ENUMERATION OF SUBWORDS IN PERMUTATIONS

SERGI ELIZALDE AND MARC NOY

Abstract

In this paper we study the distribution of the number of occurrences of a permutation σ as a subword among all permutations in $\mathcal{S}_n$. We solve the problem in several cases depending on the shape of σ by obtaining the corresponding bivariate exponential generating functions as solutions of certain linear differential equations with polynomial coefficients.

Resumen

En este artículo estudiamos la distribución del número de ocurrencias de una permutación σ como subpalabra entre todas las permutaciones de $\mathcal{S}_n$. Resolvemos el problema en varios casos, dependiendo de la forma de σ, al obtener las funciones generadoras exponenciales bivariadas correspondientes como soluciones de ciertas ecuaciones diferenciales lineales con coeficientes polinómicos.

1. Introduction

Let m, n be two positive integers with $m \leq n$, and let $\pi = (\pi_1, \pi_2, \ldots, \pi_n) \in \mathcal{S}_n$ and $\sigma \in \mathcal{S}_m$ be two permutations. We say that π *contains* σ *as a subword* if there exist m *consecutive* elements $\pi_{l+1}, \ldots, \pi_{l+m}$ such that $\rho(\pi_{l+1}, \ldots, \pi_{l+m}) = \sigma$, where ρ is the reduction consisting in relabeling the elements with $\{1, \ldots, m\}$ so that they keep the same order relationships they had in π.

For example, if $\sigma = 4132 \in \mathcal{S}_4$, then $\pi = 6725341 \in \mathcal{S}_7$ contains σ as a subword, because $\rho(7253) = 4132$. However, $\pi = 41352 \in \mathcal{S}_5$ does not contain σ as a subword (even though it contains it as a subsequence, that is, in non-consecutive positions); in this case we say that π *avoids* σ. Occurrences of a subword can be overlapped, for instance, 5716243 contains two occurrences of σ, namely 7162 and 6243.

Denote by $A_n(\sigma)$ the set of permutations of $\mathcal{S}_n$ that do not contain σ as a subword, and let $\alpha_n(\sigma) = |A_n(\sigma)|$. If we want to exclude several subwords $\sigma, \tau, \ldots$ we use the corresponding notations $A_n(\sigma, \tau, \ldots)$ and $\alpha_n(\sigma, \tau, \ldots)$. Our main purpose is to compute $\alpha_n(\sigma)$ for a given subword σ. More generally, we are also interested in the distribution of the number of occurrences of σ among all permutations in $\mathcal{S}_n$.

Some well-known counting problems in permutations can be stated in terms of forbidden subwords. For instance, occurrences of 12 correspond to ascents and are counted by Eulerian numbers; up-and-down permutations are those in $A_n(123, 321)$; permutations whose longest increasing run is at most $k-1$ correspond to $A_n(12\cdots k)$; and it can be seen that $\alpha_n(123, 132)$ is precisely the number of involutions. Related results for occurrences of subwords of length three can be found in [1].

The organization of the paper is as follows. First we present some preliminaries on the representation of permutations as increasing trees, on the symbolic method for dealing with exponential generating functions (EGF), and on asymptotic enumeration. In Section 3 we enumerate occurrences of a subword σ in three cases. First when the two smallest (or the two largest) elements of σ are at the beginning and at the end; secondly, when σ is totally increasing (or decreasing); and finally when $\sigma = 1, 2, \ldots, a-1, a, \tau, a+1$, and τ is any permutation of the elements $\{a+2, a+3, \ldots, m+2\}$. Notice that the fact that τ is arbitrary means that our result covers a very large number of subwords σ. Note also that the first case is a particular case of the last one ($a = 1$), but we have decided

Research partially supported by projects SEUI-PB98-0933 and CUR Gen. Cat. 1999SGR00356.

to include a separate proof since it illustrates the different ways of representing permutations as increased trees, as explained below.

In Section 4 we show how these results specialize in the case of subwords of length three and four. Additional results and considerations on multiple subwords of length three and on the moments and the distribution of occurrences of arbitrary subwords can be found in the full paper [3].

The basis of our work is the use of symbolic methods for specifying combinatorial classes, following the approach described in the books by Flajolet and Sedgewick [6] and [8].

The key point is the representation of permutations both as binary increasing trees and as unordered increasing trees. From these representations, using symbolic methods, we derive differential equations satisfied by the corresponding exponential generating functions. In all cases we have encountered, the differential equations become linear after a suitable substitution.

