ISSN 2590-9770
The Art of Discrete and Applied Mathematics 6 (2023) #P3.03
https://doi.org/10.26493/2590-9770.1552.b43
(Also available at http://adam-journal.eu)

Counting occurrences of subword patterns in non-crossing partitions

Toufik Mansour
Department of Mathematics, University of Haifa, 3498838 Haifa, Israel

Mark Shattuck*
Department of Mathematics, University of Tennessee, 37996 Knoxville, TN, USA

Received 8 July 2022, accepted 7 November 2022, published online 7 March 2023

Abstract

A permutation pattern in which all letters within an occurrence are required to be adjacent is known as a *subword*. In this paper, we consider the distribution of several infinite families of subword patterns on the set of non-crossing partitions of size n, denoted by NC_n, and derive formulas for the generating functions of these distributions on NC_n. As special cases of our results, we obtain formulas for the generating functions enumerating the members of NC_n according to the number of occurrences of any subword of length three with distinct letters. Simple expressions for the total number of occurrences of a pattern over all members of NC_n are also deduced. Some connections are made with the related problem of counting Dyck paths according to the number of occurrences of certain types of strings. Further, for the subwords $12\cdots m$ and $213\cdots m$ where $m \geq 3$, we consider the joint distribution with the descents statistic and make use of the kernel method to establish the results in these cases.

Keywords: Non-crossing partition, subword pattern, combinatorial statistic, set partition.

Math. Subj. Class.: 05A15, 05A18

1 Introduction

A *partition* of a set is a collection of pairwise disjoint nonempty subsets, called *blocks*, whose union is the set. The set of all partitions of $[n] = \{1, 2, \dots, n\}$ will be denoted by $\mathcal{P}_n$, with $\mathcal{P}_{n,k}$ the subset of $\mathcal{P}_n$ whose members contain exactly k blocks. Recall that

*Corresponding author.
E-mail addresses: tmansour@univ.haifa.ac.il (Toufik Mansour), mark.shattuck2@gmail.com (Mark Shattuck)

if a partition $\Pi = B_1/B_2/\cdots/B_k \in \mathcal{P}_{n,k}$ is written such that its blocks are arranged in increasing order of minimum elements, then Π is said to be in *standard form*. An equivalent sequential representation of Π, expressed in standard form, is obtained by writing $\pi = \pi_1\cdots\pi_n$, where $i \in B_{\pi_i}$ for each $i \in [n]$ (see, e.g., [15]). Then Π in standard form implies $\pi_{i+1} \leq \max(\pi_1\cdots\pi_i)+1$ for $1 \leq i \leq n-1$, which is known as the *restricted growth* condition (see, e.g., [11, 16]). The sequence π is referred to as the *canonical sequential form* of the partition Π.

Let $\tau = \tau_1\cdots\tau_m$ denote a sequence of positive integers whose set of distinct letters forms the set $[\ell]$ for some $1 \leq \ell \leq m$. Then the sequence $\rho = \rho_1\cdots\rho_n$ is said to *contain* τ if there exists a subsequence of ρ that is order-isomorphic to τ. Otherwise, ρ *avoids* τ, with τ often being referred to as a *pattern* in this context. Recall that a partition Π is said to be *non-crossing* if its canonical sequential form π avoids the pattern 1-2-1-2 in the classical sense (see, e.g., [5]). Let NC_n denote the set of non-crossing partitions of $[n]$; recall that NC_n has cardinality given by the n-th Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$ (see [14, Sequence A000108]).

If it is required further that the subsequence of ρ that is order-isomorphic to τ in the definition of pattern containment above correspond to consecutive entries of ρ (i.e., $\rho_i\rho_{i+1}\cdots\rho_{i+m-1}$ is isomorphic to τ for some $1 \leq i \leq n-m+1$), then ρ is said to contain τ as a *subword* (pattern). Here, we will be interested in counting the members of NC_n according to the number of occurrences of certain subword patterns, focusing on five infinite families of patterns. Other finite discrete structures with sequential representations that have been enumerated according to the number of subwords include k-ary words [1], set partitions [10], flattened partitions [9] and flattened involutions [7]. For examples of other types of statistics which have been studied on non-crossing partitions, we refer the reader to [6, 8, 13, 18, 19].

The organization of this paper is as follows. In the next section, we enumerate members of NC_n according to the joint distribution of the descents statistic with the number of occurrences of $12\cdots m$ for any $m \geq 3$ and provide a similar treatment for $213\cdots m$. To write a recurrence, we introduce a couple of auxiliary parameters and arrays which lead to a system of functional equations satisfied by the generating functions of the corresponding joint distributions. To solve the system in these cases, we make use of the *kernel method* [4]. In the third section, we count non-crossing partitions according to the number of occurrences of members of three additional families of subwords, namely, $23\cdots m1$, $34\cdots m21$ and $23\cdots(m-1)1m$, and make use of the *symbolic method* (see, e.g., [3]) to derive formulas for the generating functions of the corresponding distributions. Explicit formulas are found also by differentiation for the total number of occurrences of the various patterns over all members of NC_n.

As a consequence of our results, we obtain a formula for the generating function F_τ for the distribution of a subword pattern τ on NC_n for $n \geq 0$ in all cases when τ is a permutation of length three. See Table 1 below, where the corresponding functional equations for F_τ are given. The subwords 132 and 312 are trivial when considering non-crossing partitions in that they are always avoided. To see this, first note that within any set partition π, an occurrence of 132 or 312 implies the occurrence of the respective vincular pattern 1-2-132 or 1-2-312, wherein the entries of π corresponding to the first '1' and '2' in these patterns need not be adjacent to any of the other letters (this being indicated by their being separated by dashes). In either case, this would introduce a crossing, which is not allowed.

Subword	Generating function equation	Reference
123	$x((q-1)x-q)F^2+(1+(q-1)x)F-1=0$	Corollary 2.4
132	$xF^2-F+1=0$	Trivial
213	$x(q-(q-1)x)F^2-(1+x(q-1)(2-x))F+1+(q-1)x=0$	Corollary 2.9
231	$x(1+(q-1)x)F^2-(1+2(q-1)x^2)F+1+(q-1)x^2=0$	Theorem 3.1
312	$xF^2-F+1=0$	Trivial
321	$x(q-2(q-1)x)F^2-(1+x(q-1)(2-3x))F+1+x(q-1)(1-x)=0$	Theorem 3.1

Table 1: Generating functions $F=F_\tau$ for τ a permutation of length three.

2 Counting $12\cdots m$ and $213\cdots m$ subwords and descents

In this section, we consider the joint distribution on NC_n with descents of two infinite families of subword patterns.

2.1 Joint distribution of $12\cdots m$ with descents

Let $\tau=12\cdots m$, where $m\geq 3$ is fixed. We will compute a generating function (gf) formula for the joint distribution on NC_n for the parameters which track descents and the number of occurrences of τ. In order to write a recurrence for this distribution, we need to consider a further refinement of NC_n based on an apparently new parameter for the set. Given $\pi\in NC_n$, excluding the increasing partition $12\cdots n$, let $\text{rep}(\pi)$ denote the smallest repeated letter of π. Let $NC_{n,i}$ for $1\leq i\leq n-1$ denote the subset of NC_n consisting of those members π for which $\text{rep}(\pi)=i$. Define the distributions

$$a(n)=\sum_{\pi\in NC_n}p^{\text{desc}(\pi)}q^{\mu_\tau(\pi)},\qquad n\geq 0,$$

and

$$a(n,i)=\sum_{\pi\in NC_{n,i}}p^{\text{desc}(\pi)}q^{\mu_\tau(\pi)},\qquad 1\leq i\leq n-1,$$

where $\text{desc}(\pi)$ and $\mu_\tau(\pi)$ denote the number of descents and occurrences of τ in π, respectively. For example, if $n=5$ and $i=2$, then

$$NC_{5,2}=\{12222,12223,12232,12233,12234,12322,12324,12332,12342\},$$

which implies $a(5,2)=3+p+q+3pq+pq^2$ when $m=3$. Let $a(n,i)=0$ if $i\leq 0$ or $i\geq n$.

