Rationality, irrationality, and Wilf equivalence in generalized factor order

Sergey Kitaev*

The Mathematics Institute
School of Computer Science
Reykjavík University
IS-103 Reykjavík, Iceland
sergey@ru.is

Jeffrey Liese

Department of Mathematics
California Polytechnic State University
San Luis Obispo, CA 93407-0403, USA
jliese@calpoly.edu

Jeffrey Remmel†

Department of Mathematics
University of California, San Diego
La Jolla, CA 92093-0112, USA
remmel@math.ucsd.edu

Bruce E. Sagan‡

Department of Mathematics
Michigan State University
East Lansing, MI 48824-1027, USA
sagan@math.msu.edu

Submitted: Jun 1, 2008; Accepted: Nov 18, 2009; Published: Dec 2, 2009

Mathematics Subject Classifications: 05A15, 68R15, 06A07

Keywords: composition, factor order, finite state automaton, generating function, partially ordered set, rationality, transfer matrix, Wilf equivalence

Dedicated to Anders Björner on the occasion of his 60th birthday. His work has very heavily influenced ours.

Abstract

Let P be a partially ordered set and consider the free monoid P^* of all words over P. If $w, w' \in P^*$ then w' is a factor of w if there are words u, v with $w = uw'v$. Define generalized factor order on P^* by letting $u \leqslant w$ if there is a factor w' of w having the same length as u such that $u \leqslant w'$, where the comparison of u and w' is done componentwise using the partial order in P. One obtains ordinary factor order by insisting that $u = w'$ or, equivalently, by taking P to be an antichain.

Given $u \in P^*$, we prove that the language $\mathcal{F}(u) = \{w \ : \ w \geqslant u\}$ is accepted by a finite state automaton. If P is finite then it follows that the generating function $F(u) = \sum_{w \geqslant u} w$ is rational. This is an analogue of a theorem of Björner and Sagan for generalized subword order.

*The work presented here was supported by the Icelandic Research Fund, grant no. 090038011.

†Partially supported by NSF grant DMS 0654060

‡Work partially done while a Program Officer at NSF. The views expressed are not necessarily those of the NSF.

We also consider $P = \mathbb{P}$, the positive integers with the usual total order, so that P^* is the set of compositions. In this case one obtains a weight generating function $F(u; t, x)$ by substituting tx^n each time $n \in \mathbb{P}$ appears in $F(u)$. We show that this generating function is also rational by using the transfer-matrix method. Words u, v are said to be Wilf equivalent if $F(u; t, x) = F(v; t, x)$ and we prove various Wilf equivalences combinatorially.

Björner found a recursive formula for the Möbius function of ordinary factor order on P^*. It follows that one always has $\mu(u, w) = 0, \pm 1$. Using the Pumping Lemma we show that the generating function $M(u) = \sum_{w \geqslant u} |\mu(u, w)| w$ can be irrational.

1 Introduction and definitions

Let P be a set and consider the corresponding *free monoid* or *Kleene closure* of all words over P:

$$P^* = \{w = w_1 w_2 \dots w_\ell \ : \ n \geqslant 0 \text{ and } w_i \in P \text{ for all } i\}.$$

Let ϵ be the empty word and for any $w \in P^*$ we denote its cardinality or *length* by $|w|$. Given $w, w' \in P^*$, we say that w' is a *factor* of w if there are words u, v with $w = uw'v$, where adjacency denotes concatenation. For example, $w' = 322$ is a factor of $w = 12213221$ starting with the fifth element of w. *Factor order* on P^* is the partial order obtained by letting $u \leqslant_{\text{fo}} w$ if and only if there is a factor w' of w with $u = w'$.

Now suppose that we have a poset $(P, \leqslant)$. We define *generalized factor order* on P^* by letting $u \leqslant_{\text{gfo}} w$ if there is a factor w' of w such that

(a) $|u| = |w'|$, and

(b) $u_i \leqslant w'_i$ for $1 \leqslant i \leqslant |u|$.

We call w' an *embedding* of u into w, and if the first element of w' is the jth element of w, we call j an *embedding index* of u into w. We also say that, in this embedding, u_i is in *position* $j + i - 1$. To illustrate, suppose $P = \mathbb{P}$, the positive integers with the usual order relation. If $u = 322$ and $w = 12213431$ then $u \leqslant_{\text{gfo}} w$ because of the embedding factor $w' = 343$ which has embedding index 5, and the two 2's of u are in positions 6 and 7. Note that we obtain ordinary factor order by taking P to be an antichain. Also, we will henceforth drop the subscript gfo since context will make it clear what order relation is meant. Generalized factor order is the focus of this paper.

Returning to the case where P is an arbitrary set, let $\mathbb{Z}\langle\langle P \rangle\rangle$ be the algebra of formal power series with integer coefficients and having the elements of P as noncommuting variables. In other words,

$$\mathbb{Z}\langle\langle P \rangle\rangle = \left\{ f = \sum_{w \in P^*} c(w) w \ : \ c(w) \in \mathbb{Z} \text{ for all } w \right\}.$$

If $f \in \mathbb{Z}\langle\langle P \rangle\rangle$ has no constant term, i.e., $c_\epsilon = 0$, then define

$$f^* = \epsilon + f + f^2 + f^3 + \cdots = (\epsilon - f)^{-1}.$$

(We need the restriction on f to make sure that the sums are well defined as formal power series.) We say that f is *rational* if it can be constructed from the elements of P using only a finite number of applications of the algebra operations and the star operation.

A *language* is any $\mathcal{L} \subseteq P^*$. It has an associated generating function

$$f_{\mathcal{L}} = \sum_{w \in \mathcal{L}} w.$$

The language $\mathcal{L}$ is *regular* if $f_{\mathcal{L}}$ is rational.

Consider generalized factor order on P^* and fix a word $u \in P^*$. There is a corresponding language and generating function

$$\mathcal{F}(u) = \{w \ : \ w \geqslant u\} \quad \text{and} \quad F(u) = \sum_{w \geqslant u} w.$$

We begin with the following result.

Proposition 1.1. *If P is a finite poset and $u \in P^*$ then $F(u)$ is rational.*

Proof. It is easy to see directly from the definitions that P^*wP^* is a regular language for any $w \in P^*$. Also,

$$\mathcal{F}(u) = \cup_w P^*wP^*$$

where the union is over all $w \geqslant u$ with $|w| = |u|$. Since P is finite, so is the union. And finite unions of regular languages are regular, so we are done. ☐

Proposition 1.1 is an analogue of a result of Björner and Sagan [5] for generalized subword order on P^*. *Generalized subword order* is defined exactly like generalized factor order except that w' is only required to be a subword of w, i.e., the elements of w' need not be consecutive in w. For related results, also see Goyt [6].

We are going to give a second proof of Proposition 1.1 using automata. There are two reasons for doing so. The first is that this approach will allow us to generalize Propostion 1.1 so that it applies to a large class of infinite posets, see Theorem 8.2. In particular, it will apply to the infinite poset $\mathbb{P}$ which will be the focus of much of the rest of the paper. The second is that the construction of the automaton will permit us to develop an algorithm to actually compute the series in question, not only for finite posets but also for various infinite posets as well.

Given any set, P, a *nondeterministic finite automaton* or *NFA* over P is a digraph (directed graph) Δ with vertices V and arcs $\vec{E}$ having the following properties.

1. The elements of V are called *states* and $|V|$ is finite.

2. There is a designated *initial state* α and a set Ω of *final states*.

3. Each arc of $\vec{E}$ is labeled with an element of P.

Given a (directed) path in Δ starting at α, we construct a word in P^* by concatenating the elements on the arcs on the path in the order in which they are encountered. The *language accepted by* Δ is the set of all such words which are associated with paths ending in a final state. It is a well-known theorem that, for $|P|$ finite, a language $\mathcal{L} \subseteq P^*$ is regular if and only if there is a NFA accepting $\mathcal{L}$. This result is well-known and follows from the work of Kleene [8] (or see the book of Berstel and Reutenauer [2, page 37]). It was later generalized by Schützenberger [9].

We will reprove Proposition 1.1 by constructing a NFA accepting the language for $F(u)$. This will be done in the next section. In fact, the NFA still exists even if P is infinite, and we will use this fact to prove that $F(u)$ is also rational for certain infinite posets.

We are particularly interested in the case of $P = \mathbb{P}$ with the usual order relation. So $\mathbb{P}^*$ is just the set of *compositions* (ordered integer partitions). Given $w = w_1 w_2 \ldots w_\ell \in \mathbb{P}^*$, we define its *norm* to be

$$\Sigma(w) = w_1 + w_2 + \cdots + w_\ell.$$

Let t, x be commuting variables. Replacing each $n \in w$ by tx^n we get an associated monomial called the *weight* of w

$$\mathrm{wt}(w) = t^{|w|} x^{\Sigma(w)}.$$

For example, if $w = 213221$ then

$$\mathrm{wt}(w) = tx^2 \cdot tx \cdot tx^3 \cdot tx^2 \cdot tx^2 \cdot tx = t^6 x^{11}.$$

We also have the associated *weight generating function*

$$F(u; t, x) = \sum_{w \geqslant u} \mathrm{wt}(w).$$

Our NFA will demonstrate, via the transfer-matrix method, that this is also a rational function of t and x. The details will be given in Section 3.