A similar approach is taken in [4] for counting occurrences of a given subtree in binary search trees. The main difference in our case is however that patterns corresponding to forbidden subwords may be split into two subtrees in several different ways. Finally, the reader can compare the present paper with [5], where a similar approach is taken for analyzing certain geometric configurations, the difference being that here we deal with exponential GFs that are transcendental instead of being algebraic.

Note. As remarked by one of the referees, the main results in this paper can also be obtained by using the results from Section 2.8 in [7]. However, we believe that the approach used here can be valuable in the enumeration of further cases of the general problem.

2. Preliminaries

Tree representations of permutations. Following Stanley [9, Chapter 1], we represent a permutation as a tree in two different ways. Each of them establishes a bijection between permutations and certain labeled increasing trees, that is, trees in which the labels along any path from the root are increasing. These bijections allow us to translate combinatorial properties of permutations (such as avoiding a certain subword) into combinatorial properties of trees, which can be handled more conveniently.

First Representation. Let $\pi = a_1 a_2 \cdots a_n$ be a word on the alphabet of positive integers with no repeated letters. Define a binary tree $T(\pi)$ as follows. If $\pi = \emptyset$, then $T(\pi) = \emptyset$. Otherwise, let i be the least element of π, so that π can be factored uniquely in the form $\pi = \sigma i \tau$. Now define $T(\pi)$ by induction as the tree with root i, and having $T(\sigma)$ and $T(\tau)$ as left and right subtrees, respectively. This correspondence gives a bijection between $\mathcal{S}_n$ and the set of *increasing binary trees* on n vertices. In particular, we see that the number of such trees is $n!$.

Second Representation. Another correspondence, now between $\mathcal{S}_n$ and the set of *unoriented increasing trees* on $n+1$ vertices (unoriented means that the children of any node are unordered), can be described as follows. Given $\pi = a_1 a_2 \cdots a_n \in \mathcal{S}_n$, its associated tree $T'(\pi)$ has root 0, and vertices $1, \ldots, n$ are arranged by defining vertex i to be a successor of the nearest element j of π to the left of i that is less than i. If there is no such j, then i is a successor of the root 0. This is again a bijection, which implies in particular that the number of unoriented increasing trees on $n+1$ vertices is $n!$.

Labeled classes and exponential generating functions. Here we direct the reader to the books by Flajolet and Sedgewick [6] and [8] for a detailed account on combinatorial classes and the symbolic method. Let $\mathcal{A}$ be a class of labeled combinatorial objects and let $|\alpha|$ be the size of an object $\alpha \in \mathcal{A}$. If $\mathcal{A}_n$ denotes the objects in $\mathcal{A}$ of size n and $a_n = |\mathcal{A}_n|$, then the *exponential generating function*, EGF for short, of the class $\mathcal{A}$ is

$$A(z) = \sum_{\alpha \in \mathcal{A}} \frac{z^{|\alpha|}}{|\alpha|!} = \sum_{n \geq 0} a_n \frac{z^n}{n!}.$$

In our context, the size of a permutation is simply its length.

There is a direct correspondence between set-theoretic operations (or "constructions") on combinatorial classes and algebraic operations on EGFs. Table 1 summarizes this correspondence for the operations that are used in the paper. There "union" means union of disjoint copies, "labeled product" is the usual cartesian product enriched with the relabeling operation, and "set" forms sets in the usual sense. Particularly important for us is the construction "boxed product" $\mathcal{A} = \mathcal{B}^{\square} * \mathcal{C}$, which corresponds to the subset of $\mathcal{B} \star \mathcal{C}$ (the usual labeled product) formed by those pairs in which the smallest label lies in the $\mathcal{B}$ component.

TABLE 1. The basic combinatorial constructions and their translation into exponential generating functions.

Construction		*Operation on GF*
Union	$\mathcal{A} = \mathcal{B} \cup \mathcal{C}$	$A(z) = B(z) + C(z)$
Labeled product	$\mathcal{A} = \mathcal{B} \star \mathcal{C}$	$A(z) = B(z)C(z)$
Set	$\mathcal{A} = \Pi(\mathcal{B})$	$A(z) = \exp(B(z))$
Boxed product	$\mathcal{A} = \mathcal{B}^{\square} \star \mathcal{C}$	$A(z) = \int_0^z (\frac{d}{dt}B(t)) \cdot C(t)dt$

Enumerations according to size and an auxiliary parameter χ are described by bivariate generating functions, or BGFs,

$$A(u, z) = \sum_{\alpha \in \mathcal{A}} u^{\chi(\alpha)} \frac{z^{|\alpha|}}{|\alpha|!} = \sum_{n,k \geq 0} A_{n,k} u^k \frac{z^n}{n!},$$

with $A_{n,k}$ the number of objects of size n with χ-parameter equal to k. Throughout the paper the variable z is reserved for marking the length of a permutation, and the variable u is used mostly for marking occurrences of a subword.