In order to determine $a(n)$ and $a(n,i)$, we need to consider a generalization of these arrays as follows. Let $\mu_\tau^{(\ell)}(\pi)$ for $\ell\geq 0$ denote the number of occurrences of τ in $12\cdots\ell\pi$, where the partition π is represented using the letters in $\{\ell+1,\ell+2,\ldots\}$. Let $a^{(\ell)}(n)$ and $a^{(\ell)}(n,i)$ denote the respective joint distributions of $\mu_\tau^{(\ell)}$ with desc on NC_n and $NC_{n,i}$. For example, when $\ell=1$ and $m=3$, we have $a^{(1)}(3,1)=2+pq$, $a^{(1)}(3,2)=q$ and $a^{(1)}(3)=2+q+pq+q^2$. We seek to find formulas for $a(n)=a^{(0)}(n)$ and $a(n,i)=a^{(0)}(n,i)$.

The $a^{(\ell)}(n)$ and $a^{(\ell)}(n,i)$ are determined recursively as follows.

Lemma 2.1. *Let $\ell\geq 0$ and $m\geq 3$. If $n\geq 2$, then*

$$a^{(\ell)}(n,i)=p\sum_{j=i+1}^{n}a^{(\ell+i)}(j-i-1)a^{(0)}(n-j+1)\tag{2.1}$$

$$+ (1-p)a^{(\ell+i)}(0)a^{(0)}(n-i), \qquad 1 \le i \le n-1.$$

For $n \geq 1$, we have

$$a^{(\ell)}(n) = \sum_{i=1}^{n-1} a^{(\ell)}(n,i) + \begin{cases} 1, \text{ if } n+\ell < m, \\ q^{n+\ell-m+1}, \text{ if } n+\ell \geq m, \end{cases} \tag{2.2}$$

with $a^{(\ell)}(0) = \begin{cases} 1, \text{ if } \ell < m, \\ q^{\ell-m+1}, \text{ if } \ell \geq m. \end{cases}$

Proof. To show (2.1), first consider the position j of the second occurrence of the letter i within $\pi \in NC_{n,i}$. Then π may be expressed in the form $\pi = 12\cdots i\rho i\rho'$, where ρ is of length $j-i-1$ and ρ or ρ' may be empty. Note that π non-crossing implies any letters of ρ exceed i and any letters in ρ' other than i exceed $\max(\rho)$ since the smallest repeated letter of π is i. Further, ρ and $i\rho'$ are seen to be non-crossing partitions on their respective sets of letters. To determine $\mu_\tau^{(\ell)}(\pi)$, we count separately the occurrences of τ in the two sections $12\cdots(\ell+i)\rho$ and $(\ell+i)\rho'$ of $12\cdots\ell\pi$ and then combine the results. Note that ρ and ρ' here are expressed respectively using letters in $\{\ell+i+1, \ell+i+2, \ldots\}$ and $\{\ell+i\} \cup \{m+1, m+2, \ldots\}$, where $m = \max(\rho)$ if $\rho \neq \varnothing$ and $m = \ell+i$ if $\rho = \varnothing$. The distributions corresponding to the counts on the two sections are independent of one another and are seen to be given by $a^{(\ell+i)}(j-i-1)$ and $a^{(0)}(n-j+1)$, respectively. This follows from the fact that ρ contains no letters in $[\ell+i]$, with each letter in $[\ell+i-1]$ occurring only once in $12\cdots\ell\pi$ (at the beginning), and noting that the predecessor of the second $\ell+i$ cannot be strictly less than $\ell+i$. Finally, note that if $j > i+1$, then there is an additional descent due to the final letter of ρ being strictly greater than its successor.

Thus, the weight of all π of the stated form for each $j \in [i+2, n]$ equals $pa^{(\ell+i)}(j-i-1)a^{(0)}(n-j+1)$, and summing over j gives the weight of all π for which ρ is nonempty. If $j = i+1$ (i.e., $\rho = \varnothing$), then no extra descent as described above arises and there is a contribution towards the weight of $a^{(\ell+i)}(0)a^{(0)}(n-i)$. Combining this with the previous case where $j > i+1$ yields (2.1). Formula (2.2) follows from the definitions and noting that the weight of $\pi = 12\cdots n$ is given by $q^{n+\ell-m+1}$ if $n+\ell \geq m$ and 1 otherwise, since there are $n+\ell-m+1$ occurrences of τ in $12\cdots(\ell+n)$ in the first case and none in the second. Allowing n to be zero in the preceding implies the formula stated above for $a^{(\ell)}(0)$, which completes the proof. □

If $n \geq 2$ and $\ell \geq 0$, then let $A_n^{(\ell)}(v) = \sum_{i=1}^{n-1} a^{(\ell)}(n,i)v^{i-1}$ and $A_n(u,v) = \sum_{\ell\geq 0} A_n^{(\ell)}(v)u^\ell$. Let $B_n(u) = \sum_{\ell\geq 0} a^{(\ell)}(n)u^\ell$ for $n \geq 0$. Define the further gf's by

$$A(x,u,v) = \sum_{n\geq 2} A_n(u,v)x^n = \sum_{n\geq 2}\sum_{\ell\geq 0}\sum_{i=1}^{n-1} a^{(\ell)}(n,i)v^{i-1}u^\ell x^n$$

and

$$B(x,u) = \sum_{n\geq 0} B_n(u)x^n = \sum_{n\geq 0}\sum_{\ell\geq 0} a^{(\ell)}(n)u^\ell x^n.$$

Then $A(x,u,v)$ and $B(x,v)$ satisfy the following system of functional equations.

Lemma 2.2. *We have*

$$A(x,u,v)=\frac{px(B(x,0)-1)}{vx-u}(B(x,vx)-B(x,u))$$
$$+\frac{x(1-p)(B(x,0)-1)}{vx-u}\left(\frac{(1-u)(vx-(vx)^m)-(1-vx)(u-u^m)}{(1-u)(1-vx)}\right.$$
$$\left.+\frac{q(vx)^m(1-qu)-qu^m(1-qvx)}{(1-qu)(1-qvx)}\right), \tag{2.3}$$
$$B(x,u)=A(x,u,1)+\frac{x-u+u^{m+1}(1-x)-x^{m+1}(1-u)}{(1-u)(1-x)(x-u)}$$
$$+\frac{qx^{m+1}(1-qu)-qu^{m+1}(1-qx)}{(1-qu)(1-qx)(x-u)}. \tag{2.4}$$