Call $u, w \in \mathbb{P}^*$ *Wilf equivalent* if $F(u; t, x) = F(v; t, x)$. This definition is inspired by the one used in the theory of pattern avoidance, but is different since our partial order is not pattern containment. See the survey article of Wilf [11] for more information about this subject. Section 4 is devoted to proving various Wilf equivalences. Although these results were discovered by having a computer construct the corresponding generating functions, the proofs we give are purely combinatorial. In the next two sections, we investigate a stronger notion of equivalence and compute generating functions for two families of compositions.

Björner [3] gave a recursive formula for the Möbius function of (ordinary) factor order. It follows from his theorem that $\mu(u, w) = 0, \pm 1$ for all $u, w \in P^*$. Using the Pumping Lemma [7, Lemma 3.1] we show that there are finite sets P and $u \in P^*$ such that the language

$$\mathcal{M}(u) = \{w \ : \ \mu(u, w) \neq 0\}$$

is not regular. This is done in Section 7. The penultimate section is devoted to comments, conjectures, and open questions. And the final one contains tables.

2 Construction of automata

We will now introduce two other languages which are related to $\mathcal{F}(u)$ and which will be useful in our automaton proof of Proposition 1.1 and its extensions, as well as in demonstrating Wilf equivalence. We say that u is a *suffix* (respectively, *prefix*) of w if $w = vu$ (respectively, $w = uv$) for some word v. Let $\mathcal{S}(u)$ be all the $w \in \mathcal{F}(u)$ such that, in the definition of generalized factor order, the only possible choice for w' is a suffix of w. Let $S(u)$ be the corresponding generating function.

We say that $w \in P^*$ *avoids* u if $w \not\geqslant u$ in generalized factor order. Let $\mathcal{A}(u)$ be the associated language with generating function $A(u)$. The next result follows easily from the definitions and so we omit the proof. In it, we will use the notation Q to stand both for a subset of P and for the generating function $Q = \sum_{a \in Q} a$. Context will make it clear which is meant.

Lemma 2.1. *Let P be any poset and let $u \in P^*$. Then we have the following relationships:*

1. $\mathcal{F}(u) = \mathcal{S}(u)P^*$ *and* $F(u) = S(u)(\epsilon - P)^{-1}$,
2. $\mathcal{A}(u) = P^* - \mathcal{F}(u)$ *and* $A(u) = (\epsilon - P)^{-1} - F(u)$. □

We will now prove that all three of the languages we have defined are accepted by NFAs. An example follows the proof so the reader may want to read it in parallel.

Theorem 2.2. *Let P be any poset and let $u \in P^*$. Then there are NFAs accepting $\mathcal{F}(u)$, $\mathcal{S}(u)$, and $\mathcal{A}(u)$.*

Proof. We first construct an NFA, Δ, for $\mathcal{S}(u)$. Let $\ell = |u|$. The states of Δ will be all subsets T of $\{1, \ldots, \ell\}$. The initial state is $\emptyset$. The elements of T will be the lengths of prefixes of u which embedd as a suffix of a word corresponding to a path from $\emptyset$ to T. Thus the final states will be all T which contain ℓ. More precisely, let $w = w_1 \ldots w_m$ be the word corresponding to a path from $\emptyset$ to T. Then we want the only possible embedding indices to be those in the set $\{m - t + 1 \ : \ t \in T\}$. In other words, for each $t \in T$ we have

$$u_1 u_2 \ldots u_t \leqslant w_{m-t+1} w_{m-t+2} \ldots w_m, \tag{1}$$

and for each $t \in \{1, \ldots, \ell\} - T$ this inequality does not hold, and $u \not\leqslant w'$ for any factor w' of w starting at an index smaller then $m - \ell + 1$.

We now need to define the arcs of Δ in such a manner that if a path to T is continued to T' then (1) will still hold. There will be no arcs out of a final state. If T is a nonfinal state and $a \in P$ then there will be an arc from T to

$$T' = \{t + 1 \ : \ t \in T \cup \{0\} \text{ and } u_{t+1} \leqslant a\}.$$

It is easy to see that (1) continues to hold for all $t' \in T'$ once we append a to w. This finishes the construction of the NFA for $\mathcal{S}(u)$.

To obtain an automaton for $\mathcal{F}(u)$, just add loops to the final states of Δ, one for each $a \in P$. An automaton for $\mathcal{A}(u)$ is obtained by just interchanging the final and nonfinal states in the automaton for $\mathcal{F}(u)$. This is because the additional arcs in $\mathcal{F}(u)$ make it deterministic. $\square$

As an example, consider $P = \mathbb{P}$ and $u = 132$. We will do several things to simplify writing down the automaton. First of all, certain states may not be reachable by a path starting at the initial state. So we will not display such states. For example, we can not reach the state $\{2, 3\}$ since $u_1 = 1 \leqslant w_i$ for any i and so 1 will be in any state reachable from $\emptyset$. Also, given states T and U there may be many arcs from T to U, each having a different label. So we will replace them by one arc bearing the set of labels of all such arcs. Finally, set braces will be dropped for readability. The resulting digraph is displayed in Figure 1.

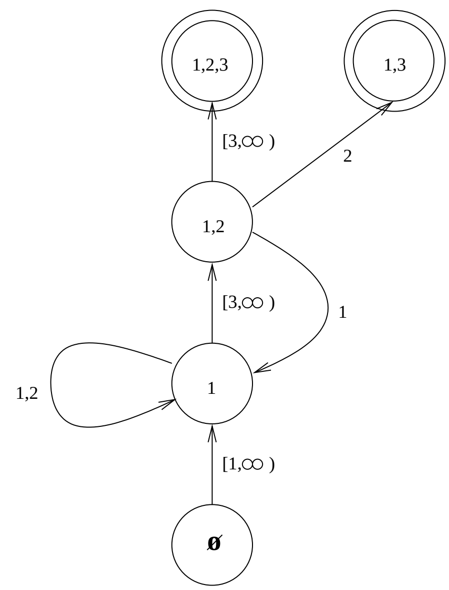

Figure 1: A NFA accepting $\mathcal{S}(132)$

Consider what happens as we build a word w starting from the initial state $\emptyset$. Since $u_1 = 1$, any element of $\mathbb{P}$ could be the first element of an embedding of u into w. That is why every element of the interval $[1, \infty) = \mathbb{P}$ produces an arrow from the initial state to the state $\{1\}$. Now if $w_2 \leqslant 2$, then an embedding of u could no longer start at w_1 and so these elements give loops at the state $\{1\}$. But if $w_2 \geqslant 3$ then an embedding could start at either w_1 or at w_2 and so the corresponding arcs all go to the state $\{1, 2\}$. The rest of the automaton is explained similarly.

As an immediate consequence of the previous theorem we get the following result which includes Proposition 1.1.

Theorem 2.3. *Let P be a finite poset and let $u \in P^*$. Then the generating functions $F(u)$, $S(u)$, and $A(u)$ are all rational.* $\square$

3 The positive integers

If $P = \mathbb{P}$ then Theorem 2.3 no longer applies to the generating functions $F(u)$, $S(u)$, and $A(u)$. However, we can still show rationality of the weight generating function $F(u;t,x)$ as defined in the introduction. Similarly, we will see that the series

$$S(u;t,x) = \sum_{w \in \mathcal{S}(u)} \mathrm{wt}(w) \quad \text{and} \quad A(u;t,x) = \sum_{w \in \mathcal{A}(u)} \mathrm{wt}(w)$$

are rational.

Note first that Lemma 2.1 still holds for $\mathbb{P}$ and can be made more explicit in this case. Extend the function wt to all of $\mathbb{Z}\langle\langle \mathbb{P} \rangle\rangle$ by letting it act linearly. Then

$$\begin{aligned} \mathrm{wt}(\epsilon - \mathbb{P})^{-1} &= \frac{1}{1 - \sum_{n \geqslant 1} tx^n} \\ &= \frac{1}{1 - tx/(1-x)} \\ &= \frac{1-x}{1-x-tx}. \end{aligned}$$

We now plug this into the lemma.

Corollary 3.1. *We have*

1. $F(u;t,x) = \dfrac{(1-x)S(u;t,x)}{1-x-tx}$ *and*

2. $A(u;t,x) = \dfrac{1-x}{1-x-tx} - F(u;t,x).$ □

It follows that if any one of these three series is rational then the other two are as well.