Generating functions for increasing trees. We apply the former machinery to increasing trees. The following derivations will be used repeatedly.

Let $\mathcal{I}^0$ be the labeled class of (possibly empty) binary increasing trees. It satisfies the recursive definition

$$\mathcal{I}^0 = \{\epsilon\} + (\{1\}^{\square} \star \mathcal{I}^0 \star \mathcal{I}^0)$$

where ϵ is the empty tree, $\{1\}$ represents the tree with one single node, and the box indicates that the root contains the smallest label. So, the equation for the EGF is

$$I^0(z) = 1 + \int_0^z I^0(t)^2 dt$$

which reduces to $I^{0\prime}(z) = I^0(z)^2$ with initial condition $I^0(0) = 1$ (derivatives are always with respect to z), admitting the solution $I^0(z) = 1/(1 - z)$. Thus, $I^0_n = n!$ as expected.

Let now $\mathcal{T}$ be the class of unoriented increasing trees. Since the subtrees attached to the root are not ordered, we use the set construction and get

$$\mathcal{T} = \{1\}^{\square} \star \Pi(\mathcal{T}) \tag{1}$$

In terms of the EGF it translates into

$$T(z) = \int_0^z e^{T(t)} dt$$

Differentiating, $T'(z) = \exp(T(z))$, with $T(0) = 0$, hence $T(z) = \ln(1/(1 - z)) = -\ln(1 - z)$. Thus $T'(z) = \frac{1}{1-z}$, in accordance with the fact that there are $n!$ unoriented increasing trees on $n + 1$ vertices.

Equivalent subwords. We say that two subwords $\sigma, \tau \in \mathcal{S}_m$ are *equivalent* if the BGF counting occurrences of σ and τ are the same. We write $\sigma \sim \tau$ to denote equivalence.

There are two simple operations that give equivalent subwords: *reversal*, which transforms $\sigma = \sigma_1 \cdots \sigma_m$ into $\overline{\sigma} = \sigma_m \cdots \sigma_1$, and *complementation*, transforming σ into $\widetilde{\sigma} = (m+1-\sigma_1)\cdots(m+1-\sigma_m)$. The explanation is that a permutation π has as many occurrences of σ as $\overline{\pi}$ has of $\overline{\sigma}$, and as $\widetilde{\pi}$ has of $\widetilde{\sigma}$.

Asymptotic enumeration. Let $A(z)$ be a meromorphic function on a domain of the complex plane including the origin, and let ρ be a pole of $A(z)$ such that $|\rho|$ is minimum. Then the following asymptotic estimate holds:

$$[z^n]A(z) \sim \gamma \cdot \rho^{-n},$$

where γ is the residue of A in ρ. See [6, Chapter 4] for a discussion.

3. Main results

This section contains the main results of the paper. We obtain the counting BGF of occurrences of a subword σ in three cases.

Theorem 3.1. *Let m be a positive integer, let $\sigma = 2, m+2, m+1, \ldots, 4, 3, 1 \in \mathcal{S}_{m+2}$, and let $P(u,z)$ be the BGF of permutations where u marks the number of occurrences of the subword σ. Then, $P(u,z) = 1/\omega(u,z)$, where ω is the solution of*

$$\omega'' + (1-u)\frac{z^m}{m!}\omega' = 0 \tag{2}$$

with $\omega(0) = 1$, $\omega'(0) = -1$, namely

$$P(u,z) = \frac{1}{1 - \int_0^z \exp((u-1)\frac{t^{m-1}}{m-1}dt)}.$$

Proof. We use the second correspondence between permutations and unoriented trees. The key observation is that occurrences of σ correspond to nodes having exactly m leaves as successors. Indeed, each occurrence of σ in the permutation produces that the node corresponding to 2 has nodes $m+2, m+1, \ldots, 4, 3$ hanging from it, and no more nodes appear in the configuration because of the 1 coming after. Reciprocally, every node with exactly m leaves hanging from it corresponds necessarily to an occurrence of σ, except in the possible case when the permutation ends in $2, m+2, m+1, \ldots, 4, 3$ (namely, σ without its last element), case in which these $m+1$ last elements of the permutation are also transformed into a configuration of this type.