Proof. By (2.1), we have

$$A(x,u,v)=p\sum_{n\geq2}\sum_{\ell\geq0}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}a^{(\ell+i)}(j-i-1)a^{(0)}(n-j+1)v^{i-1}u^\ell x^n$$
$$+(1-p)\sum_{n\geq2}\sum_{\ell\geq0}\sum_{i=1}^{n-1}a^{(\ell+i)}(0)a^{(0)}(n-i)v^{i-1}u^\ell x^n$$
$$=p\sum_{\ell\geq0}\sum_{i\geq1}\sum_{j\geq0}\sum_{n\geq1}a^{(\ell+i)}(j)a^{(0)}(n)v^{i-1}u^\ell x^{n+j+i}$$
$$+(1-p)\sum_{\ell\geq0}\sum_{i\geq1}\sum_{n\geq1}a^{(\ell+i)}(0)a^{(0)}(n)v^{i-1}u^\ell x^{n+i}$$
$$=p(B(x,0)-B(0,0))\sum_{\ell\geq0}\sum_{i\geq\ell+1}\sum_{j\geq0}a^{(i)}(j)v^{i-\ell-1}u^\ell x^{j+i-\ell}$$
$$+(1-p)(B(x,0)-B(0,0))\sum_{\ell\geq0}\sum_{i\geq\ell+1}a^{(i)}(0)v^{i-\ell-1}u^\ell x^{i-\ell}$$
$$=p(B(x,0)-1)\sum_{i\geq1}\sum_{\ell=0}^{i-1}\sum_{j\geq0}a^{(i)}(j)v^{i-\ell-1}u^\ell x^{j+i-\ell}$$
$$+(1-p)(B(x,0)-1)\sum_{i\geq1}\sum_{\ell=0}^{i-1}a^{(i)}(0)v^{i-\ell-1}u^\ell x^{i-\ell}$$
$$=p(B(x,0)-1)\sum_{i\geq1}\sum_{j\geq0}a^{(i)}(j)\frac{v^ix^{j+i+1}-x^{j+1}u^i}{vx-u}$$
$$+(1-p)(B(x,0)-1)\sum_{i\geq1}a^{(i)}(0)\frac{v^ix^{i+1}-u^ix}{vx-u}$$
$$=\frac{x(B(x,0)-1)}{vx-u}\Bigg(p(B(x,vx)-B(x,u))$$
$$+(1-p)\sum_{i=1}^{m-1}\left((vx)^i-u^i\right)+(1-p)\sum_{i\geq m}q^{i-m+1}\left((vx)^i-u^i\right)\Bigg)$$

$$\begin{aligned}&= \frac{px(B(x,0)-1)}{vx-u}(B(x,vx)-B(x,u)) \\ &\quad + \frac{x(1-p)(B(x,0)-1)}{vx-u}\left(\frac{(1-u)(vx-(vx)^m)-(1-vx)(u-u^m)}{(1-u)(1-vx)}\right. \\ &\quad \left. + \frac{q(vx)^m(1-qu)-qu^m(1-qvx)}{(1-qu)(1-qvx)}\right).\end{aligned}$$

By (2.2), we have

$$\begin{aligned}B(x,u) &= \sum_{n\geq 0}\sum_{\ell\geq 0} a^{(\ell)}(n)u^\ell x^n \\ &= \sum_{n\geq 2}\sum_{\ell\geq 0}\sum_{i=1}^{n-1} a^{(\ell)}(n,i)u^\ell x^n \\ &\quad + \sum_{\ell\geq 0} a^{(\ell)}(0)u^\ell + \sum_{\ell=0}^{m-2} u^\ell\left(\sum_{n=1}^{m-\ell-1} x^n + \sum_{n\geq m-\ell} q^{n+\ell-m+1}x^n\right) \\ &\quad + \sum_{\ell\geq m-1} u^\ell \sum_{n\geq 1} q^{n+\ell-m+1}x^n \\ &= A(x,u,1) + B(0,u) + \sum_{\ell=0}^{m-2} u^\ell\left(\frac{x-x^{m-\ell}}{1-x}\right) \\ &\quad + \sum_{\ell=0}^{m-2}\frac{qu^\ell x^{m-\ell}}{1-qx} + \sum_{\ell\geq m-1}\frac{q^{\ell-m+2}u^\ell x}{1-qx} \\ &= A(x,u,1) + B(0,u) + \frac{x(1-u^{m-1})}{(1-u)(1-x)} - \frac{x^2(x^{m-1}-u^{m-1})}{(1-x)(x-u)} \\ &\quad + \frac{qx^2(x^{m-1}-u^{m-1})}{(1-qx)(x-u)} + \frac{qxu^{m-1}}{(1-qu)(1-qx)} \\ &= A(x,u,1) + B(0,u) + \frac{x^2-xu-x^{m+1}(1-u)+x(1-x)u^m}{(1-u)(1-x)(x-u)} \\ &\quad + \frac{qx^{m+1}(1-qu)-qxu^m(1-qx)}{(1-qu)(1-qx)(x-u)}.\end{aligned}$$

From the formula for $a^{(\ell)}(0)$, we have $B(0,u)=\frac{1-u^m}{1-u}+\frac{qu^m}{1-qu}$. Combining this with the last expression above for $B(x,u)$, and simplifying, gives (2.4) and completes the proof. □

The next result expresses $B(x,0)$ for each m in terms of a root of a certain polynomial.

Theorem 2.3. *The generating function enumerating members of NC_n jointly according to the number of descents and occurrences of $12\cdots m$ for $m\geq 3$ is given by $B(x,0)=\frac{f+p-1}{p}$, where f satisfies*

$$(q-1)x^m f^m - qx^2f^3 + x(1+q)f^2 - (1+qx)f + 1 + \frac{(p-1)(f-1)(xf-1)(qxf-1)}{f+p-1} = 0. \quad (2.5)$$

Proof. Formulas (2.3) and (2.4) when $u = 0$ and $v = 1$ yield

$$\begin{aligned} A(x,0,1) &= p(B(x,0)-1)(B(x,x)-B(x,0)) \\ &\quad + (1-p)(B(x,0)-1)\left(\frac{x-x^m}{1-x}+\frac{qx^m}{1-qx}\right), \\ B(x,0) &= A(x,0,1)+\frac{1-x^m}{1-x}+\frac{qx^m}{1-qx}, \end{aligned}$$

which leads to

$$\begin{aligned} B(x,0) &= p(B(x,0)-1)(B(x,x)-B(x,0)) \\ &\quad + (1-p)(B(x,0)-1)\left(\frac{x-x^m}{1-x}+\frac{qx^m}{1-qx}\right)+\frac{1-x^m}{1-x}+\frac{qx^m}{1-qx}. \end{aligned} \tag{2.6}$$

Moreover, Lemma 2.2 gives

$$\begin{aligned} B(x,u) = \frac{px(B(x,0)-1)}{x-u}&(B(x,x)-B(x,u)) \\ &+ \frac{x(1-p)(B(x,0)-1)}{x-u}\left(\frac{(1-u)(x-x^m)-(1-x)(u-u^m)}{(1-u)(1-x)}\right. \\ &\left.+\frac{qx^m(1-qu)-qu^m(1-qx)}{(1-qu)(1-qx)}\right) \\ &+\frac{x-u+u^{m+1}(1-x)-x^{m+1}(1-u)}{(1-u)(1-x)(x-u)} \\ &+\frac{qx^{m+1}(1-qu)-qu^{m+1}(1-qx)}{(1-qu)(1-qx)(x-u)}. \end{aligned} \tag{2.7}$$

Let us denote the quantity $(1-p)x + pxB(x,0)$ by g. We apply the kernel method to (2.7) and let $u = g$, which cancels out the terms involving $B(x,u)$. Solving for $B(x,x)$ in terms of g and $B(x,0)$ in the resulting equation then gives

$$\begin{aligned} &B(x,x) \\ &= \frac{p-1}{p}\left(\frac{(1-g)(x-x^m)-(1-x)(g-g^m)}{(1-x)(1-g)}+\frac{qx^m(1-qg)-qg^m(1-qx)}{(1-qx)(1-qg)}\right) \\ &\quad+\frac{px(1-B(x,0))+g^{m+1}(1-x)-x^{m+1}(1-g)}{px(1-x)(1-g)(1-B(x,0))} \\ &\quad+\frac{qx^{m+1}(1-qg)-qg^{m+1}(1-qx)}{px(1-qx)(1-qg)(1-B(x,0))}. \end{aligned} \tag{2.8}$$