We will now use the NFA, Δ, constructed in Theorem 2.2 to show that $S(u;t,x)$ is rational. This is essentially an application of the transfer-matrix method. See the text of Stanley [10, Section 4.7] for more information about this technique. The *transfer matrix* M for Δ has rows and columns indexed by the states with

$$M_{T,U} = \sum_n \mathrm{wt}(n)$$

where the sum is over all n which appear as labels on the arcs from T to U. For example, consider the case where $w = 132$ as done at the end of the previous section. If we list the states in the order

$$\emptyset,\ \{1\},\ \{1,2\},\ \{1,3\},\ \{1,2,3\}$$

then the transfer matrix is

$$M = \begin{bmatrix} 0 & \dfrac{tx}{1-x} & 0 & 0 & 0 \\ 0 & t(x+x^2) & \dfrac{tx^3}{1-x} & 0 & 0 \\ 0 & tx & 0 & tx^2 & \dfrac{tx^3}{1-x} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Now M^k has entries $M^k_{T,U} = \sum_w \mathrm{wt}(w)$ where the sum is over all words w corresponding to a directed walk of length k from T to U. So to get the weight generating function for walks of all lengths one considers $\sum_{k\geqslant 0} M^k$. Note that this sum converges in the algebra of matrices over the formal power series algebra $\mathbb{Z}[[t,x]]$ because none of the entries of M has a constant term. It follows that

$$L := \sum_{k\geqslant 0} M^k = (I-M)^{-1} = \frac{\mathrm{adj}(I-M)}{\det(I-M)} \tag{2}$$

where adj denotes the adjoint.

Now

$$S(u;t,x) = \sum_T L_{\emptyset,T}$$

where the sum is over all final states of Δ. So it suffices to show that each entry of L is rational. From equation (2), this reduces to showing that each entry of M is rational. So consider two given states T, U. If T is final then we are done since the Tth row of M is all zeros. If T is not final, then consider

$$T' = \{t+1 \ : \ t \in T \cup \{0\}\}. \tag{3}$$

If $U = T'$ then there will be an $N \in \mathbb{P}$ such that all the arcs out of T with labels $n \geqslant N$ go to T'. So $M_{T,T'}$ will contain $\sum_{n\geqslant N} tx^n = tx^N/(1-x)$ plus a finite number of other terms of the form tx^m. Thus this entry is rational. If $U \neq T'$, then there will only be a finite number of arcs from T to U and so $M_{T,U}$ will actually be a polynomial. This shows that every entry of M is rational and we have proved, with the aid of the remark following Corollary 3.1, the following result.

Theorem 3.2. *If $u \in \mathbb{P}^*$ then $F(u;t,x)$, $S(u;t,x)$, and $A(u;t,x)$ are all rational.* □

4 Wilf equivalence

Recall that $u, v \in \mathbb{P}^*$ are *Wilf equivalent*, written $u \sim v$, if $F(u;t,x) = F(v;t,x)$. By Corollary 3.1, this is equivalent to $S(u;t,x) = S(v;t,x)$ and to $A(u;t,x) = A(v;t,x)$.

It follows that to prove Wilf equivalence, it suffices to find a weight-preserving bijection $f : \mathcal{L}(u) \to \mathcal{L}(v)$ where $\mathcal{L} = \mathcal{F}$, $\mathcal{S}$, or $\mathcal{A}$. Since $\sim$ is an equivalence relation, we can talk about the *Wilf equivalence class* of u which is $\{w \ : \ w \sim u\}$. It is worth noting that the automata for the words in a Wilf equivalence class need not bear a resemblance to each other.

Part of the motivation for this section is to try to explain as many Wilf equivalences as possible between permutations. For reference, in Section 9 the first table lists all such equivalences up through 5 elements.

First of all, we consider three operations on words in $\mathbb{P}^*$. The *reversal* of $u = u_1 \ldots u_\ell$ is $u^r = u_\ell \ldots u_1$. It will also be of interest to consider $1u$, the word gotten by prepending one to u. Finally, we will look at u^+ which is gotten by increasing each element of u by one, as well as u^- which performs the inverse operation whenever it is defined. For some of our proofs, it will also be useful to have the following factorization. Given $k \in \mathbb{P}$ and $w \in \mathbb{P}^*$ the *k-factorization* of w is the unique expression

$$w = y_1 \ z_1 \ y_2 \ z_2 \ \ldots \ z_{m-1} \ y_m$$

where $y_i \in [1, k)^*$ and $z_i \in [k, \infty)^*$ for all i, and all factors are nonempty with the possible exception of y_1 and y_m.

Lemma 4.1. *We have the following Wilf equivalences.*

(a) $u \sim u^r$,

(b) if $u \sim v$ *then* $1u \sim 1v$,

(c) if $u \sim v$ *then* $u^+ \sim v^+$.

Proof. (a) It is easy to see that the map $w \mapsto w^r$ is a weight-preserving bijection $\mathcal{F}(u) \to \mathcal{F}(u^r)$.

(b) We will show that $A(1u; t, x) = A(1v; t, x)$. Consider $w \in \mathcal{A}(1u)$. Then either u does not embed in w, or it embeds in w exactly once and that is as a prefix of w. It follows that

$$\mathcal{A}(1u) = \mathcal{A}(u) \uplus \{w^r \ : \ w \in \mathcal{S}(u^r)\}.$$

Translating this into generating functions yields

$$A(1u; t, x) = A(u; t, x) + S(u^r; t, x).$$

But the same argument shows that

$$A(1v; t, x) = A(v; t, x) + S(v^r; t, x).$$

Since $u \sim v$ we have $A(u; t, x) = A(v; t, x)$, and from part (a) we have $S(u^r; t, x) = S(v^r; t, x)$. Thus $A(1u; t, x) = A(1v; t, x)$ as desired.

(c) Now we consider a weight-preserving bijection $g : \mathcal{A}(u) \to \mathcal{A}(v)$. Given $w \in \mathbb{P}^*$, let

$$w = y_1 \ z_1 \ y_2 \ z_2 \ \ldots \ z_{m-1} \ y_m$$

be its 2-factorization. Since all elements of u^+ are at least two, $w \in \mathcal{A}(u^+)$ if and only if $z_i \in \mathcal{A}(u^+)$ for all i. This is equivalent to $z_i^- \in \mathcal{A}(u)$ for all i. Thus if we map w to

$$y_1 \; g(z_1^-)^+ \; y_2 \; g(z_2^-)^+ \; \ldots \; g(z_{m-1}^-)^+ \; y_m$$

then we will get the desired weight-preserving bijection $\mathcal{A}(u^+) \to \mathcal{A}(v^+)$. $\square$

We can combine these three operations to prove more complicated Wilf equivalences. Since a word $w \in \mathbb{P}^*$ is just a sequence of positive integers, terms like “weakly increasing” and “maximum” have their usual meanings. Also, let w^{+m} be the result of applying the $+$ operator m times. By using the previous lemma and induction, we obtain the following result. The proof is so straight forward that it is omitted.

Corollary 4.2. *Let y, y' be weakly increasing compositions and z, z' be weakly decreasing compositions such that yz is a rearrangement of $y'z'$. Then for any $u \sim v$ we have*

$$yu^{+m}z \sim y'v^{+m}z'$$

whenever $m \geqslant \max\{y, z\} - 1$. $\square$

Applying the two previous results, we can obtain the Wilf equivalences in the symmetric group $\mathfrak{S}_3$ of all the permutations of $\{1, 2, 3\}$:

$$123 \sim 321 \sim 132 \sim 231 \quad \text{and} \quad 213 \sim 312.$$

These two groups are indeed in different equivalence classes as one can use equation (2) to compute that

$$S(123; t, x) = \frac{t^3x^6}{(1-x)^2(1-x-tx+tx^3-t^2x^4)}$$

while

$$S(213; t, x) = \frac{t^3x^6(1+tx^3)}{(1-x)(1-x+t^2x^4)(1-x-tx+tx^3-t^2x^4)}.$$

However, we will need a new result to explain some of the equivalences in $\mathfrak{S}_4$ such as $2134 \sim 2143$. Let u be a composition such that $\max u$ only occurs once. Define a *pseudo-embedding* of u into w to be a factor w' of w satisfying the two conditions for an embedding except that the inequality may fail at the position(s) of $\max u$. In particular, embeddings are pseudo-embeddings.

An example of the construction used in the next theorem follows the proof and can be read in parallel.

Theorem 4.3. *Let $x, y, z \in \{1, \ldots, m\}^*$ and suppose $n > m$. Then*

$$xmynz \sim xnymz.$$

Proof. Let $u = xmynz$ and $v = xnymz$. We will construct a weight-preserving bijection $\mathcal{A}(u) \to \mathcal{A}(v)$. To do this, it suffices to construct such a bijection between the set differences $\mathcal{A}(u) - \mathcal{A}(v) \to \mathcal{A}(v) - \mathcal{A}(u)$ since the identity map can be used on $\mathcal{A}(u) \cap \mathcal{A}(v)$.

Given $w \in \mathcal{A}(u) - \mathcal{A}(v)$, consider the set

$$\eta(w) = \{i \ : \ \text{there is an embedding of } v \text{ into } w \text{ with the } n \text{ in position } i\}.$$

For such i, $w_i \geqslant n$. It must also be that w_{i+k} is in the interval $[m, n)$ where $k = |y| + 1$: Certainly $w_{i+k} \geqslant m$ because of the embedding. But if $w_{i+k} \geqslant n$ then there would also be an embedding of u at the same position as the one for v, contradicting $w \in \mathcal{A}(u)$.