We can decompose (1) into

$$\mathcal{T} = \{1\}^{\square} \star [\{1\}^m + (\Pi(\mathcal{T}) - \{1\}^m)],$$

where $\{1\}^m$ represents m leaves. Let $T(u,z)$ be the BGF of unoriented increasing trees, where u marks the number of nodes with m leaves as successors. Since the BGF for the m leaves is z^m/m, T satisfies

$$T'(u,z) = u\frac{z^m}{m!} + e^{T(u,z)} - \frac{z^m}{m!}, \qquad T(u,0) = 0,$$

where the derivative is with respect to z. Setting $S = e^T$, we get

$$S'(u,z) = (u-1)\frac{z^m}{m}S(u,z) + S(u,z)^2, \qquad S(u,0) = 1.$$

Next we need to exclude the possible counting of a configuration corresponding to an ending in $2, m+2, m+1, \ldots, 4, 3$. If we take a closer look at how the correspondence between permutations and trees works, we realize that when we go through the elements of the permutation from left to right building the corresponding tree, the successors of a fixed element always appear in decreasing order, and we first go over all the subtree rooted by one of these successors before passing through another successor with a smaller value. Therefore, we know from the increasing tree which are the last elements of the permutation by just descending along the tree starting from the root and visiting always the smallest successor until we arrive at a leaf, which is necessarily the last element. Rewrite (1) as

$$\mathcal{T} = \{1\} + \{1\}^{\square} \star [\mathcal{T}^{\square} \star \Pi(\mathcal{T})],$$

separating the subtree $\mathcal{T}^{\square}$ with the smallest label, where the last element of the permutation lays. If $U(u,z)$ is the BGF of increasing trees where u marks the number of nodes with exactly m leaves hanging from them, and that do not correspond to the last $m+1$ elements, the above decomposition implies that

$$U(u,z) = z + \int_0^z \int_0^t U'(u,s) e^{T(u,s)} ds dt,$$

which reduces to $U'' = U'e^T$. Finally, let $P(u,z)$ be the BGF of permutations where u marks occurrences of σ. Since the number of nodes of the corresponding tree is one more than the length of the permutation, we have $P = U'$. Therefore P satisfies

$$P'(u,z) = P(u,z)e^{T(u,z)} = P(u,z)S(u,z), \quad P(u,0) = 1.$$

The result stated in the theorem follows easily with the substitution $P(u,z) = 1/\omega(u,z)$. $\square$

The above result admits a significant generalization. The proof uses a direct combinatorial reasoning, but it is be omitted for lack of space.

Theorem 3.2. *Let m be a positive integer, and let $\sigma = 1\tau 2 \in \mathcal{S}_{m+2}$, where τ is any permutation of the elements $\{3,4,\dots,m+2\}$. Then, the distribution of the number of occurrences of σ in permutations does not depend on τ, and the BGF is given by $P(u,z) = 1/\omega(u,z)$, where $\omega(u,z)$ is the solution of*

$$\omega'' + (1-u)\frac{z^m}{m!}\omega' = 0 \tag{3}$$

with $\omega(0) = 1$, $\omega'(0) = -1$.

By reversal and complementation, the last theorem gives the distribution of occurrences of any subword having either their two smallest or their two largest elements in the extremes.

Our next result treats increasing (or decreasing) subwords.

Theorem 3.3. *Let m be a positive integer, let $\sigma = 1,2,\dots,m+1,m+2 \in \mathcal{S}_{m+2}$, and let $P(u,z)$ be the BGF of permutations where u marks the number of occurrences of the subword σ. Then, $P(u,z) = 1/\omega(u,z)$, where ω is the solution of*

$$\omega^{(m+1)} + (1-u)(\omega^{(m)} + \omega^{(m-1)} + \ldots + \omega' + \omega) = 0 \tag{4}$$

with $\omega(0) = 1$, $\omega'(0) = -1$, and $\omega^{(k)} = 0$ for $2 \le k \le m$.

Proof. We use the first correspondence of permutations as binary increasing trees. We get for $P(u,z)$ a system of $m+1$ first order differential equations, which will be reduced to a differential equation of order $m+1$ with the substitution $P(u,z) = 1/\omega(u,z)$. Let $\mathcal{P}$ be the class of all permutations, let $\mathcal{K}_i$ be the class of permutations not beginning with $1,2,\dots,m+2-i$, and let $K_i(u,z)$ be the BGF of $\mathcal{K}_i$ where u marks occurrences of σ. With some abuse of notation, we introduce the parameter u in the equations for classes meaning that it will be placed there when we write the corresponding differential equations for the BGF. With this notation, we can write

$$\mathcal{P} = \{\epsilon\} + \{1\}^{\square} \star \mathcal{P} \star [\mathcal{K}_1 + u(\mathcal{P} - \mathcal{K}_1)].$$

This is because occurrences of σ in a permutation (seen as a binary increasing tree) can be separated into occurrences on the left subtree and occurrences on the right subtree, taking into account that if the permutation on the right subtree begins with $1,2,\dots,m+1$ (that is, belongs to $\mathcal{P} - \mathcal{K}_1$), then there is an additional occurrence of σ beginning at the root.

