

# Pattern statistics in faro words and permutations

Jean-Luc Baril<sup>a,\*</sup>, Alexander Burstein<sup>b</sup>, Sergey Kirgizov<sup>a</sup>

<sup>a</sup> LIB, Université de Bourgogne Franche-Comté, B.P. 47 870, 21078 Dijon Cedex, France

<sup>b</sup> Howard University, Department of Mathematics, Washington, DC 20059, USA



## ARTICLE INFO

### Article history:

Received 13 October 2020

Received in revised form 22 April 2021

Accepted 3 May 2021

Available online xxxx

### Keywords:

Word

Permutation

Faro shuffle

Pattern statistics

Popularity

Equidistribution

## ABSTRACT

We study the distribution and the popularity of some patterns in  $k$ -ary faro words, i.e. words over the alphabet  $\{1, 2, \dots, k\}$  obtained by interlacing the letters of two nondecreasing words of lengths differing by at most one. We present a bijection between these words and dispersed Dyck paths (i.e. Motzkin paths with all level steps on the  $x$ -axis) with a given number of peaks. We show how the bijection maps statistics of consecutive patterns of faro words into linear combinations of other pattern statistics on paths. Then, we deduce enumerative results by providing multivariate generating functions for the distribution and the popularity of patterns of length at most three. Finally, we consider some interesting subclasses of faro words that are permutations, involutions, derangements, or subexcedent words.

© 2021 Elsevier B.V. All rights reserved.

## 1. Introduction and notations

The faro shuffle is a well-known technique to shuffle a deck of cards. The deck is split in two at the middle, and the cards from the two halves are combined back by taking alternatively the bottoms of stacks. Certain mathematical questions about the faro shuffle are considered for example in the works of Morris [23], Diaconis, Graham and Kantor [15]. Inspired by these studies and a solid body of modern combinatorial literature (see for instance Lothaire [21], Stanley [27], Bóna [13] and Kitaev [19] books) that explores enumerative and bijective aspects of patterns in various discrete structures, the present paper considers an unexpectedly overlooked combinatorial objects, which we call *faro words*. They are special kind of word shuffles, which are important in several algorithmic and combinatorial settings (see for example Barnes work [7] and references therein). In this paper, we present enumerative results and show how faro words and patterns therein are related to other structures such as Dyck paths, Motzkin paths and Dumont permutations.

### 1.1. Faro words and permutations

We deal with  $k$ -ary words  $u_1 u_2 \dots u_n$  over the integer alphabet  $[1, k] = \{1, 2, \dots, k\}$  endowed with the usual total order. A  $k$ -ary word is called *nondecreasing* if  $u_i \leq u_{i+1}$  for all  $i \in [1, n - 1]$ .

**Definition 1.1.** For two  $k$ -ary words  $u$  and  $v$  such that  $0 \leq |u| - |v| \leq 1$ , the *faro shuffle* of  $u$  and  $v$  is the  $k$ -ary word of length  $|u| + |v|$  obtained by interlacing the letters of  $u$  and  $v$  as follows:  $u_1 v_1 u_2 v_2 u_3 v_3 \dots$ . A  $k$ -ary faro word is a faro shuffle of two nondecreasing  $k$ -ary words.

\* Corresponding author.

E-mail addresses: [barjl@u-bourgogne.fr](mailto:barjl@u-bourgogne.fr) (J.-L. Baril), [aburstein@howard.edu](mailto:aburstein@howard.edu) (A. Burstein), [sergey.kirgizov@u-bourgogne.fr](mailto:sergey.kirgizov@u-bourgogne.fr) (S. Kirgizov).

Let  $\mathcal{S}_{n,k}$  be the set of  $k$ -ary faro words of length  $n$ . Its cardinality equals the product of two binomial coefficients  $\binom{\lfloor n/2 \rfloor + k - 1}{k-1} \binom{\lceil n/2 \rceil + k - 1}{k-1}$ , each of them being, respectively, the number of  $m$ -multisets of  $[1, k]$  for  $m = \lfloor \frac{n}{2} \rfloor$  and  $m = \lceil \frac{n}{2} \rceil$ . For example, we have  $\mathcal{S}_{4,2} = \{1111, 1112, 1121, 1122, 1212, 1222, 2121, 2122, 2222\}$  and  $|\mathcal{S}_{4,2}| = 9$ .

**Definition 1.2.** A faro permutation of length  $n$  is an  $n$ -ary faro word of length  $n$  that contains every letter in  $[1, n]$  exactly once.