Now for each $i \in \eta(w)$ we define the *sequence* beginning at i as

$$\sigma(i) = \{i,\ i+k,\ i+2k,\ \ldots,\ i+\ell k\}$$

where ℓ is the least nonnegative integer such that there is no pseudo-embedding of v into w with the n in position $i+\ell k$. Note that ℓ depends on i even though this is not reflected in our notation. Also, $\ell \geqslant 1$ since there is embedding of v into w with the n in position i. Finally, it is easy to see that $w_{i+k}, w_{i+2k}, \ldots, w_{i+\ell k} \in [m, n)$ by an argument similar to that for w_{i+k}. This implies that any two sequences are disjoint since $w_i \geqslant n$ for $i \in \eta(w)$.

Now map w to $\bar{w}$ which is constructed by switching the values of w_i and $w_{i+\ell k}$ for every $i \in \eta(w)$. Since sequences are disjoint, the switchings are well defined. We must show that $\bar{w} \in \mathcal{A}(v) - \mathcal{A}(u)$. We prove that $\bar{w} \in \mathcal{A}(v)$ by contradiction. The switching operation removes every embedding of v in w. If a new embedding was created then, because only elements of size at least m move, the n in v must correspond to $\bar{w}_{i+\ell k}$ for some $i \in \eta(w)$. But now there is a pseudo-embedding of v into w with the n in position $i + \ell k$, contradicting the definition of ℓ.

To show $\bar{w} \notin \mathcal{A}(u)$, we will actually prove the stronger statement that there is an embedding of u in $\bar{w}$ with the n in position $i + \ell k$ for each $i \in \eta(w)$ and these are the only embeddings. These embeddings exist because there is a pseudo-embedding of v into w with the n in position $i+(\ell-1)k$, $\bar{w}_{i+\ell k} \geqslant n$, and only elements of size at least m move in passing from w to $\bar{w}$. They are the only ones because $w \in \mathcal{A}(u)$ and so any embedding of u in $\bar{w}$ would have to have the n in a position of the form $i + \ell k$.

Finally, we need to show that this map is bijective. But modifying the above construction by exchanging the roles of u and v and building the sequences from right to left gives an inverse. This completes the proof. $\square$

By way of illustration, suppose $u = 1\ 3\ 5\ 2\ 4\ 6\ 3$ and $v = 1\ 3\ 6\ 2\ 4\ 5\ 3$ so that $m = 5$, $n = 6$, and $k = 3$. We will write our example w in two line form with the upper line being the positions:

$$w = \begin{array}{cccccccccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 \\ 1 & 1 & 2 & 4 & 8 & 3 & 9 & 5 & 4 & 5 & 5 & 4 & 5 & 5 & 3 & 3 & 3 & 6 & 6 & 5 & 5 & 3. \end{array}$$

Now there are three embeddings of v (and none of u) into w with the 6 in positions $\eta(w) = \{5, 7, 18\}$. For $i = 5$ we have the sequence $\sigma(5) = \{5, 8, 11, 14\}$ since there are pseudo-embeddings of v with the n in positions $5, 8, 11$ but not in position 14. Similarly $\sigma(7) = \{7, 10, 13\}$ and $\sigma(18) = \{18, 21\}$. So $\bar{w}$ is obtained by switching w_5 with w_{14}, w_7 with w_{13}, and w_{18} with w_{21} to obtain

$$\bar{w} = \begin{array}{cccccccccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 \\ 1 & 1 & 2 & 4 & 5 & 3 & 5 & 5 & 4 & 5 & 5 & 4 & 9 & 8 & 3 & 3 & 3 & 5 & 6 & 5 & 6 & 3. \end{array}$$

It is now easy to verify that our results so far suffice to explain all the Wilf equivalences in symmetric groups up through $\mathfrak{S}_4$. They also explain most, but not all, of the ones in $\mathfrak{S}_5$. We will return to the $n = 5$ case in the section on open questions.

One might wonder about the necessity of the requirement that the two equivalent words in Theorem 4.3 have a unique maximum. However, one can see from Table 2 in Section 9 that 122 and 212 are not Wilf equivalent. So if there is an analogue of this theorem for more general words, another condition will have to be imposed.

One might also hope that it would be possible to do without the sequences in the proof and merely switch w_i and w_{i+k} for all $i \in \eta(w)$ to get $\bar{w}$. This would only be invertible if the embedding indices for v in w would be the same as those for u in $\bar{w}$. Unfortunately, this does not always work as the following example shows. Consider $u = 231$, $v = 321$, and all w which are permutations of 1223. Then the members of $\mathcal{A}(u) - \mathcal{A}(v)$ are 1322, 3212, and 3221; while those of $\mathcal{A}(v) - \mathcal{A}(u)$ are 1232, 2313, and 2231. The embedding indices of v in the first three compositions are 2, 1, and 1 (respectively); while those of u in the second three are 2, 1, and 2. Thus preservation of the indices is not possible in this case. However, it would be interesting to know when one can leave the indices invariant and this will be investigated in the next section.

The reader may have noted that a number of the maps constructed in proving the results of this section involve rearrangement of the letters of the word (which makes the map automatically weight preserving). We will now show that if one strengthens the hypothesis of Lemma 4.1 (c) by adding a rearrangement assumption, then one can also strengthen the conclusion by applying any strictly increasing function to u and v. To state and prove this result, we first need some definitions.

Say that a map $f : P^* \to P^*$ is a *rearrangement* if $f(w)$ is a rearrangement of w for all $w \in P^*$. Now let $u, v \in \mathbb{P}^*$ be given. If $f : \mathbb{P}^* \to \mathbb{P}^*$ is a weight-preserving bijection such that, for all $w \in \mathbb{P}^*$,

$$u \leqslant w \iff v \leqslant f(w) \tag{4}$$

then we say that f *witnesses* the Wilf equivalence $u \sim v$.

Given any function $\iota : \mathbb{P} \to \mathbb{P}$ we extend ι to $\mathbb{P}^*$ by letting

$$\iota(u_1 u_2 \ldots u_n) = \iota(u_1)\iota(u_2)\ldots\iota(u_n).$$

Now assume that ι is strictly increasing on $\mathbb{P}$ with range $\{k_1 < k_2 < \ldots\}$. Given a word $w = w_1 \ldots w_m$ in $(\mathbb{P} - [1, k_1))^*$ we form its *collapse*, $\operatorname{clp}(w)$, by replacing each letter of w in the interval $[k_j, k_{j+1})$ by j for all $j \in \mathbb{P}$. For example, if $\iota(1) = 3$, $\iota(2) = 5$, $\iota(3) = 8$, and $\iota(4) = 13$ then $\operatorname{clp}(356749438) = 122213113$. For any $u, w \in \mathbb{P}^*$, we have

$$\iota(u) \leqslant w \iff u \leqslant \operatorname{clp}(w). \tag{5}$$

We now have everything in place for proof of the next result which resembles the proof of Lemma 4.1 (c).

Theorem 4.4. *Suppose $u, v \in \mathbb{P}^*$ such that there is a rearrangement $f : \mathbb{P}^* \to \mathbb{P}^*$ witnessing $u \sim v$. Then for any strictly increasing function $\iota : \mathbb{P} \to \mathbb{P}$ there is a rearrangement $g : \mathbb{P}^* \to \mathbb{P}^*$ witnessing $\iota(u) \sim \iota(v)$.*

Proof. It suffices to construct a bijective rearrangement g satisfying (4) since then it must also be weight preserving. Given $w \in \mathbb{P}^*$, let

$$w = y_1\ z_1\ y_2\ z_2\ \ldots\ z_{m-1} y_m$$

be its k_1-factorization where $k_1 = \iota(1)$. Clearly $\iota(u) \leqslant w$ if and only if $\iota(u) \leqslant z_i$ for some i. For each i, define

$$z_i' = f(\mathrm{clp}(z_i)).$$

By our assumptions and (5) we have

$$\iota(u) \leqslant z_i \iff u \leqslant \mathrm{clp}(z_i) \iff v \leqslant z_i'.$$

Now fix $j \geqslant 1$ and let $z_i(1) \ldots z_i(r_j)$ be the elements of z_i in $[k_j, k_{j+1})$, reading from left to right. These are the elements of z_i which get replaced by j when passing from z_i to $\mathrm{clp}(z_i)$. Since $z_i' = f(\mathrm{clp}(z_i))$ is a rearrangement of $\mathrm{clp}(z_i)$, there must be r_j occurrences of j in z_i'. Replace these j's by $z_i(1) \ldots z_i(r_j)$, reading from left to right. Do this for each $j \in \mathbb{P}$ and call the result $g(z_i)$. Then $g(z_i)$ is a rearrangement of z_i and $\mathrm{clp}(g(z_i)) = z_i'$. It follows from (5) and the previous displayed equation that

$$\iota(v) \leqslant g(z_i) \iff v \leqslant z_i' \iff \iota(u) \leqslant z_i.$$

Now let

$$g(w) = y_1\ g(z_1)\ y_2\ g(z_2)\ \ldots\ g(z_{m-1}) y_m.$$

This map is a rearrangement by construction and satisfies (4) because of the last displayed equation in the previous paragraph. One can construct g^{-1} from f^{-1} in the same way that we constructed g from f. So we are done. $\square$