The next step is to find an equation for $\mathcal{K}_1$. Note that for a permutation not to begin with a $1,2,\dots,m+1$, it is not enough that the permutation on the left subtree does not begin with this subword. We have to exclude also the case in which the left subtree is empty and the right subtree begins with a $1, 2, \ldots, m$. This gives us

$$\mathcal{K}_1 = \{\epsilon\} + \{1\}^{\square} \star (\mathcal{K}_1 - \{\epsilon\}) \star [\mathcal{K}_1 + u(\mathcal{P} - \mathcal{K}_1)] + \{1\}^{\square} \star \mathcal{K}_2.$$

Now it is clear how to find an equation for $\mathcal{K}_2$, in which $\mathcal{K}_3$ will appear, and so on, until we arrive to $\mathcal{K}_m$ (permutations not beginning with $1, 2$), which satisfies

$$\mathcal{K}_m = \{\epsilon\} + \{1\}^{\square} \star (\mathcal{K}_m - \{\epsilon\}) \star [\mathcal{K}_1 + u(\mathcal{P} - \mathcal{K}_1)] + \{1\}.$$

All these equations yield a system of differential equations involving the corresponding BGFs. After the substitution

$$R = uP + (1-u)K_1$$

we obtain

$$(5) \qquad \begin{cases} R' = R^2 + (u-1)(R - K_2) \\ {K_2}' = (K_2 - 1)R + K_3 \\ {K_3}' = (K_3 - 1)R + K_4 \\ \vdots \\ {K_{m-1}}' = (K_{m-1} - 1)R + K_m \\ {K_m}' = (K_m - 1)R + 1 \\ P' = PR \end{cases} \qquad \text{with } P(0) = R(0) = 1, K_i(0) = 1 \quad \text{for all } i$$

Finally, it is easy to check that setting $P(u,z) = 1/\omega(u,z)$, then ω satisfies (4). $\square$

Note that for $u = 0$ the differential equation for ω has a very simple solution, a linear combination of exponentials $e^{\lambda_k z}$ where λ_k are $(m-2)$-th roots of unity, so we can easily get an explicit expression for $P(0,z)$, the GF of permutations avoiding the subword σ.

Now comes the last result in this section, which generalizes the theorems 3.1 and 3.2.

Theorem 3.4. *Let m, a be positive integers with $a \leq m$, let $\sigma = 1, 2, \ldots, a-1, a, \tau, a+1 \in \mathcal{S}_{m+2}$, where τ is any permutation of the elements $\{a+2, a+3, \ldots, m+2\}$, and let $P(u,z)$ be the BGF of permutations where u marks the number of occurrences of the subword σ. Then, $P(u,z) = 1/\omega(u,z)$, where ω is the solution of*

$$(6) \qquad \omega^{(a+1)} + (1-u)\frac{z^{m-a+1}}{(m-a+1)!}\omega' = 0$$

with $\omega(0) = 1$, $\omega'(0) = -1$ and $\omega^{(k)} = 0$ for $2 \leq k \leq a$. In particular, the distribution does not depend on τ.

Proof. Again we find a system of $a+1$ differential equations for $P(u,z)$ that, after the substitution $P(u,z) = 1/\omega(u,z)$, yield a differential equation of order $a+1$. Let $\mathcal{P}$ be as before the class of all permutations. For $1 \leq i \leq m+1$, we denote $\sigma_{>i} = \rho(\sigma_{i+1}, \sigma_{i+2}, \ldots, \sigma_{m+2})$. Note that for $i < a$, $\sigma_{>i}$ has its smallest element in the first position, while for $i \geq a$, the smallest element of $\sigma_{>i}$ is the last one. Now let $\mathcal{K}_i$ be the class of permutations not beginning with any of the following: $\sigma_{>1}, \sigma_{>2}, \ldots, \sigma_{>i}$, and let $K_i(u,z)$ be the BGF of $\mathcal{K}_i$ where u marks occurrences of σ. With this notation, we can write

$$\mathcal{P} = \{\epsilon\} + \{1\}^{\square} \star \mathcal{P} \star [\mathcal{K}_1 + u(\mathcal{P} - \mathcal{K}_1)].$$

The explanation, as in previous proofs, is that in the decomposition of the increasing tree, new occurrences of σ, apart from the ones on the left and right subtrees, can appear beginning in the root, when the permutation on the right subtree begins with $\sigma_{>1}$ (that is, belongs to $\mathcal{P} - \mathcal{K}_1$).