Let  $\mathcal{P}_n$  be the set of length  $n$  faro permutations. For instance, we have  $\mathcal{P}_3 = \{123, 132, 213\}$ . Since a faro permutation is entirely determined by the choice of its values on the odd indices, the cardinality of  $\mathcal{P}_n$  is  $\binom{n}{\lfloor n/2 \rfloor}$ .

A  $k$ -ary word  $w = w_1 w_2 \dots w_n$  avoids a classical pattern (resp. consecutive pattern)  $p = p_1 p_2 \dots p_k$  (resp.  $p = p_1 p_2 \dots p_k$ ) if there does not exist a strictly increasing sequence of indices  $i_1 i_2 \dots i_k$  (resp. with  $i_{j+1} = i_j + 1$  for  $1 \leq j \leq k-1$ ) such that  $w_{i_1} w_{i_2} \dots w_{i_k}$  is order-isomorphic to  $p$  (see [19] for instance). Obviously, any faro word avoids the classical pattern 3-2-1. Let  $Av_n(\sigma)$  denote the set of permutations avoiding a classical pattern  $\sigma$ , then we have  $\mathcal{P}_n \subseteq Av_n(3-2-1)$  for  $n \geq 0$ , and  $\mathcal{P}_n \neq Av_n(3-2-1)$  for  $n \geq 3$  since  $(n-1)n12 \dots (n-2) \in Av_n(3-2-1)$  is not a faro word. Note that a faro permutation can contain all classical patterns of length 3 except 3-2-1 (e.g., 31425).

**Remark 1.3.** A  $k$ -ary word  $w = w_1 w_2 \dots w_n$  is a faro word if and only if  $w_i \leq w_{i+2}$  for any  $i \in [1, n-2]$ , which means that faro permutations are precisely those avoiding the three consecutive patterns 231, 321 and 312.

### 1.2. Dyck and dispersed Dyck paths

In order to study the distribution of patterns in faro words, we will exhibit one-to-one correspondences between these objects and some specific lattice paths in the first quadrant of the plane. Hence, we provide basic necessary definitions on lattice paths.

**Definition 1.4.** Dispersed Dyck paths (see [17]) are lattice paths starting at  $(0, 0)$ , ending at  $(n, 0)$ , consisting of level steps  $F = (1, 0)$ , up step  $U = (1, 1)$  and down steps  $D = (1, -1)$ , and never going below the  $x$ -axis and where all level steps are on the  $x$ -axis.

Let  $\mathcal{B}_n$  be the set of dispersed Dyck paths of length  $n$  (or, equivalently, consisting of  $n$  steps) and set  $\mathcal{B} = \cup_{n \geq 0} \mathcal{B}_n$ , where the empty path is denoted by  $\epsilon$ . A Dyck path of semilength  $n \geq 0$  is a dispersed Dyck path of length  $2n$  with no level steps. Let  $\mathcal{D}_n$  be the set of Dyck paths of semilength  $n$  and let  $\mathcal{D} = \cup_{n \geq 0} \mathcal{D}_n$ . Dispersed Dyck paths of length  $n$  are in straightforward bijection with prefixes of Dyck paths of length  $n$ , also known as ballot paths [8,28]. Indeed, we can obtain a ballot path from a dispersed Dyck path by replacing all level steps with up steps. Dyck and dispersed Dyck paths are counted by the Catalan and ballot numbers, respectively (see A000108 and A001405 in the Online Encyclopedia of Integer Sequences of N.J.A. Sloane [26], where the general terms are  $c_n = \frac{1}{n+1} \binom{2n}{n}$  and  $b_n = \binom{n}{\lfloor n/2 \rfloor}$ , respectively).

A path  $P$  avoids a pattern  $X$  if and only if  $P$  does not contain  $X$  as a sequence of consecutive steps (see for instance [14,22]). Note that other pattern definitions exist in the literature where steps are not necessarily consecutive [3]. We also need some notations similar to Kleene star and plus symbols of formal language theory. For a nonempty pattern  $X$ , an occurrence of the pattern  $X^+$  in a path  $P$  is a maximal sequence of consecutive repetitions of  $X$ , i.e. a maximal subword of the form  $X^k$  for  $k \geq 1$ . The pattern  $X^*$  will be either an empty pattern or a pattern  $X^+$ . More generally, for two possibly empty patterns  $Y$  and  $Z$  such that  $Y$  does not end with  $X$  and  $Z$  does not start with  $X$ , the pattern  $YX^+Z$  (resp.  $YX^*Z$ ) corresponds to an occurrence obtained by concatenation of  $Y$ ,  $X^+$  and  $Z$  (resp.  $Y$ ,  $X^*$  and  $Z$ ). For instance, the path  $FUDUDFFUDF$  contains two occurrences of the pattern  $F(UD)^+F$  and three occurrences of  $F(UD)^*F$ .