Then (2.6), together with (2.8), yields

$$\begin{aligned} &(1-p)B(x,0)+pB^2(x,0) \\ &= p(B(x,0)-1)B(x,x)+(1-p)(B(x,0)-1)\left(\frac{x-x^m}{1-x}+\frac{qx^m}{1-qx}\right) \\ &\quad+\frac{1-x^m}{1-x}+\frac{qx^m}{1-qx} \end{aligned}$$

$$
\begin{aligned}
&= (1-p)(B(x,0)-1)\left(\frac{g-g^m}{1-g}+\frac{qg^m}{1-qg}\right)\\
&\quad -\frac{p(1-B(x,0))+x^m(1-p+pB(x,0))^{m+1}(1-x)-x^m(1-g)}{(1-x)(1-g)}\\
&\quad -\frac{qx^m(1-qg)-qx^m(1-p+pB(x,0))^{m+1}(1-qx)}{(1-qx)(1-qg)}+\frac{1-x^m}{1-x}+\frac{qx^m}{1-qx}\\
&= (1-p)(B(x,0)-1)\left(\frac{g-g^m}{1-g}+\frac{qg^m}{1-qg}\right)\\
&\quad +\frac{1-g-p(1-B(x,0))-x^m(1-p+pB(x,0))^{m+1}(1-x)}{(1-x)(1-g)}\\
&\quad +\frac{qx^m(1-p+pB(x,0))^{m+1}}{1-qg}.
\end{aligned}
$$

Let $g = xf$ so that f satisfies $B(x,0) = \frac{f+p-1}{p}$. Then the last equation may be rewritten as

$$
\begin{aligned}
\frac{f(f+p-1)}{p} &= \frac{(1-p)(f-1)}{p}\left(\frac{xf-(xf)^m}{1-xf}+\frac{q(xf)^m}{1-qxf}\right)\\
&\quad +\frac{1-xf-(1-f)-(1-x)x^mf^{m+1}}{(1-x)(1-xf)}+\frac{qx^mf^{m+1}}{1-qxf}\\
&= \frac{(1-p)(f-1)}{p}\left(\frac{xf-(xf)^m}{1-xf}\right)+\frac{f+p-1}{p}\left(\frac{q(xf)^m}{1-qxf}\right)\\
&\quad +\frac{f-x^mf^{m+1}}{1-xf}\\
&= \frac{f+p-1}{p}\left(\frac{1-(xf)^m}{1-xf}+\frac{q(xf)^m}{1-qxf}\right)+\frac{(p-1)(f-1)}{p}.
\end{aligned}
$$

Dividing both sides of the last equation by $\frac{f+p-1}{p}$, we get

$$
f = \frac{1-(xf)^m}{1-xf}+\frac{q(xf)^m}{1-qxf}+\frac{(p-1)(f-1)}{f+p-1},
$$

which leads to (2.5). □

Remarks: Note that since $xf-1$ divides the polynomial $(q-1)x^mf^m - qx^2f^3 + x(1+q)f^2 - (1+qx)f + 1$ for all $m \geq 3$, it is seen that f, and hence $B(x,0)$, satisfies a polynomial of degree $m-1$. By (2.7), we have

$$
\begin{aligned}
B(x,u) &= \frac{px(B(x,0)-1)}{pxB(x,0)+(1-p)x-u}B(x,x)\\
&\quad +\frac{x(1-p)(B(x,0)-1)}{pxB(x,0)+(1-p)x-u}\left(\frac{(1-u)(x-x^m)-(1-x)(u-u^m)}{(1-u)(1-x)}\right.\\
&\qquad\qquad \left.+\frac{qx^m(1-qu)-qu^m(1-qx)}{(1-qu)(1-qx)}\right)\\
&\quad +\frac{x-u}{pxB(x,0)+(1-p)x-u}\left(\frac{x-u+u^{m+1}(1-x)-x^{m+1}(1-u)}{(1-u)(1-x)(x-u)}\right.
\end{aligned}
$$

$$+\frac{qx^{m+1}(1-qu)-qu^{m+1}(1-qx)}{(1-qu)(1-qx)(x-u)}\Bigg), \tag{2.9}$$

where $B(x,0)$ and $B(x,x)$ are determined by (2.5) and (2.8), respectively. By (2.9) and (2.3), one then has an explicit formula for $A(x,u,v)$, and hence a general solution to the system of functional equations in Lemma 2.2.

Letting $m=3$ in Theorem 2.3 yields the following explicit formula.

Corollary 2.4. *The $m=3$ case of the generating function $B(x,0)$ is given by*

$$\frac{\begin{array}{l}1+(2pq-q-1)x-(p-1)(q-1)x^2\\ \qquad\qquad -\sqrt{1-2(q+1)x-2(pq-q+p+1)x^2+x^2(q+1+(p-1)(q-1)x)^2}\end{array}}{2px(x+q-qx)}$$
$$=1+x+2x^2+(p+q+3)x^3+(2pq+q^2+4p+2q+5)x^4$$
$$+(p^2q+3pq^2+q^3+p^2+9pq+2q^2+12p+5q+8)x^5+\cdots.$$

Proof. By (2.5) and the factorization

$$x^2((q-1)x-q)f^3+x(1+q)f^2-(1+qx)f+1$$
$$=(xf-1)(x((q-1)x-q)f^2+(1+(q-1)x)f-1),$$

we have that f when $m=3$ satisfies

$$(f+p-1)(x((q-1)x-q)f^2+(1+(q-1)x)f-1)$$
$$+(p-1)(f-1)(qxf-1)=0,$$

which implies

$$x((q-1)x-q)f^2+(1+(q-1)x+(p-1)(q-1)x^2)f-1-(p-1)x=0.$$

Solving for f in this last equation, noting $f(0)=1$ and simplifying, we get

$$f=\frac{\begin{array}{l}1+(q-1)x+(p-1)(q-1)x^2\\ \qquad\qquad -\sqrt{1-2(q+1)x-2(pq-q+p+1)x^2+x^2(q+1+(p-1)(q-1)x)^2}\end{array}}{2x(x+q-qx)}.$$

The desired formula now follows from $B(x,0)=\frac{f+p-1}{p}$. □

Letting $q=1$ and $p=1$ in Corollary 2.4 gives the univariate gf formulas

$$\sum_{n\geq 0}x^n\sum_{\pi\in NC_n}p^{\mathrm{desc}(\pi)}=\frac{1+2(p-1)x-\sqrt{1-4x-4(p-1)x^2}}{2px} \tag{2.10}$$

and

$$\sum_{n\geq 0}x^n\sum_{\pi\in NC_n}q^{\tau_{123}(\pi)}=\frac{1+(q-1)x-\sqrt{(q+1)^2x^2-2(q+1)x-4x^2+1}}{2x(x+q-qx)}. \tag{2.11}$$

We have the following formula for the total number of occurrences of $12\cdots m$ subwords in all non-crossing partitions of a given length.

Corollary 2.5. *The total number of* $12\cdots m$ *subwords where* $m \geq 3$ *in all the members of* NC_n *is given by* $\binom{2n-m+1}{n+1}$ *for* $n \geq m$.