For $\mathcal{K}_1$ we derive the equation

$$\mathcal{K}_1 = \{\epsilon\} + \{1\}^{\square} \star (\mathcal{K}_1 - \{\epsilon\}) \star [\mathcal{K}_1 + u(\mathcal{P} - \mathcal{K}_1)] + \{1\}^{\square} \star [\mathcal{K}_2 + u(\mathcal{P} - \mathcal{K}_1)].$$

Note that the last summand corresponds to the case where the left subtree is empty. Then, the permutation on the right subtree cannot begin with $\sigma_{>2}$, which would produce, together with the root, a permutation beginning with $\sigma_{>1}$. So, after separating on the right subtree the permutations beginning with $\sigma_{>1}$ $(\mathcal{P}-\mathcal{K}_1)$, we are left with those beginning with neither $\sigma_{>1}$ nor $\sigma_{>2}$, that is, $\mathcal{K}_2$.

Analogous equations can be found for $\mathcal{K}_i$, $1<i<a$:

$$\mathcal{K}_i = \{\epsilon\} + \{1\}^\Box \star (\mathcal{K}_i - \{\epsilon\}) \star [\mathcal{K}_1 + u(\mathcal{P}-\mathcal{K}_1)] + \{1\}^\Box \star [\mathcal{K}_{i+1} + u(\mathcal{P}-\mathcal{K}_1)]$$

For $\mathcal{K}_a$ (note that $\sigma_{>a} = (\rho(\tau),1))$ we have

$$\mathcal{K}_a = \{\epsilon\} + \{1\}^\Box \star (\mathcal{K}_a - \{\epsilon\} - \{\rho(\tau)\}) \star [\mathcal{K}_1 + u(\mathcal{P}-\mathcal{K}_1)] + \{1\}^\Box \star [\mathcal{K}_a + u(\mathcal{P}-\mathcal{K}_1)].$$

The difference now is that in order to avoid a permutation beginning with $\sigma_{>a} = \rho(\tau, a+1)$, the left subtree cannot be $\rho(\tau)$. (Remember that the BGF corresponding to $\{\rho(\tau)\}$ is $z^{m-a+1}/(m-a+1)!$ since it is a permutation of size $m-a+1$ and does not contain the subword σ.) Another difference is that no new variables appear, since when the left subtree is empty there is no danger of beginning with $\sigma_{>a}$, and so there are no additional restrictions for the right subtree.

¿From all these equations, we get a system of differential equations. It is convenient to apply the substitutions $R = uP + (1-u)K_1$ as before, and also $S_i = K_{i-1} - K_i$ for $2 \le i \le a$. Then

$$\left\{\begin{array}{l} R' = R^2 + (u-1)S_2 \\ {S_2}' = S_2 R + S_3 \\ {S_3}' = S_3 R + S_4 \\ \vdots \\ {S_{a-1}}' = S_{m-1} R + S_a \\ {S_a}' = (S_a + \frac{z^{m-a+1}}{(m-a+1)!})R \\ P' = PR \end{array}\right. \qquad \text{with } P(0) = R(0) = 1, S_i(0) = 0 \quad \text{for all } i \tag{7}$$

Finally, it only remains to set $P(u,z) = 1/\omega(u,z)$ and to check that ω is the solution of (6). $\Box$

Note that the fact that $a \le m$ ensures that τ is not empty. The case in which τ is empty corresponds to the case where σ is the increasing permutation, and has already been treated in the previous theorem.

4. Subwords of length at most four

Occurrences of the two subwords 12 and 21 of length two correspond, respectively, to ascents and descents in permutations, giving rise to the well-known Eulerian numbers [2].