5 Strong Wilf equivalence

Given $v, w \in \mathbb{P}^*$ we let

$$\mathrm{Em}(v, w) = \{j\ :\ j \text{ is an embedding index of } v \text{ into } w\}.$$

Call compositions u, v *strongly Wilf equivalent*, written $u \sim_s v$, if there is a weight-preserving bijection $f : \mathbb{P}^* \to \mathbb{P}^*$ such that

$$\mathrm{Em}(u, w) = \mathrm{Em}(v, f(w)) \tag{6}$$

for all $w \in \mathbb{P}^*$. In this case we say that f *witnesses* the strong Wilf equivalence $u \sim_s v$. Clearly strong Wilf equivalence implies Wilf equivalence. In addition to being a natural notion, our interest in this concept is motivated by the fact that we were able to prove Theorem 5.3 below only under the assumption of strong Wilf equivalence, although we suspect it is true for ordinary Wilf equivalence. First, however, we will prove analogues of some of our results from the previous section in this setting.

Lemma 5.1. *If $u \sim_s v$ then*

(a) $1u \sim_s 1v$,

(b) $1u \sim_s v1$,

(c) $u^+ \sim_s v^+$.

Proof. Let $f : \mathbb{P}^* \to \mathbb{P}^*$ be a weight-preserving map satisfying (6). Define maps $g : \mathbb{P}^* \to \mathbb{P}^*$ and $h : \mathbb{P}^* \to \mathbb{P}^*$ by $g(\epsilon) = h(\epsilon) = \epsilon$ and, for $w = by$ with $b \in \mathbb{P}$,

$$g(by) = bf(y) \quad \text{and} \quad h(by) = f(y)b.$$

It follows easily that these functions establish (a) and (b). Finally, the construction used in the proof of (c) in Lemma 4.1 can be carried over to prove the analogous case here. That is, if one assumes that the function g given there also satisfies (6) then the derived map will demonstrate that $u^+ \sim_s v^+$. □

As before, we can combine the previous result and induction to get a more general equivalence.

Corollary 5.2. *Let y, y' be weakly increasing compositions and z, z' be weakly decreasing compositions such that yz is a rearrangement of $y'z'$. Then for any $u \sim_s v$ we have*

$$yu^{+m}z \sim_s y'v^{+m}z'$$

whenever $m \geqslant \max\{y, z\} - 1$. □

Not every Wilf equivalence is a strong Wilf equivalence. From Lemma 4.1 (a) we know that $w \sim w^r$. But we can show that $2143 \not\sim_s 3412$ as follows. Consider how one could construct a word w of length 7 with $\Sigma(w)$ minimum and $\mathrm{Em}(2143, w) = \{1, 3, 4\}$. Construct a table with a copy of 2143 starting in the first, third, and fourth positions in rows 1, 2, and 3, respectively. Then take the maximum value in each column for the corresponding entries of w:

$$\begin{array}{ccccccccc} & & 2 & 1 & 4 & 3 & & & \\ & & & & 2 & 1 & 4 & 3 & \\ & & & & & 2 & 1 & 4 & 3 \\ \hline w & = & 2 & 1 & 4 & 3 & 4 & 4 & 3. \end{array}$$

By construction, w has the desired embedding indices and one sees immediately that it has no others. Note that this is the unique w satisfying the given restrictions and that $\mathrm{wt}(w) = t^7x^{21}$. But applying the same process to 3412 gives $\bar{w} = 3434422$ with $\mathrm{wt}(\bar{w}) = t^7x^{22}$. Since the weights do not agree, we can not have strong Wilf equivalence.

Finally, we come to the result alluded to at the beginning of this section. Given $b \in \mathbb{P}$ we let b^k denote the composition consisting of k copies of b.

Theorem 5.3. *Suppose $u = u_1 \ldots u_n \sim_s v = v_1 \ldots v_n$. Then for any $k \in \mathbb{P}$*

$$u_1^k \ldots u_n^k \sim_s v_1^k \ldots v_n^k.$$

Proof. Let $f : \mathbb{P}^* \to \mathbb{P}^*$ be a map satisfying (6). Given any $w \in \mathbb{P}^*$ and i with $1 \leqslant i \leqslant k$, consider the subword $w[i] = w_i w_{i+k} w_{i+2k} \ldots$ of w. Then the embeddings of $u_1^k \ldots u_n^k$ in w are completely determined by the embeddings of u in the $w[i]$ and vice-versa. So replacing each subword $w[i]$ by the subword $f(w[i])$ yields the desired map. $\square$

Just as in the previous section, we can get an interesting result by imposing the rearrangement condition on maps. Here is an analogue of Corollary 5.2 in this setting without the weakly increasing assumption.

Theorem 5.4. *Fix $k \in \mathbb{P}$ and suppose $u, v \in [k, \infty)^*$ such that there is a rearrangement $f : \mathbb{P}^* \to \mathbb{P}^*$ witnessing $u \sim_s v$. Then for any two words $y, z \in [1, k]^*$ there is a rearrangement $g : \mathbb{P}^* \to \mathbb{P}^*$ witnessing $yuz \sim_s yvz$.*

Proof. It suffices to construct a bijective rearrangement g satisfying (6) since then it must also be weight preserving. Given $w \in \mathbb{P}^*$, let

$$w = \psi_1 \ \omega_1 \ \psi_2 \ \omega_2 \ \ldots \ \omega_{m-1}\psi_m$$

be its k-factorization. Define

$$w' = g(w) = \psi_1 \ f(\omega_1) \ \psi_2 \ f(\omega_2) \ \ldots \ f(\omega_{m-1})\psi_m.$$

This is clearly a bijective rearrangement, so we just need to verify (6).

If yuz embeds in w at some index, then we must show yvz embeds in w' at the same index. (Showing the converse is similar.) Now $yuz \leqslant w$ if and only if $u \leqslant \omega_i$ for some i. By assumption, v embeds in $f(\omega_i)$ at the same index. We will show that y embeds in w' just before this embedding of v. (The proof that z embeds just after is similar.) So consider any element y_p with $y_p \leqslant w_q$ in the embedding of yuz in w. If $w_q \in \psi_j$ for some j, then $w'_q = w_q \geqslant y_p$. If $w_q \in \omega_j$ for some j, then $w'_q \geqslant k \geqslant y_p$ because f is a rearrangement. So y_q will still embed at index q in w'. Thus yvz embeds in w' as desired and we have completed the proof. $\square$

As an application of this theorem, we will derive a strong Wilf equivalence in $\mathfrak{S}_5$ which we could not obtain from our previous results alone. The proof of Lemma 5.1 (b) shows that $123 \sim_s 231$ is witnessed by a rearrangement. From this and the proof of Lemma 5.1 (c), it follows that $345 \sim_s 453$ is witnessed by a rearrangement. So the theorem just proved shows that $34512 \sim_s 45312$.

6 Computations

We will now explicitly calculate the generating functions $S(u;t,x)$ for two families of words u. Aside from providing an application of the ideas from the previous sections, these particular power series are of interest because they have numerators which are single monomials. This is not always the case. For example,

$$S(212;t,x) = \frac{t^3x^5(1+tx^2)}{(1-x)(1-x+t^2x^3)(1-x-tx+tx^2-t^2x^3)}.$$

One can use the theory of Gröbner bases to show that $(1-x)(1-x+t^2x^3)(1-x-tx+tx^2-t^2x^3)$ is not in the ideal generated by $1+tx^2$. So $1+tx^2$ does not divide $(1-x)(1-x+t^2x^3)(1-x-tx+tx^2-t^2x^3)$ and we can not write $S(212;t,x)$ in the form $t^ax^b/Q(x,t)$ for some polynomial $Q(x,t)$.

We first determine the generating function for increasing permutations. It will be convenient to have the standard notation that, for a nonnegative integer k,

$$[k]_x = 1 + x + x^2 + \cdots + x^{k-1}.$$

Theorem 6.1. *For $n \geqslant 2$, define polynomials $B_n(t,x)$ by*

$$\begin{aligned} B_2(t,x) &= tx(1-x)^2, \\ B_{n+1}(t,x) &= tx^{n+1}B_n(t,x) + tx(1-x)^n(1-x^n). \end{aligned}$$

Then

$$S(12\ldots n;t,x) = \frac{t^n x^{\binom{n+1}{2}}}{(1-x)^n - B_n(t,x).}$$

Proof. Since $12\ldots n \sim n\ldots 21$, it suffices to compute the generating function for the latter. In that case, one can simplify the automaton Δ constructed in Theorem 2.2.

