), Springer-Verlag, Berlin.

6. L. Comtet, *Advanced Combinatorics*, D. Reidel, Dordrecht, 1974.

7. M. Božek, K. Drechsler, Singularities on flag spaces, Math Slovaca **49** (1999), 503–514.

8. D. Foata, Inversions de Möbius, unpublished notes, Science Research Council meeting, Aberdeen, 1975.

9. I. Gessel, Generating functions and enumeration of sequences, Ph.D. Thesis, M.I.T., 1977.

10. M. Golumbic, *Perfect Graphs and Algorithms*, Academic Press, New York, 1980.

11. I. P. Goulden and D. M. Jackson, *Combinatorial Enumeration*, Wiley, New York, 1983.

12. L. J. Guibas and A. M. Odlyzko, String overlaps, pattern matching, and nontransitive games, J. Comb. Theory A **30** (1981), 183–208.

13. J. Herzog, T. Hibi, and X. Zheng, Dirac's theorem on chordal graphs and Alexander duality, preprint, 2003.

14. J. Herzog, V. Reiner, V. Welker, The Koszul property in the affine semigroup rings, Pacific J. Math. **186** (1998), 39-65.

15. D. Kelly, I. Rival, Crowns, fences, and dismantlable lattices, Canad. J. Math. **26** (1974), 1257–1271.

16. H.A. Kierstead, W.T. Trotter and J. Qin, The dimension of cycle-free orders, Order **9** (1992), 103–110.

17. D. E. Knuth, The Art of Computer Programming, III, Addison-Wesley, Reading, 1973.

18. T.-H. Ma and J.P. Spinrad, Cycle-free partial orders and chordal comparability graphs, Order **8** (1991), 49–61.

19. I. Peeva, V. Reiner, B. Sturmfels, How to shell a monoid, Math. Ann. **310** (1998), 379–393.

20. P. A. MacMahon, Combinatory Analysis, Chelsea, New York, 1960.

21. A. Postnikov, Intransitive Trees, Jour. Combinatorial Theory A **79** (1997), 360-366.

22. A. Postnikov and R. P. Stanley, Deformations of Coxeter Hyperplane Arrangements, Jour. Combinatorial Theory A **91** (2000), 544–597.

23. N.V. Smirnov, O.V. Sarmanov and V.K. Zaharov, A local limit theorem for transition numbers in a Markov chain, and its applications, Sovi. Math. Dokl. **7** (1966), 563–566.

24. R.P. Stanley, *Enumerative combinatorics, Volume 1*, Wadsworth and Brooks/Cole, Monterey, 1986.

25. R.P. Stanley, A symmetric function generalization of the chromatic polynomial of a graph, Advances in Math. **111** (1995), 166–194.

26. W. T. Trotter, *Combinatorics and partially ordered sets: Dimension theory*, Johns Hopkins University Press, Baltimore, 1992.