Proof. Let $F = \frac{\partial f}{\partial q}\mid_{p=q=1}$ and $C = C(x) = \frac{1-\sqrt{1-4x}}{2x}$. Then, by differentiation of both sides of (2.5) with respect to q, setting $p = q = 1$ and observing $f\mid_{p=q=1} = C$, we get

$$x^mC^m - 3x^2C^2F - x^2C^3 + 4xCF + xC^2 - (1+x)F - xC = 0.$$

Noting $xC^2 = C - 1$, and solving for F, gives

$$\begin{aligned} F &= \frac{x^mC^m}{3x^2C^2 - 4xC + 1 + x} = \frac{x^mC^m}{1 - 2x - xC} = \frac{x^mC^{m+1}}{1 - 2xC} = \frac{x^mC^{m+1}}{\sqrt{1-4x}} \\ &= x^m \sum_{n\geq 0} \binom{2n+m+1}{n} x^n = \sum_{n\geq m} \binom{2n-m+1}{n+1} x^n, \end{aligned}$$

which implies the result, where in the second-to-last equality we have made use of [17, Equation 2.5.15]. □

Remarks: By differentiating formula (2.10), one can show that there are $\binom{2n-2}{n+1}$ descents altogether in the members of NC_n, which agrees with the $m = 3$ case of Corollary 2.5. Thus, the totals for descents and 123 subwords on NC_n are the same, despite the respective distributions being different, and it would be interesting to find a bijective proof demonstrating this fact. Perhaps one could make use of a "marked" members of NC_n construction wherein a particular descent or 123 is distinguished from all others in attempting to do so. Furthermore, the formula in Corollary 2.5 continues to hold for $m = 1$ and $m = 2$ (i.e., there are $\binom{2n}{n+1}$ letters and $\binom{2n-1}{n+1}$ ascents altogether in NC_n), though these cases can be shown by more direct arguments.

Next, it should be observed that the $p = 1$ case of (2.5) when $m = 3$ and $m = 4$ coincides with equations from [14, Sequence A092107] and [12] for the gf's which enumerate the set of Dyck paths $\mathcal{D}_n$ of semilength n according to the number of occurrences of u^3 and u^4, respectively, where u denotes an up-step. Further, it is seen that the method employed in [12] can be extended to show that the corresponding gf for the statistic on $\mathcal{D}_n$ recording the number of strings of u^m where $m \geq 3$ is given by the $p = 1$ case of (2.5). This implies a general equidistribution result between NC_n and $\mathcal{D}_n$, which can also be argued directly. This argument (described below), when taken together with the method from [12], constitutes an alternative demonstration of the $p = 1$ case of (2.5).

However, the technique employed in [12], which does not make use of the kernel method, is not readily extended to prove (2.5) in its full generality where p is arbitrary. To see this, first note that the descents statistic on NC_n and the statistic recording the number of duu's on $\mathcal{D}_n$ are equally distributed, based on a comparison of the gf's. However, the joint distribution of descents and 123's on NC_n (given by the $m = 3$ case of $a(n)$ above) is different than the joint distribution for the number of duu's and uuu's on $\mathcal{D}_n$ (note, for example, when $n = 4$, one gets $5 + 4p + 2q + 2pq + q^2$ for the former and $4 + 5p + 3q + pq + q^2$ for the latter). Thus, the question remains if there exists a readily defined statistic on $\mathcal{D}_n$ which when taken jointly with the number of uuu's has the distribution given by $a(n)$ and it is unclear whether such a statistic would fit nicely into the combinatorial framework utilized in [12].

Finally, as previously mentioned, it is possible to prove the equidistribution of u^m on $\mathcal{D}_n$ with $12\cdots m$ on NC_n directly without recourse of gf's. To do so, one can demonstrate more generally for all $\ell \geq 0$ that the u^m statistic on paths of the form $u^\ell \lambda d^\ell$ where $\lambda \in \mathcal{D}_n$ has the same distribution as $12\cdots m$ on partitions $12\cdots\ell\pi$ where $\pi \in NC_n$ is expressed using letters in $\{\ell+1, \ell+2, \ldots\}$. This can be implemented by an induction on n, the $n=0$ case (for all ℓ) being clear, where one makes use of the decompositions of members of $\mathcal{D}_n$ and NC_n in terms of returns to the x-axis and runs of the letter 1. We leave the details of this argument to the interested reader.

2.2 Joint distribution of $213\cdots m$ with descents

A comparable argument can be applied to the case of $\tau = 213\cdots m$, where again we consider the joint distribution with descents. Let $d(n)$ and $d(n,i)$ be defined exactly as $a(n)$ and $a(n,i)$ above except now we are tracking occurrences of $213\cdots m$. Given $\ell \geq 0$, let $e^{(\ell)}(n)$ enumerate $\pi \in NC_n$ according to the number of descents and occurrences of τ in $\alpha 12\cdots\ell\pi$, where π is expressed using $\{\ell+1, \ell+2, \ldots\}$ and α denotes a letter that assumes the value 1.5. Let $e^\ell(n,i)$ be the restriction of $e^\ell(n)$ to members of $NC_{n,i}$. For example, when $n=3$ and $m=4$, one may verify

$$e^{(\ell)}(3,1) = \begin{cases} p^2+2p, \text{ if } \ell=0, \\ p^2q+2p, \text{ if } \ell=1, \\ p^2q+2pq, \text{ if } \ell\geq 2, \end{cases} \qquad e^{(\ell)}(3,2) = \begin{cases} p, \text{ if } \ell=0, \\ pq, \text{ if } \ell\geq 1, \end{cases}$$

and

$$e^{(\ell)}(3) = \begin{cases} p^2+pq+3p, \text{ if } \ell=0, \\ p^2q+2pq+2p, \text{ if } \ell=1, \\ p^2q+4pq, \text{ if } \ell\geq 2. \end{cases}$$

We have the following system of linear recurrences satisfied by the various arrays.

Lemma 2.6. *Let $\ell \geq 0$ and $m \geq 3$. If $n \geq 2$, then*

$$d(n,i) = \sum_{j=i+2}^{n} d(j-i-1)e^{(0)}(n-j+1) + d(n-i), \quad 1 \leq i \leq n-1, \tag{2.12}$$

with $d(n) = \sum_{i=1}^{n-1} d(n,i) + 1$ for $n \geq 1$ and $d(0)=1$. If $n \geq 2$, then

$$e^{(\ell)}(n,i) = \sum_{j=i+2}^{n} e^{(\ell+i)}(j-i-1)e^{(0)}(n-j+1) + e^{(\ell+i)}(0)d(n-i), \quad 1 \leq i \leq n-1, \tag{2.13}$$

with

$$e^{(\ell)}(n) = \sum_{i=1}^{n-1} e^{(\ell)}(n,i) + \begin{cases} p, \text{ if } n+\ell < m-1, \\ pq, \text{ if } n+\ell \geq m-1, \end{cases}$$

for $n \geq 1$ and $e^{(\ell)}(0) = \begin{cases} 1, \text{ if } \ell=0, \\ p, \text{ if } 1\leq \ell < m-1, \\ pq, \text{ if } \ell \geq m-1. \end{cases}$

Proof. We modify somewhat the proof given above for Lemma 2.1. For (2.12), note that $\pi \in NC_{n,i}$ may be decomposed as $\pi = 12\cdots i\rho i\rho'$, where ρ is of length $j-i-1$ for some $j \in [i+1,n]$ and contains no 1's. If $j > i+1$, then ρ is nonempty and hence the section $i\rho'$ of π is enumerated by $e^{(0)}(n-j+1)$, as there is an extra descent between the second i and the letter directly preceding it. If $j = i+1$, i.e., $\rho = \varnothing$, then there is no such extra descent and thus there are $d(n-i)$ possible π in this case. Considering all possible j implies (2.12). The formula for $d(n)$ follows from the definitions, upon including the increasing partition which has weight 1. A comparable argument to that given for (2.12) applies to (2.13), upon noting that the section ρ in this case is accounted for by $e^{(\ell+i)}(j-i-1)$, instead of $d(j-i-1)$, when $j > i+1$. Further, if $j = i+1$, then an initial increasing run of length $\ell+i$ is directly followed by $\ell+i$, which yields $e^{(\ell+i)}(0)d(n-i)$ possibilities. The formula for $e^{(\ell)}(n)$ for $n \geq 1$ follows from the definitions, upon taking into account the sequence $\alpha 12\cdots(\ell+n)$, which has a single descent for all $\ell \geq 0$ as $n \geq 1$ and contains an occurrence of $213\cdots m$ if and only if $\ell+n \geq m-1$. If $n = 0$, then the same reasoning applies except that one must consider separately the cases of zero and nonzero ℓ. □

Define $D(x,v) = \sum_{n\geq 2}\sum_{i=1}^{n-1} d(n,i)v^{i-1}x^n$, $D(x) = \sum_{n\geq 0} d(n)x^n$, $E(x,u,v) = \sum_{n\geq 2}\sum_{\ell\geq 0}\sum_{i=1}^{n-1} e^{(\ell)}(n,i)v^{i-1}u^{\ell}x^n$ and $E(x,u) = \sum_{n\geq 0}\sum_{\ell\geq 0} e^{(\ell)}(n)u^{\ell}x^n$. Rewriting the recurrences in Lemma 2.6 in terms of the gf's (we omit the details, the proof being similar to that of Lemma 2.2 above) yields the following system of functional equations.