Subwords of length three. Among the 6 permutations of three elements, there are only two different classes regarding its distribution as subwords of permutations since, by reversal and complementation, we have that

$$\begin{aligned} &123 \sim 321 \\ &132 \sim 231 \sim 312 \sim 213 \end{aligned}$$

By the results in the previous section we get:

Theorem 4.1. *Let $P(u,z)$ and $Q(u,z)$ be the BGFs of permutations where u marks, respectively, the number of occurrences of the subword 123 and 132. Then*

$$\begin{aligned} Q(u,z) &= \frac{1}{1-\int_0^z e^{(u-1)t^2/2}dt} \\ P(u,z) &= \frac{2e^{\frac{1}{2}(1-u+\sqrt{(u-1)(u+3)})z}\sqrt{(u-1)(u+3)}}{1+u+\sqrt{(u-1)(u+3)}+e^{\sqrt{(u-1)(u+3)}z}(-1-u+\sqrt{(u-1)(u+3)})} \\ P(0,z) &= \frac{\sqrt{3}}{2}\frac{e^{z/2}}{\cos(\frac{\sqrt{3}}{2}z+\frac{\pi}{6})} \end{aligned}$$

Furthermore, the numbers $\alpha_n(123)$ and $\alpha_n(132)$ of permutations avoiding, respectively, the subwords 123 and 132, satisfy

$$\alpha_n(123) \sim \gamma_1 \cdot (\rho_1)^n \cdot n!$$
$$\alpha_n(132) \sim \gamma_2 \cdot (\rho_2)^n \cdot n!$$

where $\rho_1 = \frac{3\sqrt{3}}{2\pi}$, $\gamma_1 = e^{3\sqrt{3}\pi}$, $(\rho_2)^{-1}$ is the unique positive root of $\int_0^z e^{-t^2/2}dt = 1$, and $\gamma_2 = \exp((\rho_2)^{-2}/2)$, the approximate values being

$$\rho_1 = 0.8269933, \qquad \gamma_1 = 1.8305194$$
$$\rho_2 = 0.7839769, \qquad \gamma_2 = 2.2558142$$

Table 2 indicates the number of permutations of length n with k occurrences of the subwords 123 (top) and 132 (bottom).

TABLE 2. Occurrences of subwords of length 3 in permutations.

$n\backslash k$	0	1	2	3	4
1	1 1				
2	2 2				
3	5 5	1 1			
4	17 16	6 8	1		
5	70 63	41 54	8 3	1	
6	349 296	274 368	86 56	10	1

The asymptotic estimates are obtained as an application of the result quoted at the end of Section 2. (The computation of ρ_2 has been done numerically using the computer algebra system MAPLE.) Since $\rho_1 > \rho_2$ in the previous theorem, we see that $\alpha_n(123)$ is asymptotically larger than $\alpha_n(132)$. Looking at Table 2 one is led to conjecture that this is always so.

Proposition 4.2. *For every $n \geq 4$, we have*

$$\alpha_n(123) > \alpha_n(132).$$

There is a simple combinatorial proof of the last result, that we omit for lack of space.

Subwords of length four. By reversal and complementation, and by Theorem 3.2, they fall into seven classes, and the results in the previous section give the BGFs for the distribution of occurrences of subwords I, VI and VII.

I. $1234 \sim 4321$
II. $2413 \sim 3142$
III. $2143 \sim 3412$
IV. $1324 \sim 4231$
V. $1423 \sim 3241 \sim 4132 \sim 2314$
VI. $1342 \sim 2431 \sim 4213 \sim 3124 \sim 1432 \sim 2341 \sim 4123 \sim 3214$
VII. $1243 \sim 3421 \sim 4312 \sim 2134$

Theorem 4.3. *In each of the following cases, let $P(u,z)$ be the BGF of permutations where u marks the number of occurrences of the corresponding subword.*

Case 1342. $P(u,z) = \frac{1}{1-\int_0^z e^{(u-1)t^3/6}dt}$.

Case 1234. $P = 1/\omega$, *where ω is the solution of*

$$\omega''' + (1-u)(\omega'' + \omega' + \omega) = 0$$

with $\omega(0) = 1$, $\omega'(0) = -1$, $\omega''(0) = 0$. For $u = 0$, the solution is

$$P(0,z) = \frac{2}{\cos z - \sin z + e^{-z}}$$

Case 1243. $P = 1/\omega$, *where ω is the solution of*

$$\omega''' + (1-u)z\omega' = 0$$

with $\omega(0) = 1$, $\omega'(0) = -1$, $\omega''(0) = 0$.