### 1.3. Statistics on words and lattice paths

**Definition 1.5.** A statistic  $\mathbf{s}$  is an integer-valued function from a set  $\mathcal{A}$  of words or paths.

To a given pattern  $p$ , we associate the pattern statistic  $\mathbf{p} : \mathcal{A} \rightarrow \mathbb{N}$  such that  $\mathbf{p}(a)$  is the number of occurrences of the pattern  $p$  in the object  $a \in \mathcal{A}$  (we use the boldface to denote statistics). For example, the statistic giving the number of occurrences of the consecutive pattern 123 (resp.  $UDUD$ ) in a word (resp. a lattice path) is denoted by **123** (resp. **UDUD**). We denote by  $\hat{\mathbf{1}}$  (resp.  $\hat{\mathbf{2}}, \hat{\mathbf{n}}$ ) the constant statistic returning the value 1 (resp. 2,  $n$ ).

**Definition 1.6.** The popularity of a pattern  $p$  in  $\mathcal{A}$  is the total number of occurrences of  $p$  over all objects of  $\mathcal{A}$ , that is  $\mathbf{p}(\mathcal{A}) = \sum_{a \in \mathcal{A}} \mathbf{p}(a)$  (see [5,10,18,19]).

For instance, for a dispersed Dyck path  $P = FFUDFUUDUUUDDDD$  we have  $\mathbf{FF}(P) = 1$ ,  $\mathbf{DDD}(P) = 2$ ,  $\mathbf{UD}(P) = 3$ ,  $\mathbf{UUUU}(P) = 0$  and  $\hat{\mathbf{1}}(P) = 1$ . Moreover, if  $\mathcal{A} = \{UUDD, UDUD\}$  then the popularity of the pattern  $UD$  in  $\mathcal{A}$  is  $\mathbf{UD}(\mathcal{A}) = 3$ .

Let  $\mathbf{T}_{\mathcal{A}}$  be the set of all statistics defined on a set  $\mathcal{A}$ . For any pair of statistics  $\mathbf{s}, \mathbf{t} \in \mathbf{T}_{\mathcal{A}}$ , we define the statistic  $\mathbf{s} + \mathbf{t}$  by  $(\mathbf{s} + \mathbf{t})(a) = \mathbf{s}(a) + \mathbf{t}(a)$  for any  $a \in \mathcal{A}$ , which endows  $\mathbf{T}_{\mathcal{A}}$  with a  $\mathbb{Z}$ -module structure. Let  $\mathcal{B}$  be a set of combinatorial objects, and let  $\mathbf{T}_{\mathcal{B}}$  be the corresponding set of statistics. We say that two statistics  $\mathbf{s} \in \mathbf{T}_{\mathcal{A}}$  and  $\mathbf{t} \in \mathbf{T}_{\mathcal{B}}$  have the *same distribution*, or are *equidistributed*, if there exists a bijection  $f : \mathcal{A} \rightarrow \mathcal{B}$  such that  $\mathbf{s}(a) = \mathbf{t}(f(a))$  for any  $a \in \mathcal{A}$ . In this case, with a slight abuse of the notation already used in [4], we write shortly  $f(\mathbf{s}) = \mathbf{t}$  or  $\mathbf{s} = \mathbf{t}$  whenever  $f$  is the identity. As a byproduct, for any constant statistic  $\hat{\mathbf{n}}$ , we have  $f(\hat{\mathbf{n}}) = \hat{\mathbf{n}}$ .

### 1.4. Outline of the paper

The paper is organized as follows. In Section 2, we present a constructive bijection  $f$  between the set  $\mathcal{S}_{n,k}$  of  $k$ -ary faro words of length  $n$  and the set of dispersed Dyck paths of length  $n + 2k - 2$  with  $k - 1$  peaks. We show where pattern statistics are transported by  $f$ , which provides a more suitable ground for studying the distribution of consecutive patterns. Thus, we derive enumerating results on the distribution and popularity of patterns in  $\mathcal{S}_{n,k}$  by giving multivariate generating functions where the coefficient of  $x^n y^k z^t$  is the number of  $k$ -ary faro words of length  $n$  having exactly  $t$  occurrences of a given pattern. In Section 3, we present a similar study for faro permutations. More precisely, we provide a bijection  $g$  between  $\mathcal{P}_n$  and the set of dispersed Dyck paths of length  $n$  and show how  $g$  acts on pattern statistics of length at most three. Consequently, we deduce enumerative results for the distribution and the popularity of these patterns in  $\mathcal{P}_n$ . We also present a bijection between  $\mathcal{P}_n$  and involutions avoiding the classical pattern 3-2-1. Finally, in Section 4, we prove that the set of subexcedent words in  $\mathcal{S}_{n,n}$  is related to ternary trees and Dumont permutations of the second kind [12] avoiding the classical pattern 2-1-4-3, and we show why faro involutions and faro derangements are respectively enumerated by the Fibonacci and Catalan numbers.