Lemma 2.7. *We have*

$$D(x) = \frac{1}{1-x} + D(x,1), \tag{2.14}$$

$$D(x,v) = \frac{x}{1-vx}E(x,0)(D(x)-1), \tag{2.15}$$

$$\begin{aligned} E(x,u,v) &= \frac{x}{vx-u}(E(x,0)-1)(E(x,vx)-E(x,u)) \\ &\quad + \frac{px(D(x)-E(x,0))}{1-vx}\cdot\frac{vx-u+(1-q)(u^{m-1}(1-vx)-(vx)^{m-1}(1-u))}{(1-u)(vx-u)}, \end{aligned} \tag{2.16}$$

$$\begin{aligned} E(x,u) &= E(x,u,1) + \frac{px(1+(q-1)u^{m-2})}{(1-u)(1-x)} \\ &\quad + \frac{p(q-1)x^2(x^{m-2}-u^{m-2})}{(1-x)(x-u)} + \frac{1+(p-1)u+p(q-1)u^{m-1}}{1-u}. \end{aligned} \tag{2.17}$$

Solving explicitly the system of functional equations in the preceding lemma gives the following formula for $D(x)$.

Theorem 2.8. *The generating function $D(x)$ enumerating members of NC_n jointly according to the number of descents and occurrences of $213\cdots m$ for $m \geq 3$ is given by $D(x) = \frac{1-xf}{1-x-xf}$, where $f = E(x,0)$ satisfies*

$$p(q-1)x^{m-1}f^{m-2} + xf^2 - f + 1 + (p-1)x = 0. \tag{2.18}$$

Proof. By (2.14) and (2.15) with $v = 1$, we obtain $D(x) - \frac{1}{1-x} = \frac{x}{1-x}E(x,0)(D(x)-1)$, which implies the first statement. By (2.16) with $v = 1$ and (2.17), one obtains an equation relating $E(x,u)$ and $E(x,x)$. We apply the kernel method to this equation and set $u =$

$xE(x,0)$, which cancels out the $E(x,u)$ terms and leads to the formula

$$E(x,x) = \frac{px(1+(q-1)(xf)^{m-2})}{(1-x)(1-xf)} + \frac{p(q-1)x^{m-1}(1-f^{m-2})}{(1-x)(1-f)} + \frac{1+(p-1)xf+p(q-1)(xf)^{m-1}}{1-xf} + \frac{px(1-f+xf^2)(1-f+(1-q)x^{m-2}((1-x)f^{m-1}+xf-1))}{(1-x)(1-f)(1-xf)(1-x-xf)}. \tag{2.19}$$

By (2.16) and (2.17) when $u=0, v=1$, we have

$$f^2 = (f-1)E(x,x) + \frac{px(1-f+xf^2)(1+(q-1)x^{m-2})}{(1-x)(1-x-xf)} + \frac{1+(p-1)x+p(q-1)x^{m-1}}{1-x}. \tag{2.20}$$

By (2.19) and (2.20), one obtains

$$\begin{aligned} f^2 &= \frac{px(f-1)(1+(q-1)(xf)^{m-2})}{(1-x)(1-xf)} - \frac{p(q-1)x^{m-1}(1-f^{m-2})}{1-x} \\ &\quad + \frac{(f-1)(1+(p-1)xf+p(q-1)(xf)^{m-1})}{1-xf} \\ &\quad - \frac{px(1-f+xf^2)(1-f+(1-q)x^{m-2}((1-x)f^{m-1}+xf-1))}{(1-x)(1-xf)(1-x-xf)} \\ &\quad + \frac{px(1-f+xf^2)(1+(q-1)x^{m-2})}{(1-x)(1-x-xf)} \\ &\quad + \frac{1+(p-1)x+p(q-1)x^{m-1}}{1-x} \\ &= \frac{px(f-1)(1+(q-1)(xf)^{m-2})}{(1-x)(1-xf)} + \frac{1+(p-1)x+p(q-1)x^{m-1}f^{m-2}}{1-x} \\ &\quad + \frac{(f-1)(1+(p-1)xf+p(q-1)(xf)^{m-1})}{1-xf} \\ &\quad + \frac{px(1-f+xf^2)(f+(q-1)x^{m-2}f^{m-1})}{(1-xf)(1-x-xf)} \\ &= \frac{f+(p-1)xf^2+p(q-1)x^{m-1}f^m}{1-xf} \\ &\quad + \frac{px(1-f+xf^2)(f+(q-1)x^{m-2}f^{m-1})}{(1-xf)(1-x-xf)}, \end{aligned}$$

after several algebraic steps, and hence

$$(1-xf)f = 1+(p-1)xf+p(q-1)(xf)^{m-1} + \frac{px(1-f+xf^2)(1+(q-1)(xf)^{m-2})}{1-x-xf}.$$

The last equation may be rewritten as

$$\begin{aligned}&p(q-1)x^{m-1}f^{m-2}(1-xf)+1+(p-1)x-(1+x+(p-1)x^2)f+2xf^2-x^2f^3\\&=(1-xf)(p(q-1)x^{m-1}f^{m-2}+1+(p-1)x-f+xf^2)=0,\end{aligned}$$

which implies (2.18) and completes the proof. □

Taking $m = 3$ and $m = 4$ in Theorem 2.8, and solving for $D(x)$ explicitly, yields the following result.

Corollary 2.9. *The $m = 3$ and $m = 4$ cases of the generating function $D(x)$ are given respectively by*

$$\begin{aligned}&\frac{1+2(pq-1)x+p(1-q)x^2-\sqrt{1-4x+2(2-p-pq)x^2+p^2(1-q)^2x^4}}{2px(q+(1-q)x)}\\&=1+x+2x^2+(p+4)x^3+(pq+5p+8)x^4+(p^2q+p^2+6pq+18p+16)x^5+\cdots\end{aligned}$$

and

$$\begin{aligned}&\frac{1-2(1-pq)x+2p(1-q)x^2-\sqrt{1-4x+4(1-p)x^2+4p(1-q)(1-(1-p)x)x^3}}{2px(q+(1-q)(2-x)x)}\\&=1+x+2x^2+(p+4)x^3+(6p+8)x^4+(2p^2+pq+23p+16)x^5+\cdots.\end{aligned}$$

Note that the $p = 1$ case of the formula above for $D(x)$ when $m = 3$, which corresponds to counting 213 subwords in members of NC_n, is obtained as a particular case of a different family of patterns below (see Theorem 3.3).

Differentiation of the formula for $D(x)$ given in Theorem 2.8 with respect to q leads to the following result.