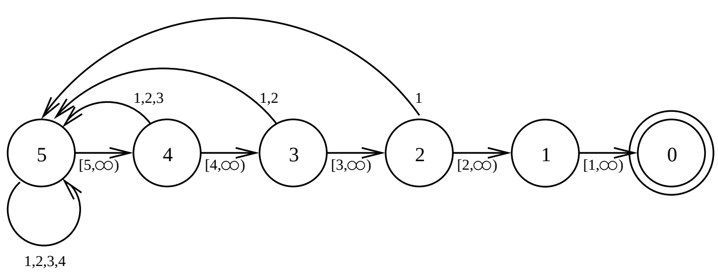

Figure 2: An automaton accepting $\mathcal{S}(54321)$.

Note that T is an accepting state for Δ if and only if $\max T = n$ (where we define $\max\emptyset = 0$). Furthermore, because of our choice of permutation, if there is an arc from T to U labeled a, then $\max U$ is completely determined by $\max T$ and a. So we can contract all the states with the same maximum into one. And when we do so, arcs of the same label will collapse together. The result for $n = 5$ is shown in Figure 2. For convenience in later indexing, the state labeled k is the one resulting from amalgamating those with maximum $n-k$.

Let $\mathcal{L}_k$ be the language of all words u such that the path for w starting at state k leads to the accepting state 0. Consider the corresponding generating function $L_k = \sum_{u \in \mathcal{L}_k} \mathrm{wt}(u)$. Directly from the automaton, we have $L_0 = 1$ and

$$L_k = \frac{tx^k}{(1-x)} L_{k-1} + tx[k-1]_x L_n$$

for $k \geqslant 1$. It is now easy to prove by induction that, for $k \geqslant 2$,

$$L_k = \frac{t^k x^{\binom{k+1}{2}} + B_k(t,x) L_n}{(1-x)^k}.$$

Plugging in $k = n$ and solving for $L_n = S(n \ldots 21; t, x)$ completes the proof. □

Theorem 6.2. *For any integers $k \geqslant 0$, $\ell \geqslant 1$, and $b \geqslant 2$ we have*

$$S(1^k b^\ell; t, x) = \frac{t^{k+\ell} x^{k+b\ell}}{(1-x)^{k+1} \left((tx^b)^{\ell-1}(1 - tx[b-1]_x) + (1-x-tx) \sum_{i=0}^{\ell-2} (1-x)^i (tx^b)^{\ell-2-i} \right)}.$$

Proof. Suppose $w = w_1 \ldots w_n \in \mathcal{S}(1^k b^\ell)$. Then to have $1^k b^\ell$ as a suffix, we must have $w_n, \ldots, w_{n-\ell+1} \geqslant b$.

There are now two cases depending on the length of w. If $|w| = k + \ell$ then $w_1, \ldots, w_k$ are arbitrary positive integers. If $|w| > k + \ell$ then write $w = yaz$ where $|z| = \ell$ and $a \in \mathbb{P}$. In order to make sure that $1^k b^\ell$ does not have another embedding intersecting z it is necessary and sufficient that $a < b$. And ruling out any embeddings inside y is equivalent to $y \in \mathcal{A}(1^k b^\ell)$. We must also make sure that $|y| \geqslant k$ in order to have $|w| > k + \ell$.

Let $S = S(1^k b^\ell; t, x)$ and $A = A(1^k b^\ell; t, x)$. Turning all the information about w into a generating function identity gives

$$S = \left(\frac{tx^b}{1-x} \right)^\ell \left[\left(\frac{tx}{1-x} \right)^k + tx[b-1]_x \left(A - [k]_{tx/(1-x)} \right) \right].$$

Also, combining the two parts of Corollary 3.1 gives

$$A = \frac{(1-x)(1-S)}{1-x-tx}.$$

Substituting this expression for A into our previous equation, one can easily solve for S to obtain that

$$S(1^k b^\ell; t, x)) = \frac{t^{k+\ell} x^{k+b\ell}(1-x-tx^b)}{(1-x)^{k+1}((1-x)^{\ell-1}(1-x-tx) + t^{\ell+1} x^{b\ell+1}[b-1]_x)}.$$

Thus to finish the proof, one need only show that

$$\frac{(1-x)^{\ell-1}(1-x-tx)+t^{\ell+1}x^{b\ell+1}[b-1]_x}{(1-x-tx^b)}$$
$$=(tx^b)^{\ell-1}(1-tx[b-1]_x)+(1-x-tx)\sum_{i=0}^{\ell-2}(1-x)^i(tx^b)^{\ell-2-i}$$

which can be easily verified by cross multiplication. □

7 The Möbius function

We will now show that the language for the Möbius function of ordinary factor order is not regular. This is somewhat surprising because Björner and Reutenauer [4] showed that this language is regular if one considers ordinary subword order, and then Björner and Sagan [5] extended this result to generalized subword order. We will begin by reviewing some basic facts about Möbius functions. The reader wishing more details can consult [10, Chapter 3].

For any poset P, the *incidence algebra* of P over the integers is

$$I(P)=\{\alpha : P\times P\to\mathbb{Z}\ :\ \alpha(a,b)=0 \text{ if } a\not\leqslant b\}.$$

This set is an algebra whose multiplication is given by *convolution*

$$(\alpha*\beta)(a,b)=\sum_{c\in P}\alpha(a,c)\beta(c,b).$$

It is easy to see that the identity for this operation is the *Kronecker delta*

$$\delta(a,b)=\begin{cases}1 & \text{if } a=b,\\ 0 & \text{else.}\end{cases}$$

So it is possible for incidence algebra elements to have multiplicative inverses.

One of the simplest elements of $I(P)$ is the *zeta function*

$$\zeta(a,b)=\begin{cases}1 & \text{if } a\geqslant b,\\ 0 & \text{else.}\end{cases}$$

Note that $F(u)$ can be rewritten as

$$F(u)=\sum_{w\in P^*}\zeta(u,w)w.$$

It turns out that ζ has a convolutional inverse μ in $I(P)$. This function is important in enumerative and algebraic combinatorics. Björner [3] has given a formula for μ in ordinary factor order which we will need. To describe this result, we must make some definitions. The *dominant outer factor* or *border* of w, denoted $o(w)$, is the longest word other than w which is both a prefix and a suffix of w. Note that we may have $o(w)=\epsilon$. The *dominant inner factor* of $w=w_1\dots w_\ell$, written $i(w)$, is $w_2\dots w_{\ell-1}$. Finally, a word is *flat* if all its elements are equal. For example, $w=abbaabb$ has $o(w)=abb$ and $i(w)=bbaab$.

Theorem 7.1 (Björner)**.** *In (ordinary) factor order, if $u \leqslant w$ then*

$$\mu(u,w) = \begin{cases} \mu(u, o(w)) & \text{if } |w|-|u|>2 \text{ and } u \leqslant o(w) \not\leqslant i(w), \\ 1 & \text{if } |w|-|u|=2,\ w \text{ is not flat, and } u = o(w) \text{ or } i(w), \\ (-1)^{|w|-|u|} & \text{if } |w|-|u|<2, \\ 0 & \text{otherwise.} \end{cases}$$

□

Continuing the example

$$\mu(b, abbaabb) = \mu(b, abb) = 1.$$

Note that this description is inductive. It also implies that $\mu(u,w)$ is ± 1 or 0 for all u, w in factor order.

We will show that the language $\mathcal{M}(u) = \{w \ : \ \mu(u,w) \neq 0\}$ need not be regular. To do this, we will need the Pumping Lemma which we now state. A proof can be found in the text of Hopcroft and Ullman [7, pp. 55–56].

Lemma 7.2 (Pumping Lemma)**.** *Let $\mathcal{L}$ be a regular language. Then there is a constant $n \geqslant 1$ such that any $z \in \mathcal{L}$ can be written as $z = uvw$ satisfying*

1. *$|uv| \leqslant n$ and $|v| \geqslant 1$,*
2. *$uv^i w \in \mathcal{L}$ for all $i \geqslant 0$.* □

Roughly speaking, any word in a regular language has a prefix of bounded length such that pumping up the end of the prefix keeps one in the language.

Theorem 7.3. *Consider (ordinary) factor order where $P = \{a, b\}$. Then $\mathcal{M}(a)$ is not regular.*

Proof. Suppose, to the contrary, that $\mathcal{M}(a)$ is regular and let n be the constant guaranteed by the pumping lemma. We will derive a contradiction by letting $z = ab^n ab^n a$ where, as usual, b^n represents the letter b repeated n times.

First we show that $z \in \mathcal{M}(a)$. Indeed, $o(z) = ab^n a$ and $i(z) = b^n ab^n$ which implies that $a \leqslant o(z) \not\leqslant i(z)$. So we are in the first case of Björner's formula and $\mu(a,z) = \mu(a, ab^n a)$. Repeating this analysis with $ab^n a$ in place of z gives $\mu(a,z) = \mu(a,a) = 1$. Hence $z \in \mathcal{M}(a)$ as promised.