## 2. Patterns in faro words

In this section we construct a bijection  $f$  between the set  $\mathcal{S}_{n,k}$  of  $k$ -ary faro words of length  $n$  and a subset of dispersed Dyck paths, and show how  $f$  transports pattern statistics. Then, we deduce generating functions for the distribution and popularity of some patterns.

A pair in a faro word  $w$  is an occurrence  $w_i w_{i+1}$  with  $w_i > w_{i+1}$ . Remark 1.3 implies that a letter cannot be part of two pairs since a faro word avoids the consecutive pattern 321. A singleton in  $w$  is a letter  $w_i$  not in any pair of  $w$ . Any faro word can be uniquely decomposed as a sequence of pairs and singletons, which are called blocks of faro words. For instance, the block decomposition of 111212131333 is  $1^3(21)^2(31)3^3$ .

Let  $\mathcal{L}_k$  be the set of all possible blocks of a decomposition of a  $k$ -ary faro word, that is

$$\mathcal{L}_k = \{1, 2, \dots, k\} \cup \{ji : 1 \leq i < j \leq k\}.$$

**Definition 2.1.** We define an order relation  $\leq$  on  $\mathcal{L}_k$  as follows: for  $g, h, i, j \in \{1, 2, \dots, k\}$ ,

$$\begin{cases} i \leq j, & \text{if } i \leq j, \\ i \leq jh, & \text{if } i \leq h < j, \\ ig \leq j, & \text{if } g < i \leq j, \\ ig \leq jh, & \text{if } g < i \leq j \text{ and } g \leq h < j. \end{cases}$$

**Remark 2.2.** The order relation  $\leq$  can be defined less technically as follows: for  $p, q \in \mathcal{L}_k$ ,

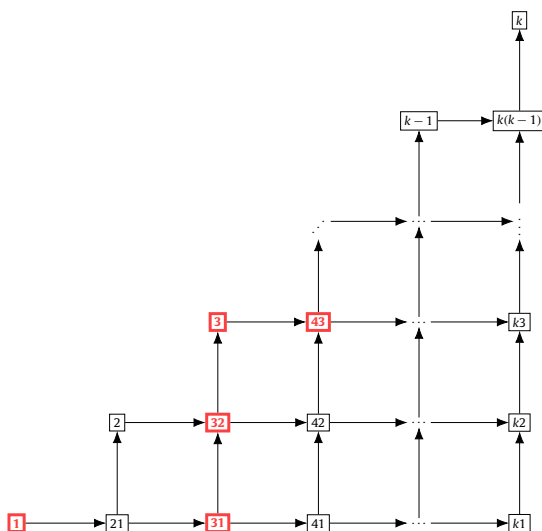
$$p \leq q \iff pq \text{ is a faro word different from a pair.}$$

This order relation endows the set  $\mathcal{L}_k$  with a poset structure, which we call *faro poset*. See Fig. 2.1 for an illustration of the Hasse diagram of  $(\mathcal{L}_k, \leq)$ .

A *multichain* in a poset is a chain, i.e. a *totally ordered subset*, with repetitions allowed. Due to the simple structure of the faro poset, we easily deduce the following remarks.

**Remark 2.3.** There is a one-to-one correspondence between  $k$ -ary faro words and the multichains of  $\mathcal{L}_k$ . Indeed, Remark 2.2 implies that the block decomposition of a  $k$ -ary faro word  $w$  into pairs and singletons  $w = b_1 b_2 \dots b_\ell$  unambiguously corresponds to the multichain  $b_1 \leq b_2 \leq \dots \leq b_\ell$  in  $\mathcal{L}_k$ , and vice versa. For instance, the faro word 11313232343 = 11(31)(32)(32)3(43) corresponds to the multichain  $1 \leq 1 \leq 31 \leq 32 \leq 32 \leq 3 \leq 43$  (see Fig. 2.1).