Corollary 2.10. *The total number of $213\cdots m$ subwords where $m \geq 3$ in all the members of NC_n is given by $\binom{2n-m}{n+1}$ for $n \geq m+1$.*

3 Other subword patterns

In this section, we consider the distribution of three further infinite families of subword patterns, namely, $23\cdots m1$, $34\cdots m21$ and $23\cdots(m-1)1m$ where $m \geq 3$. To establish our results in these cases, we make use of the symbolic enumeration method. In each case, we let $F_\tau(x)$ denote the gf which enumerates the non-crossing partitions of length n for all $n \geq 0$ according to the number of occurrences of the subword pattern τ in question, where x marks the length of a partition and q the number of occurrences of τ.

Theorem 3.1. *The generating functions that enumerate the non-crossing partitions according to the number of occurrences of $\tau = 23\cdots m1$ and $\tau = 34\cdots m21$ where $m \geq 3$ are given respectively by*

$$\frac{1+2(q-1)x^{m-1}-\sqrt{1-4x-4(q-1)x^m}}{2x(1+(q-1)x^{m-2})}$$

and

$$\frac{1+(q-1)(2-3x)x^{m-2}-\sqrt{(1-(q-1)x^{m-1})(1-4x-(q-1)x^{m-1})}}{2x(1+(q-1)(1-2x)x^{m-3})}.$$

Proof. Let $F = F_\tau(x)$ for the pattern τ at hand. To aid in finding F, we define the auxiliary generating function $G = G_\tau(x)$ which counts non-crossing partitions π according to the number of occurrences of τ in $\pi 0$. For either pattern τ, we have

$$F = \frac{1}{1-x} + \frac{x}{1-x}(F-1) + \frac{x}{1-x}(G-1)(F-1). \tag{3.1}$$

To see this, note that $\pi \in NC_n$ for some $n \geq 0$ may be expressed as

(i) $\pi = 1^n$,

(ii) $\pi = 1^m \rho$, where $m \geq 1$ and $\rho \neq \varnothing$ contains no 1's, or

(iii) $\pi = 1^m \rho 1 \sigma$, where $m \geq 1$, ρ is as before and 1σ is a nonempty non-crossing partition on the letters in $\{1\} \cup \{r+1, r+2, \ldots\}$ where $r = \max(\rho)$.

Then the three terms on the right-hand side of (3.1) account for the cases (i) – (iii), respectively; note that in (iii), the $G-1$ factor accounts for the nonempty section ρ being directly followed by a 1.

Now assume $\tau = 23\cdots m1$. Then F and G in this case assign all partitions the same weight except for those of the form $\pi = \pi'(i+1)(i+2)\cdots(i+m-2)$ wherein π' is a partition with i blocks for some $i \geq 1$, which implies

$$G = F + x^{m-2}(q-1)(F-1). \tag{3.2}$$

By (3.1), we have $G = \frac{(1-x)F-1}{x(F-1)}$. Equating this with (3.2) gives

$$x(1+(q-1)x^{m-2})F^2 - (1+2(q-1)x^{m-1})F + 1 + (q-1)x^{m-1} = 0,$$

and solving for F yields the formula above for $\tau = 23\cdots m1$.

To prove the formula for $\tau = 34\cdots m21$, first note the gf counting the members of NC_n for $n \geq 3$ whose last two letters form a descent according to the number of occurrences of τ is given by $F - 1 - x - 2x(F-1) = (1-2x)(F-1) - x$, by subtraction. Consider inserting an increasing string of length $m-3$ directly prior to the last letter within a non-crossing partition whose last two letters form a descent, and note that F and G differ (by one) with respect to the assigned q-weights on the partitions resulting from such an insertion and no others. This implies the relation

$$G = F + x^{m-3}(q-1)((1-2x)(F-1) - x). \tag{3.3}$$

To solve (3.1) and (3.3) simultaneously, it is easier to first find $U := F - 1$. Note that (3.1) and (3.3) imply

$$U = \frac{x}{1-x}(U+1) + \frac{x(1+(q-1)(1-2x)x^{m-3})}{1-x}U^2 - \frac{(q-1)x^{m-1}}{1-x}U,$$

i.e., $x(1+(q-1)(1-2x)x^{m-3})U^2 - (1-2x+(q-1)x^{m-1})U + x = 0$. We then get

$$\begin{aligned} U &= \frac{1-2x+(q-1)x^{m-1} - \sqrt{(1-2x+(q-1)x^{m-1})^2 - 4x^2(1+(q-1)(1-2x)x^{m-3})}}{2x(1+(q-1)(1-2x)x^{m-3})} \\ &= \frac{1-2x+(q-1)x^{m-1} - \sqrt{(1-(q-1)x^{m-1})(1-4x-(q-1)x^{m-1})}}{2x(1+(q-1)(1-2x)x^{m-3})}. \end{aligned}$$

Adding 1 to the last expression yields F and hence the second formula above, which completes the proof. □

We have the following corresponding formulas for the totals on NC_n of the preceding subword patterns.

Corollary 3.2. *If $n \geq m-1$, then there are $\binom{2n-2m+2}{n-m+3}$ and $\binom{2n-2m+2}{n-m+4}$ total occurrences of $23\cdots m1$ and $34\cdots m21$, respectively, within all the members of NC_n.*

Proof. Let $C(x) = \sum_{n\geq 0} C_n x^n = \frac{1-\sqrt{1-4x}}{2x}$. For $\tau = 34\cdots m21$, we have

$$\begin{aligned}\frac{\partial}{\partial q}F(x)\mid_{q=1} &= \left(\frac{(2-3x)x^{m-2}+\frac{x^{m-1}(1-2x)}{\sqrt{1-4x}}}{1-\sqrt{1-4x}} - (1-2x)x^{m-3}\right)C(x)\\ &= \frac{1}{2}x^{m-2} + \frac{x^{m-2}(1-2x)}{2\sqrt{1-4x}} - x^{m-3}(1-2x)(C(x)-1).\end{aligned}$$

If $r = n-m+2$, then

$$\begin{aligned}[x^n]\left(\frac{\partial}{\partial q}F(x)\mid_{q=1}\right) &= \frac{1}{2}\binom{2r}{r} - \binom{2r-2}{r-1} - C_{r+1} + 2C_r\\ &= \frac{(r-1)(r-2)(r-3)}{r(r+1)(r+2)}\binom{2r-2}{r-1} = \binom{2r-2}{r+2},\end{aligned}$$

which establishes the second formula. A similar proof applies to the first formula where there are only two terms contributing to the total in this case. □

Remarks: Note that for a fixed m, the sequences of nonzero values obtained from the expressions for the totals in Corollary 3.2 coincide respectively with entries A002694 and A002696 in [14]. Further, the first formula in Theorem 3.1 is seen to hold also for $m = 2$ and agrees with the formula (2.10) found above for the descents statistic. As mentioned above, descents on NC_n has distribution equal to that of the statistic on $\mathcal{D}_n$ recording the number of duu's; to see this, compare the functional equation satisfied by the gf in (2.10) with that found for the duu parameter on $\mathcal{D}_n$ in [2, Section 6.17]. Indeed, both parameters have what is known as the Touchard distribution on their respective structures, which is given by $\sum_{k=0}^{\lfloor (n-1)/2\rfloor}\binom{n-1}{2k}C_k 2^{n-1-2k}p^k$. We leave it as a question for the reader to find a direct bijective proof of this equivalence.

We now consider the distribution for the $23\cdots(m-1)1m$ family of subwords.