Now pick any prefix uv of z as in the Pumping Lemma. There are two cases. The first is if $u \neq \epsilon$. So $v = b^j$ for some j with $1 \leqslant j < n$. Picking $i = 2$, we conclude that $z' = uv^2 w = ab^{n+j} ab^n a$ is in $\mathcal{M}(a)$. But $o(z') = a$ and $i(z') = b^{n+j} ab^n$. Thus $|z'| - |a| > 2$ and $a \leqslant o(z') \leqslant i(z')$, so z' does not fall into any of the first three cases of Björner's formula. This implies that $\mu(a, z') = 0$ and hence $z' \notin \mathcal{M}(a)$, which is a contradiction in this case.

The second possibility is that $u = \epsilon$ and $v = ab^j$ for some $0 \leqslant j < n$. Similar considerations to those in the previous paragraph show that if we take $z' = uv^2 w$ then $\mu(a, z') = 0$ again. So we have a contradiction as before and the theorem is proved. □

8 Comments, conjectures, and open questions

8.1 Mixing factors and subwords

It is possible to create languages using combinations of factors and subwords. This is an idea that was first studied by Babson and Steingrímsson [1] in the context of pattern avoidance in permutations. Many of the results we have proved can be generalized in this way. We will indicate how this can be done for Theorem 2.2.

A *pattern* p over P is a word in P^* where certain pairs of adjacent elements have been overlined (barred). For example, in the pattern $p = 1\overline{133}24\overline{61}$ the pairs 13, 33, and 61 have been overlined. If $w \in P^*$ we will write $\overline{w}$ for the pattern where every pair of adjacent elements in w is overlined. So every pattern has a unique factorization of the form $p = \overline{y_1}\ \overline{y_2}\ \ldots\ \overline{y_k}$. In the preceding example, the factors are $y_1 = 1$, $y_2 = 133$, $y_3 = 2$, $y_4 = 4$, and $y_5 = 61$.

If $p = \overline{y_1}\ \overline{y_2}\ \ldots\ \overline{y_k}$ is a pattern and $w \in P^*$ then p *embeds* into w, written $p \to w$, if there is a subword $w' = z_1 z_2 \ldots z_k$ of w where, for all i,

1. z_i is a factor of w with $|z_i| = |y_i|$, and
2. $y_i \leqslant z_i$ in generalized factor order.

For example $\overline{32}4 \to 14235$ and there is only one embedding, namely 425. For any pattern p, define the language

$$\mathcal{F}(p) = \{w \in P^* \ : \ p \to w\}$$

and similarly for $\mathcal{S}(p)$ and $\mathcal{A}(p)$. The next result generalizes Theorem 2.2 to an arbitrary pattern.

Theorem 8.1. *Let P be any poset and let p be a pattern over P. Then there are NFAs accepting $\mathcal{F}(p)$, $\mathcal{S}(p)$, and $\mathcal{A}(p)$.*

Proof. As before, it suffices to build an NFA, Δ, for $\mathcal{S}(p)$. It will be simplest to construct an NFA with ϵ-moves, i.e., with certain arcs labeled ϵ whose traversal does not append anything to the word being constructed. It is well known that the set of languages accepted by NFAs with ϵ-moves is still the set of regular languages.

Let $p = \overline{y_1}\ \overline{y_2}\ \ldots\ \overline{y_k}$ be the factorization of p and, for all i, let Δ_i be the automaton constructed in Theorem 2.2 for $\mathcal{S}(y_i)$. We can paste these automata together to get Δ as follows. For each i with $1 \leqslant i < k$, add an ϵ-arc from every final state of Δ_i to the initial state of Δ_{i+1}. Now let the initial state of Δ be the initial state of Δ_1 and the final states of Δ be the final states of Δ_k. It is easy to see that the resulting NFA accepts the language $\mathcal{S}(p)$. □

8.2 Rationality for infinite posets

It would be nice to have a criterion that would imply rationality even for some infinite posets P. To this end, let $\mathbf{x} = \{x_1, \ldots, x_m\}$ be a set of commuting variables and consider the formal power series algebra $\mathbb{Z}[[\mathbf{x}]]$. Suppose we are given a function

$$\mathrm{wt} : P \to \mathbb{Z}[[\mathbf{x}]]$$

which then defines a weighting of words $w = w_1 \dots w_\ell \in P^*$ by

$$\mathrm{wt}(w) = \prod_{i=1}^{m} \mathrm{wt}(w_i).$$

To make sure our summations will be defined in $\mathbb{Z}[[\mathbf{x}]]$, we assume that there are only finitely many w of any given weight and call such a weight function *regular*.

For $u \in P^*$, let

$$F(u; \mathbf{x}) = \sum_{w \geqslant u} \mathrm{wt}(w)$$

and similarly for $S(u; \mathbf{x})$ and $A(u; \mathbf{x})$. Suppose we want to make sure that $S(u; \mathbf{x})$ is rational. As done in Section 3, we can consider a transfer matrix with entries

$$M_{T,U} = \sum_{a} \mathrm{wt}(a)$$

where the sum is over all $a \in P$ occurring on arcs from T to U. Equation (2) remains the same, so it suffices to make sure that $M_{T,U}$ is always rational.

If there is an arc labeled a from T to U then we must have $U \subseteq T'$ where T' is given in equation (3). Recalling the definition of Δ from the proof of Theorem 2.2, we see that the a's appearing in the previous sum are exactly those satisfying

1. $a \geqslant u_{t+1}$ for $t + 1 \in U$, and

2. $a \not\geqslant u_{t+1}$ for $t + 1 \in T' - U$.

To state these criteria succinctly, for any subword y of u we write $a \geqslant y$ (respectively, $a \not\geqslant y$) if $a \geqslant b$ (respectively, $a \not\geqslant b$) for all $b \in y$. Finally, note that, from the proof of Theorem 2.2, similar transfer matrices can be constructed for $F(u; \mathbf{x})$ and $A(u; \mathbf{x})$. We have proved the following result which generalizes Theorem 3.2.

Theorem 8.2. *Let P be a poset with a regular weight function $\mathrm{wt} : P^* \to \mathbb{Z}[[\mathbf{x}]]$, and let $u \in P^*$. Suppose that for any two subwords y and z of u we have*

$$\sum_{\substack{a \geqslant y \\ a \not\geqslant z}} \mathrm{wt}(a)$$

is a rational function. Then so are $F(u; \mathbf{x})$, $S(u; \mathbf{x})$, and $A(u; \mathbf{x})$. □

8.3 Irrationality for infinite posets

When P is countably infinite it is possible for the generating functions we have considered to be irrational. As an example, pick a distinguished element $a \in P$. For any $A \subseteq P$ with $a \in A$, we define an order $\leqslant_A$ by insisting that the elements of $P - \{a\}$ form an antichain, and that $a \leqslant_A b$ if and only if $b \in A$. Consider the corresponding language $\mathcal{S}_A$. Clearly $\mathcal{S}_A = (P - A)^* A$ and so no two of these languages are equal. It follows that the mapping $A \to \mathcal{S}_A$ is injective. So one of the $\mathcal{S}_A$ must be irrational since there are uncountably many possible A but only countably many rational functions in $\mathbb{Z}\langle\langle P \rangle\rangle$.

8.4 Wilf equivalence and strong equivalence

There are a number of open problems and questions raised by our work on Wilf equivalence.

(1) **If $u \sim v$, then must v be a rearrangement of u?** This is the case for all the Wilf equivalences we have proved. Note that if the answer is "yes," then the Wilf equivalences for the symmetric groups given in Table 1 of Section 9 are actually Wilf equivalence classes.

(2) **What about Wilf equivalence in $[m]^*$ where $[m] = \{1, 2, \ldots, m\}$?** Given a positive integer m, one can define Wilf equivalence of words $u, v \in [m]^*$ in the same way that we did for $\mathbb{P}^*$. We write $u \sim_m v$ for this relation. Is it true that $u \sim_m v$ if and only if $u \sim v$?

(3) **If $u^+ \sim v^+$ then is $u \sim v$?** In other words, does the converse of Lemma 4.1 (c) hold? We note that the converse of (b) is true. For suppose $1u \sim 1v$ and let $f : \mathcal{S}(1u) \to \mathcal{S}(1v)$ be a corresponding map. Then to construct $g : \mathcal{S}(u) \to \mathcal{S}(v)$ we consider two cases for $w \in \mathcal{S}(u)$. If $|w| > |u|$ then $w \in \mathcal{S}(1u)$ so let $g(w) = f(w)$. Otherwise $|w| = |u|$ and so let $g(w) = v + (w - u)$ where addition and subtraction is done componentwise. It is easy to check that g is well defined and weight preserving.