**Remark 2.4.** If a  $k$ -ary faro word  $w$  contains a singleton  $x$  in its decomposition into blocks, then it satisfies the following property: the set of pairs of the form  $ab$ ,  $b < a \leq x$ , equals the set of pairs of the form  $cd$ ,  $d \leq x - 1$ .



**Fig. 2.1.** The faro poset  $(\mathcal{L}_k, \leq)$ . Red blocks represent the multichain associated to the  $k$ -ary faro word  $11313232343 = 1^2(31)(32)^23(43)$ . (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

2.1. A bijection to the set of dispersed Dyck paths

As mentioned by E. Deutsch in [26] (see sequence A124428), the number of dispersed paths of length  $n$  with  $k$  peaks (a peak is an occurrence of the pattern  $UD$ ) is given by

$$|\mathcal{B}_{n,k}| = \binom{\lfloor \frac{n}{2} \rfloor}{k} \binom{\lceil \frac{n}{2} \rceil}{k}.$$

Thus, we present a bijection  $f$  from the set  $\mathcal{S}_{n,k}$  of  $k$ -ary faro words of length  $n$  to the set  $\mathcal{B}_{n+2(k-1),k-1}$  of dispersed Dyck paths of length  $n + 2(k - 1)$  with exactly  $k - 1$  peaks. For a given  $w \in \mathcal{S}_{n,k}$ , we set

$$f(w) = F^{T_0} U^{T_1} D^{T_2} F^{T_3} \dots F^{T_{3(k-2)}} U^{T_{3(k-2)+1}} D^{T_{3(k-2)+2}} F^{T_{3(k-1)}},$$

where  $T_i$  is defined for  $0 \leq i \leq 3(k - 1)$  as follows:

- if  $i = 3(x - 1)$  then  $T_i$  is the number of occurrences of the singleton  $x$  in  $w$ ;
- if  $i = 3(x - 1) - 1$  then  $T_i$  is one plus the number of pairs  $xy, y < x$ , in  $w$ ;
- if  $i = 3(x - 1) + 1$  then  $T_i$  is one plus the number of pairs  $yx, y > x$ , in  $w$ .

It is worth noting that the image of a faro word  $w \in \mathcal{S}_{n,k}$  depends on the arity  $k$  that we consider. Indeed, the image of the empty word  $\epsilon$  is  $UD$  when  $k = 2$ , while  $f(\epsilon) = UDUD$  for  $k = 3$ . We refer to Fig. 2.2 for one detailed example of this bijection, while Fig. 2.3 provides more additional examples. For instance, the images by  $f$  of the 5-ary words  $\epsilon, 12345, 3141, 11111212222$  are, respectively,  $UDUDUDUD, FUDFUDFUDFUDF, UUUDUDDUDDUD$  and  $FFFFFFUDDFFFUDUDUD$ .

**Remark 2.5.** Clearly, the values  $T_i, 0 \leq i \leq 3(k - 1)$ , can be obtained from  $w$  by reading it from left to right and by determining if the current entry  $x$  belongs to either a pair  $xy$  or  $yx$ , or a singleton  $x$ . Moreover, values of  $T$  at indices  $i = 0 \pmod 3$  correspond to the lengths of maximal runs of consecutive level steps, and values at indices  $i = 1 \pmod 3$  (resp.  $i = 2 \pmod 3$ ) correspond to the lengths of maximal runs of consecutive up (resp. down) steps, which means that the sequence  $T = T_0 T_1 \dots T_{3(k-1)}$  is a run-length-like encoding of the path  $f(w)$ . Thus,  $f(w)$  can be constructed from  $w$  using a linear time algorithm.

**Lemma 2.6.** The path  $f(w)$  is necessarily a dispersed Dyck path of length  $n + 2(k - 1)$  with exactly  $k - 1$  peaks.

**Proof.** Since for any  $i \not\equiv 0 \pmod 3, 1 \leq i \leq 3(k - 1) - 1$  we have  $T_i \geq 1$ , the path  $w$  contains exactly  $k - 1$  peaks  $UD$ . Interpreting Remark 2.4 on the path  $f(w)$ , the number of up steps before a given level step equals the number of down steps before the same level step, which implies that any level step belongs to the  $x$ -axis. Let  $d_x = \sum_{i=2}^{x+2} T_{3(i-1)-1}$  (resp.  $u_x = \sum_{i=1}^{x+1} T_{3(x-1)+1}$ ) be the total number of down steps (resp. up steps) in the first  $x + 1$  maximal runs of down steps (resp. up steps). Due to the definition of  $f$ ,  $d_x$  equals the number of pairs  $ij, 1 \leq j < i \leq x + 2$ , in  $w$ , and  $u_x$  equals the