Theorem 3.3. *The generating function that enumerates the non-crossing partitions according to the number of occurrences of $\tau = 23\cdots(m-1)1m$ where $m \geq 3$ is given by*

$$\frac{1+(q-1)(2-x)x^{m-2} - \sqrt{(1-(q-1)x^{m-1})^2 - 4x}}{2x(1+(q-1)(1-x)x^{m-3})}.$$

Proof. To find F for $\tau = 23\cdots(m-1)1m$, let G denote the gf that counts $\pi \in NC_n$ for some $n \geq 0$ according to the number of occurrences of τ in $\pi 0(n+1)$. Note that F and G differ in their assignment of q-weights only on partitions of the form $\rho = \rho'(i+1)(i+2)\cdots(i+m-3)$, where i is a partition with i blocks for some $i \geq 1$, which implies

$$G = F + x^{m-3}(q-1)(F-1). \tag{3.4}$$

To write a formula for F, note that $\pi \in NC_n$ for some $n \geq 0$ implies it may be expressed as

(i) $\pi = 1^n$,

(ii) $\pi = 1^m\rho$, where $1 \leq m \leq n-1$,

(iii) $\pi = 1^m\rho 1$, where $1 \leq m \leq n-2$,

(iv) $\pi = 1^m\rho 1\rho'$, where $1 \leq m \leq n-3$, or

(v) $\pi = 1^m\rho 1^j\rho'$, where $1 \leq m \leq n-3$ and $2 \leq j \leq n-m-1$, and it is assumed that ρ contains no 1's and the first letter of ρ' is not 1.

Further, ρ and ρ' are nonempty in all cases, except for possibly (v) where ρ' may be empty. Note that the gf enumerating the section $1\rho'$ of π described in case (iv) in which the smallest letter occurs in an initial run of length one is given by $F-1-x-x(F-1) = (1-x)F-1$, by subtraction. Also, the nonempty section ρ in (iv) is accounted for by $G-1$, as it is directly followed by a 1 followed by $\max(\rho)+1$. Thus, combining (i) – (v) implies F satisfies

$$F = \frac{1}{1-x} + \frac{x+x^2}{1-x}(F-1) + \frac{x}{1-x}(G-1)((1-x)F-1) + \frac{x^2}{1-x}(F-1)^2. \quad (3.5)$$

By (3.4) and (3.5), we have that F satisfies

$$x(1+(q-1)(1-x)x^{m-3})F^2 - (1+(q-1)(2-x)x^{m-2})F + 1 + (q-1)x^{m-2} = 0.$$

Solving for F, and simplifying, yields the stated formula. □

Differentiation of the preceding gf formula with respect to q gives the following result.

Corollary 3.4. *If $n \geq m-1$, then the total number of occurrences of $23\cdots(m-1)1m$ within all the members of NC_n is given by $\binom{2n-2m+3}{n-m+4}$.*

Remarks: The distribution of 213 on NC_n is the same as that of the number of $dduu$'s on $\mathcal{D}_n$, which follows from comparing the $m=3$ case of Theorem 3.3 with the gf formula given for entry [14, Seq. 114492]. This can also be shown directly by an inductive argument making use of the decompositions of NC_n and $\mathcal{D}_n$ according to the occurrences of 1 and returns to the x-axis, respectively. Further, the sequence for the total number of occurrences of $23\cdots(m-1)1m$ in members of NC_n for $n \geq m+1$ corresponds to [14, Sequence A003516]. Let $t(\tau)$ denote the number of occurrences of the subword τ in all the members of NC_n. Then a comparison of the results above reveals $t(12\cdots m) > t(213\cdots m) \geq t(23\cdots(m-1)1m) > t(23\cdots m1) > t(34\cdots m21)$ for all $m \geq 3$ and $n \geq m+2$, which is roughly in accord with intuition concerning the relative frequency of occurrence of the various patterns. Note that there is equality in the second inequality if and only if $m=3$.

ORCID iDs

Toufik Mansour https://orcid.org/0000-0001-8028-2391
Mark Shattuck https://orcid.org/0000-0001-7441-1759

References

[1] A. Burstein and T. Mansour, Counting occurrences of some subword patterns, *Discrete Math. Theor. Comput. Sci.* **6** (2003), 1–11.

[2] E. Deutsch, Dyck path enumeration, *Discrete Math.* **204** (1999), 167–202, doi: 10.1016/S0012-365X(98)00371-9, https://doi.org/10.1016/S0012-365X(98)00371-9.

[3] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2009, doi:10.1017/CBO9780511801655, https://doi.org/10.1017/CBO9780511801655.

[4] Q.-H. Hou and T. Mansour, Kernel method and linear recurrence system, *J. Comput. Appl. Math.* **216** (2008), 227–242, doi:10.1016/j.cam.2007.05.001, https://doi.org/10.1016/j.cam.2007.05.001.

[5] M. Klazar, On $abab$-free and $abba$-free set partitions, *European J. Comb.* **17** (1996), 53–68, doi:10.1006/eujc.1996.0005, https://doi.org/10.1006/eujc.1996.0005.

[6] Z. Lin and S. Fu, On 1212-avoiding restricted growth functions, *Electron. J. Comb.* **24** (2017), Paper No. 1.53, 20.

[7] T. Mansour and M. Shattuck, Counting subwords in flattened involutions and Kummer functions, *J. Difference Equ. Appl.* **22** (2016), 1404–1425, doi:10.1080/10236198.2016.1199689, https://doi.org/10.1080/10236198.2016.1199689.

[8] T. Mansour and M. Shattuck, Visibility in non-crossing and non-nesting partitions, *J. Difference Equ. Appl.* **27** (2021), 354–375, doi:10.1080/10236198.2021.1897117, https://doi.org/10.1080/10236198.2021.1897117.

[9] T. Mansour, M. Shattuck and S. Wagner, Counting subwords in flattened partitions of sets, *Discrete Math.* **338** (2015), 1989–2005, doi:10.1016/j.disc.2015.04.023, https://doi.org/10.1016/j.disc.2015.04.023.

[10] T. Mansour, M. Shattuck and S. H. F. Yan, Counting subwords in a partition of a set, *Electron. J. Comb.* **17** (2010), Research Paper 19, 21, doi:10.37236/291, https://doi.org/10.37236/291.

[11] S. C. Milne, A q-analog of restricted growth functions, Dobinski's equality, and Charlier polynomials, *Trans. Am. Math. Soc.* **245** (1978), 89–118, doi:10.2307/1998858, https://doi.org/10.2307/1998858.

[12] A. Sapounakis, I. Tasoulas and P. Tsikouras, Counting strings in Dyck paths, *Discrete Math.* **307** (2007), 2909–2924, doi:10.1016/j.disc.2007.03.005, https://doi.org/10.1016/j.disc.2007.03.005.

[13] R. Simion, Combinatorial statistics on non-crossing partitions, *J. Comb. Theory Ser. A* **66** (1994), 270–301, doi:10.1016/0097-3165(94)90066-3, https://doi.org/10.1016/0097-3165(94)90066-3.

[14] N. J. A. Sloane et al., *The On-Line Encyclopedia of Integer Sequences*, 2019, https://oeis.org.

[15] D. Stanton and D. White, *Constructive Combinatorics*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1986, doi:10.1007/978-1-4612-4968-9, https://doi.org/10.1007/978-1-4612-4968-9.

[16] C. G. Wagner, Generalized Stirling and Lah numbers, *Discrete Math.* **160** (1996), 199–218, doi:10.1016/0012-365X(95)00112-A, https://doi.org/10.1016/0012-365X(95)00112-A.

[17] H. S. Wilf, *generatingfunctionology*, Academic Press, Inc., Boston, MA, 1990.

[18] F. Yano and H. Yoshida, Some set partition statistics in non-crossing partitions and generating functions, *Discrete Math.* **307** (2007), 3147–3160, doi:10.1016/j.disc.2007.03.050, https://doi.org/10.1016/j.disc.2007.03.050.

[19] H. Zhao and Z. Zhong, Two statistics linking Dyck paths and non-crossing partitions, *Electron. J. Comb.* **18** (2011), Paper 83, 12.