(4) **Find a theorem which, together with the results already proved, explains all the Wilf equivalences in $\mathfrak{S}_5$.** In particular, the results of Section 4 and the last paragraph of Section 5 generate all of the Wilf equivalences in Table 1 with one exception. In particular, our results show that

$$31425 \sim 31524 \sim 42513 \sim 52413 \quad \text{and} \quad 32415 \sim 32514 \sim 41523 \sim 51423.$$

but not why a permutation of the first group is Wilf equivalent to one of the second. However, we do have a conjecture which has been verified by computer in a large number of examples and which would connect these two groups.

Conjecture 8.3. *For any $a, b, c \in [2, \infty)$ we have*

$$a1b2c \sim a2b1c.$$

(5) **Is it always the case that the number of elements of $\mathfrak{S}_n$ Wilf equivalent to a given permutation is a power of 2?** This is always true in Table 1.

(6) **Is it true that** $312 \sim_s 213$**?** From our results on strong Wilf equivalence it follows that $12 \sim_s 21$ and $123 \sim_s 132 \sim_s 231 \sim_s 321$. So all the Wilf equivalent elements in $\mathfrak{S}_2$ and $\mathfrak{S}_3$ are actually strongly Wilf equivalent with the possible exception of the pair in the question. Of course, this breaks down in $\mathfrak{S}_4$ as noted in Section 5.

(7) **Does Theorem 5.3 remain true if one replaces strong Wilf equivalence with ordinary Wilf equivalence throughout?** If so, a completely different proof will have to be found for that case.

8.5 The language $\mathcal{M}(u)$

We have shown that $\mathcal{M}(u)$ is not always regular and so the corresponding generating function $M(u)$ is not always rational. But this leaves open whether $\mathcal{M}(u)$ might fall into a more general class of languages such as context free grammars. A *context free grammar* or *CFG* is a quadruple $G = (V, S, T, P)$ where

1. V is a finite set of *variables*,
2. S is a special variable called the *start symbol*,
3. T is a finite set of terminals disjoint from V, and
4. P is a finite set of *productions* of the form $A \to \alpha$ where $A \in V$ and $\alpha \in (V \cup T)^*$.

There is a Pumping Lemma for CFGs, see [7, Section 6.1]. So it is tempting to try and modify the proof of Theorem 7.3 to show that $\mathcal{M}(u)$ is not even a CFG. However, all our attempts in that direction have failed. Is $\mathcal{M}(u)$ a CFG or not?

9 Tables

The following two tables were constructed by having a computer calculate, for each composition u, the generating functions $S(u; t, x)$. This was done with the aid of the corresponding automaton from Section 2.

Acknowledgement. We would like to thank the two anonymous referees for suggestions which greatly improved the exposition, as well as for providing the simpler proof of Proposition 1.1.

12, 21
123, 132, 231, 321
213, 312
1234, 1243, 1342, 1432, 2341, 2431, 3421, 4321
1324, 1423, 3241, 4231
2134, 2143, 3412, 4312
3124, 3214, 4123, 4213
2314, 2413, 3142, 4132
12345, 12354, 12453, 12543, 13452, 13542, 14532, 15432, 23451, 23541, 24531, 25431, 34521, 35421, 45321, 54321
12435, 12534, 14352, 15342, 24351, 25341, 43521, 53421
13245, 13254, 14523, 15423, 32451, 32541, 45231, 54231
21345, 21354, 21453, 21543, 34512, 35412, 45312, 54312
23145, 23154, 45132, 54132
32145, 32154, 45123, 54123
24153, 25143, 34152, 35142
14235, 14325, 15234, 15324, 42351, 43251, 52341, 53241
31425, 31524, 32415, 32514, 41523, 42513, 51423, 52413
24315, 25314, 41352, 51342
24135, 25134, 43152, 53142
34215, 35214, 41253, 51243
34125, 35124, 42153, 52143
41325, 42315, 51324, 52314
41235, 43215, 51234, 53214
42135, 43125, 52134, 53124
13425, 13524, 14253, 15243, 34251, 35241, 42531, 52431
21435, 21534, 43512, 53412
24513, 25413, 31452, 31542
23415, 23514, 41532, 51432
31245, 31254, 45213, 54213

Table 1: Wilf equivalences for permutations of at most 5 elements

Equivalences	$S(u;t,x)$
1	$\frac{tx}{1-x}$
2	$\frac{tx^2}{(1-x)(1-tx)}$
3	$\frac{tx^3}{(1-x-tx+tx^3)}$
11	$\frac{t^2x^2}{(1-x)^2}$
12,21	$\frac{t^2x^3}{(1-x)^2(1-tx)}$
13,31	$\frac{t^2x^4}{(1-x)^2(1-tx-tx^2)}$
22	$\frac{t^2x^4}{(1-x)(1-x-tx+tx^2-t^2x^3)}$
23,32	$\frac{t^2x^5}{(1-x)(1-x-tx+tx^3-t^2x^4)}$
33	$\frac{t^2x^6}{(1-x)(1-x-tx+tx^3-t^2x^4-t^2x^5)}$
111	$\frac{t^3x^3}{(1-x)^3}$
112,121,211	$\frac{t^3x^4}{(1-x)^3(1-tx)}$
122,221	$\frac{t^3x^5}{(1-x)^2(1-x-tx+tx^2-t^2x^3)}$
212	$\frac{t^3x^5(1+tx^2)}{(1-x)(1-x+t^2x^3)(1-x-tx+tx^2-t^2x^3)}$
113,131,311	$\frac{t^3x^5}{(1-x)^3(1-tx-tx^2)}$
213,312	$\frac{t^3x^6(1+tx^3)}{(1-x)(1-x+t^2x^4)(1-x-tx+tx^3-t^2x^4)}$
123,132,231,321	$\frac{t^3x^6}{(1-x)^2(1-x-tx+tx^3-t^2x^4)}$
222	$\frac{t^3x^6}{(1-x)(1-2x-tx+x^2+2tx^2-tx^3-t^2x^3+t^2x^4-t^3x^5)}$
133,331	$\frac{t^3x^7}{(1-x)^2(1-x-tx+tx^3-t^2x^4-t^2x^5)}$
313	$\frac{t^3x^7(1+tx^3+tx^4)}{(1-x)(1-x+t^2x^4+t^2x^5)(1-x-tx+tx^3-t^2x^4-t^2x^5)}$
223,232,322	$\frac{t^3x^7}{(1-x)(1-2x-tx+x^2+tx^2+tx^3-tx^4-t^2x^4+t^2x^5-t^3x^6)}$
323	$\frac{t^3x^8(1+tx^3)}{(1-x)(1-2x-tx+x^2+tx^2+tx^3-tx^4-t^2x^4+t^2x^5-t^3x^6-t^3x^7+t^3x^8-t^4x^9-t^4x^{10})}$
233,332	$\frac{t^3x^8}{(1-x)(1-2x-tx+x^2+tx^2+tx^3-tx^4-t^2x^4+t^2x^6-t^3x^7)}$
333	$\frac{t^3x^9}{(1-x)(1-2x-tx+x^2+tx^2+tx^3-tx^4-t^2x^4+t^2x^6-t^3x^7-t^3x^8)}$

Table 2: Wilf equivalences for u with $|u| \leqslant 3$ and $u_i \leqslant 3$ for all i.

References

[1] Babson, E., and Steingrímsson, E. Generalized permutation patterns and a classification of the Mahonian statistics. *Sém. Lothar. Combin. 44* (2000), Art. B44b, 18 pp. (electronic).

[2] Berstel, J., and Reutenauer, C. *Rational series and their languages*, vol. 12 of *EATCS Monographs on Theoretical Computer Science*. Springer-Verlag, Berlin, 1988.

[3] Björner, A. The Möbius function of factor order. *Theoret. Comput. Sci. 117*, 1-2 (1993), 91–98.

[4] Björner, A., and Reutenauer, C. Rationality of the Möbius function of subword order. *Theoret. Comput. Sci. 98*, 1 (1992), 53–63. Second Workshop on Algebraic and Computer-theoretic Aspects of Formal Power Series (Paris, 1990).

[5] Björner, A., and Sagan, B. E. Rationality of the Möbius function of a composition poset. *Theoret. Comput. Sci. 359*, 1-3 (2006), 282–298.

[6] Goyt, A. The Möbius function of a restricted composition poset. Preprint can be found at `arXiv:0806.1500`.

[7] Hopcroft, J. E., and Ullman, J. D. *Introduction to automata theory, languages, and computation*. Addison-Wesley Publishing Co., Reading, Mass., 1979. Addison-Wesley Series in Computer Science.

[8] Kleene, S. C. Representation of events in nerve nets and finite automata. In *Automata studies*, Annals of mathematics studies, no. 34. Princeton University Press, Princeton, N. J., 1956, pp. 3–41.

[9] Schützenberger, M. P. On the definition of a family of automata. *Information and Control 4* (1961), 245–270.

[10] Stanley, R. P. *Enumerative Combinatorics. Vol. 1*, vol. 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997. With a foreword by Gian-Carlo Rota, Corrected reprint of the 1986 original.

[11] Wilf, H. S. The patterns of permutations. *Discrete Math. 257*, 2-3 (2002), 575–583.