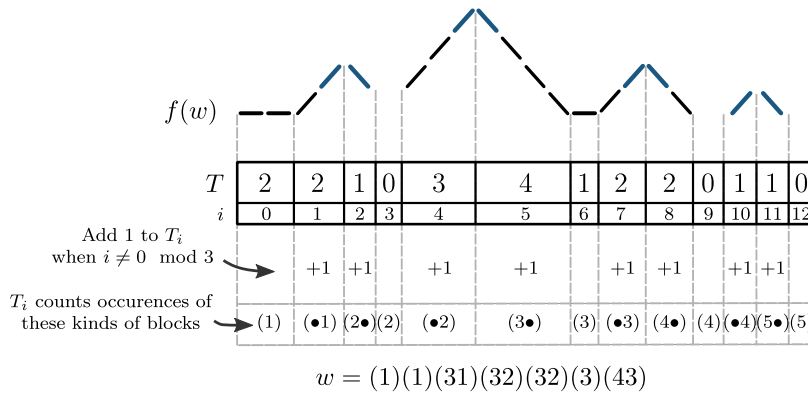


Fig. 2.2. The image by  $f$  of the 5-ary faro word  $w = 11313232343$  is  $f(w) = FFUUDUUUDDDDFFUDDUD$ .

number of pairs  $ij, 1 \leq j \leq x + 1, i \geq j + 1$ , which implies that  $d_x \leq u_x$ . Also by definition, the total number of up steps (resp. down steps) in  $f(w)$  equals the total number of pairs in  $w$ , which completes the proof.  $\square$

**Theorem 2.7.** *The map  $f$  is a bijection from  $S_{n,k}$  to the set  $\mathcal{B}_{n+2(k-1),k-1}$  of dispersed Dyck paths of length  $n + 2(k - 1)$  with exactly  $k - 1$  peaks.*

**Proof.** Let us prove that if  $w$  and  $w'$  are two distinct  $k$ -ary faro words then we have  $f(w) \neq f(w')$ . Let  $i \geq 1$  be the smallest positive integer such that  $w_i \neq w'_i$ . Without loss of generality, we assume  $w_i < w'_i$ . Let us consider the positions of  $w_i$  and  $w'_i$  in the block decomposition of  $w$ .

If  $w_i$  and  $w'_i$  are both in the pairs  $w_i w_{i+1}$  and  $w'_i w'_{i+1}$ , then Remark 2.3 implies that a pair  $w_i x, w_i > x$ , cannot appear to the right of  $w'_i$  in  $w'$ , which implies that  $T_{3(w_i-1)-1} \neq T'_{3(w'_i-1)-1}$ , and thus  $f(w) \neq f(w')$ .

There remain the following cases:

- (i)  $w_i$  or  $w'_i$  is a singleton in  $w$ ,
- (ii)  $w_i$  and  $w'_i$  are both in the pairs  $w_{i-1} w_i$  and  $w'_{i-1} w'_i = w_{i-1} w'_i$ ,
- (iii)  $w_i$  belongs to the pair  $w_{i-1} w_i$  and  $w'_i$  belongs to the pair  $w'_i w'_{i+1}$ ,
- (iv)  $w_i$  and  $w'_i$  are both in the pairs  $w_i w_{i+1}$  and  $w'_{i-1} w'_i$ .

The fact that a faro word avoids 231 in case (i) and Remark 2.3 for cases (ii), (iii) (iv), imply that  $w_i$  cannot appear to the right of  $w'_i$  in  $w'$ . Then the number of  $w_i$  in  $w$ , i.e.  $T_{3(w_i-1)} + T_{3(w_i-1)+1} + T_{3(w_i-1)-1}$ , is different from the number of  $w_i$  in  $w'$ , which is  $T'_{3(w'_i-1)} + T'_{3(w'_i-1)+1} + T'_{3(w'_i-1)-1}$ . Therefore, there is  $\delta \in \{-1, 0, 1\}$  such that  $T_{3(w_i-1)+\delta} \neq T'_{3(w'_i-1)+\delta}$ , which implies that  $f(w) \neq f(w')$ .

Thus,  $f$  is an injective map, and using a cardinality argument (see A124428 in [26]), we conclude that  $f$  is a bijection from  $S_{n,k}$  to  $\mathcal{B}_{n+2(k-1),k-1}$ .  $\square$

Although it is not used in the paper, we could prove that from a given dispersed Dyck path  $P \in \mathcal{B}_{n+2(k-1),k-1}$ ,  $f^{-1}(P)$  can be obtained after applying the following procedure. We refer to Fig. 2.3 for several examples.