Furthermore, the numbers $\alpha_n(1342), \alpha_n(1234)$ and $\alpha_n(1243)$ satisfy

$$\begin{aligned}\alpha_n(1342) &\sim \gamma_1 \cdot (\rho_1)^n \cdot n! \\ \alpha_n(1234) &\sim \gamma_2 \cdot (\rho_2)^n \cdot n! \\ \alpha_n(1243) &\sim \gamma_3 \cdot (\rho_3)^n \cdot n!\end{aligned}$$

where $(\rho_1)^{-1}$ is the smallest postitive solution of $\int_0^z e^{(u-1)t^3/6}dt = 1$, $(\rho_2)^{-1}$ is the smallest positive solution of $\cos z - \sin z + e^{-z} = 0$, and ρ_3 is the solution of a certain equation involving Airy functions. The approximate values are

$$\begin{aligned}\rho_1 &= 0.954611, & \gamma_1 &= 1.8305194 \\ \rho_2 &= 0.963005, & \gamma_2 &= 2.2558142 \\ \rho_3 &= 0.952891, & \gamma_3 &= 1.6043282\end{aligned}$$

In the last case the equation $\omega''' + (1-u)z\omega' = 0$, for $u = 0$ and $v = w'$, can be solved in terms of Bessel functions (note that the equation for v is actually a slight variant of the Airy equation). The computations have been performed with the use of the MAPLE system. Table 3 shows the number of occurrences in each of the seven classes for $n \leq 7$, showing that no two of them have the same distribution. (Entries II, III, IV, V in Table 3 have been computed directly.)

As in the case of length 3, one might expect that again for any two permutations $\sigma, \tau \in \mathcal{S}_4$, if $\alpha_{n_0}(\sigma) > \alpha_{n_0}(\tau)$, then $\alpha_n(\sigma) > \alpha_n(\tau)$ for all $n > n_0$. However, inequality $\alpha_n(1324) \geq \alpha_n(2143)$ holds for $n \leq 11$ but does not hold for $n = 12$. Indeed,

$$\alpha_{11}(1324) = 27959880 > 27954521 = \alpha_{11}(2143),$$

but

$$\alpha_{12}(1324) = 320706444 < 320752991 = \alpha_{12}(2143).$$

This results have been obtained by exhaustive computation, as we do not know the associated EGFs. There is however one relation among classes that we have been able to establish: for every $n \geq 7$, we have $\alpha_n(1342) > \alpha_n(1243)$.

Acknowledgements. We are grateful to Juan-José Morales for showing us how to linearize the various differential equations, and to the referees for useful comments.

TABLE 3. Occurrences of subwords of length 4 in permutations.

$n\backslash k$		0	1	2	3	4
4	I	23	1			
	II	23	1			
	III	23	1			
	IV	23	1			
	V	23	1			
	VI	23	1			
	VII	23	1			
5	I	111	8	1		
	II	110	10			
	III	110	10			
	IV	110	10			
	V	110	10			
	VI	110	10			
	VII	110	10			
6	I	642	67	10	1	
	II	632	86	2		
	III	631	88	1		
	IV	632	86	2		
	V	631	88	1		
	VI	630	90			
	VII	630	90			
7	I	4326	602	99	12	1
	II	4237	766	37		
	III	4223	794	23		
	IV	4229	782	29		
	V	4218	804	18		
	VI	4210	820	10		
	VII	4204	832	4		

REFERENCES

[1] L. Carlitz, R. Scoville, Enumeration of permutations by rises, falls, rising maxima and falling maxima, *Acta Math. Scient. Hungar.* 25 (1974), 269–277.

[2] L. Comtet, *Advanced Combinatorics*, Reidel (1974).

[3] S. Elizalde, M. Noy, Forbidden subwords in permutations, Tech. Report MAII-IR-21, Universitat Politècnica de Catalunya, 2000 (20 pages).

[4] P. Flajolet, X. Gourdon, C. Martínez, Patterns in random binary search trees, *Random Structures Algorithms* 11 (1997), 223–244.

[5] P. Flajolet, M. Noy, Analytic combinatorics of non-crossing configurations, *Discrete Math.* 204 (1999), 203–229.

[6] P. Flajolet, P. and R. Sedgewick, *Analytic combinatorics* (book in preparation) 1998. (Individual chapters are available as INRIA Research Reports 1888, 2026, 2376, 2956, 3162.).

[7] I.P. Goulden and D.M. Jackson, *Combinatorial Enumeration*, John Wiley (1983).

[8] R. Sedgewick and P. Flajolet, *An introduction to the analysis of algorithms*, Addison-Wesley (1996).

[9] R. Stanley, *Enumerative Combinatorics*, vol. I, Cambridge Univ. Press, Cambridge (1997).

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 77 MASSACHUSETTS AVENUE, CAMBRIDGE MA 02139.

E-mail address: `sergi@math.mit.edu`

DEPARTAMENT DE MATEMÀTICA APLICADA II, UNIVERSITAT POLITÈCNICA DE CATALUNYA, PAU GARGALLO 5, 08028 BARCELONA, SPAIN.

E-mail address: `noy@ma2.upc.es`