We set  $s = 1$  as the initial value. We mark all  $D$ -steps preceded by an  $U$ -step and all the other  $D$ -steps are left unmarked. Reading the steps of  $P$  from left to right:

- If a  $D$ -step is encountered, then skip it.
- If an  $F$ -step is encountered, then write the singleton  $s$ . If the next step is not an  $F$ -step, then update  $s = s + 1$ .
- If an  $U$ -step is encountered in the  $i$ th run of  $U$ -steps, then we distinguish two cases:
  - (i) the next step is  $D$ ; then we skip this  $UD$ -pattern by continuing from the step after  $D$ , if it exists.
  - (ii) the next step is  $U$ ; then we write the pair  $ji$ , where  $j$  is the least integer such that the  $(j - 1)$ -th run of  $D$ -steps has at least one unmarked  $D$ -step. Mark the first unmarked  $D$ -step from the  $(j - 1)$ -th run of  $D$ -steps.

## 2.2. Distribution and popularity of patterns

In this part, we first show how the bijection  $f$  transports pattern statistics on  $S_{n,k}$  into the context of dispersed Dyck paths. After, we deduce multivariate generating functions for the distribution and the popularity of patterns of length two by exploiting the classic recursive decomposition of dispersed Dyck paths.

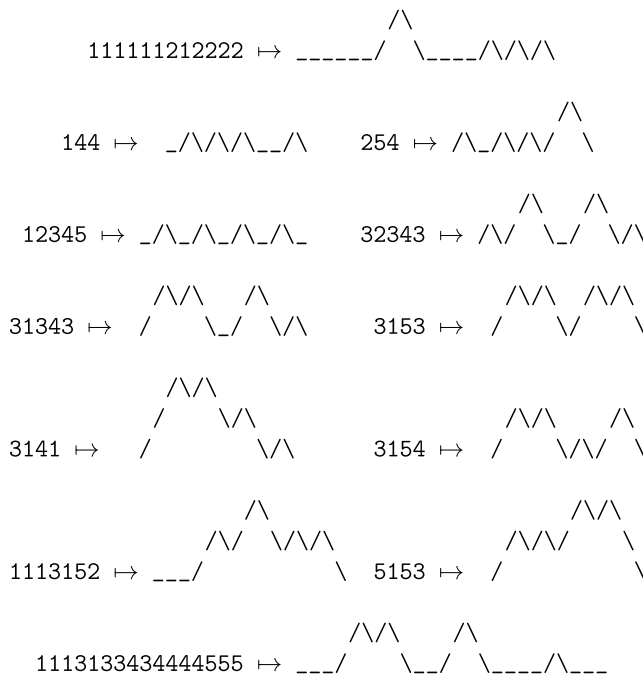


Fig. 2.3. Images of several 5-ary words under bijection  $f$ .

**Theorem 2.8.** For  $n \geq 0$ , the bijection  $f$  from  $\mathcal{S}_{n,k}$  to  $\mathcal{B}_{n+2(k-1),k-1}$  maps statistics associated to patterns of length 2 as follows:

$$\begin{aligned}
 f(\mathbf{11}) &= \mathbf{FF}, \\
 f(\mathbf{21}) &= \mathbf{UU} = \mathbf{DD}, \\
 f(\mathbf{12}) &= \mathbf{DD(UD)^*UU} + \mathbf{DD(UD)^*D} + \mathbf{DD(UD)^*F} + \mathbf{F(UD)^*F} + \mathbf{F(UD)^*UU} \\
 &= \hat{n} - \hat{1} - \mathbf{UU} - \mathbf{FF}.
 \end{aligned}$$

**Proof.** By Remark 2.3, any occurrence of the pattern 11 in a faro word  $w$  is formed by two consecutive singletons  $xx$ . From the definition of the bijection  $f$ , it follows that the number of occurrences of 11 in  $w$  equals the number of occurrences of  $FF$  in  $f(w)$ , that is  $f(\mathbf{11}) = \mathbf{FF}$ .

An occurrence of the pattern 21 in  $w$  is necessarily a pair in the decomposition of  $w$ . Since the length of a maximal run of consecutive up steps is equal to one plus the number of pairs  $yx$  in  $w$  for a given  $x \in [1, n]$ , the number of occurrences of 21 in  $w$  equals the number of occurrences of  $UU$  in  $f(w)$ . On the other hand, any nonempty dispersed Dyck path  $P$  is of the form either  $P = FR$  or  $P = UQDR$  where  $Q$  is a Dyck path and  $R$  a dispersed Dyck path. Reasoning by induction, we obtain that the number of occurrences of  $DD$  equals those of  $UU$  in any dispersed Dyck path, which implies  $f(\mathbf{21}) = \mathbf{UU} = \mathbf{DD}$ .

Now, let us prove the equation  $f(\mathbf{12}) = \mathbf{DD(UD)^*UU} + \mathbf{DD(UD)^*D} + \mathbf{DD(UD)^*F} + \mathbf{F(UD)^*F} + \mathbf{F(UD)^*UU}$ . An occurrence  $xy$  of the pattern 12 occurs in  $w$  as a subblock of one of the following:

- (i) two distinct consecutive pairs  $(ax)(yb)$ ,
- (ii) two equal consecutive pairs  $(yx)(yx)$ ,
- (iii) a pair followed by a singleton  $(ax)(y)$ ,
- (iv) a singleton followed by a pair  $(x)(ya)$ ,
- (v) two distinct singletons  $(x)(y)$ .

For the case (i), we distinguish three subcases.

*Subcase 1.* The occurrence  $xy$  appears in a factor of the form  $(ax)(yb)$  with  $b \geq a$ . This implies that neither a singleton  $s \in [a, b]$  nor a pair  $pq$  with  $p \in (a, b]$  or  $q \in [a, b)$  can appear in  $w$ . Therefore,  $T_{3(s-1)} = 0$  for  $s \in [a, b]$ ,  $T_{3(p-1)-1} = 1$  for  $p \in (a, b]$  and  $T_{3(q-1)+1} = 1$  for any  $q \in [a, b)$ . Thus, between the run of  $D$ -steps associated to  $T_{3(a-1)-1} \geq 2$  and the run of  $U$ -steps associated to  $T_{3(b-1)+1} \geq 2$ , there are no level steps, and the runs of  $D$ -steps and  $U$ -steps are of length one, which creates  $m = b - a \geq 0$  peaks  $UD$ . Hence, the occurrence  $xy$  is associated to an occurrence of the pattern  $DD(UD)^*UU$ .

*Subcase 2.* The occurrence  $xy$  appears in a factor of the form  $(ax)(yb)$  with  $b < a$  and  $a < y$ . This implies that neither a singleton  $x \in [a, y)$  nor a pair  $pq$  with  $p \in (a, y)$  or  $q \in [a, y)$  can appear in the word  $w$ . Therefore,  $T_{3(x-1)} = 0$  for

$x \in [a, y]$ ,  $T_{3(p-1)-1} = 1$  for  $p \in (a, y)$  and  $T_{3(q-1)+1} = 1$  for any  $q \in [a, y)$ . Thus, between the run of  $D$ -steps associated to  $T_{3(a-1)-1} \geq 2$  and the run of  $D$ -steps associated to  $T_{3(y-1)+1} \geq 2$ , there are no level steps, and the runs of  $D$ -steps and  $U$ -steps are of length one, which creates  $m = y - a > 0$  peaks  $UD$ . Hence, the occurrence  $xy$  is associated to an occurrence of the pattern  $DD(UD)^+D$ .

*Subcase 3.* The occurrence  $xy$  appears in a factor of the form  $(ax)(yb)$  with  $b < a$  and  $a \geq y$ . By definition of a faro word, we necessarily have  $a \leq y$ . Thus, we deduce  $a = y$ . So, we have  $T_{3(a-1)-1} \geq 3$ , which counts all consecutive pairs  $az$ ,  $a > z$  in  $w$ . Due to Remark 2.3, all these pairs appear consecutively in  $w$ . Thus, the number of occurrences of the form  $(ax)(ab)$ , for  $x, b$  such that  $x \leq b < a$  is equal to the number of  $DDD = DD(UD)^0D$  patterns in the  $(a - 1)$ -th run of  $D$ -steps in the corresponding dispersed Dyck path. Combining to the subcase 2, the occurrence  $xy$  is associated to an occurrence of the pattern  $DD(UD)^*D$ .

In the case (ii), we have a factor of the form  $(ax)(yb)$  with  $a = y$  and  $x = b$  and the argument from Subcase 3 of case (i) applies. For the remaining cases, (iii) through (v), the occurrence  $xy$  of the pattern 12 is either created by a pair followed by a singleton  $(ax)(y)$ , or by a singleton followed by a pair  $(x)(ya)$ , or by two different singletons  $(x)(y)$ . Arguments similar to the ones given above, allow us to prove that an occurrence  $xy$  in  $w$  corresponds to an occurrence of:

- $DD(UD)^{y-a}F$  for the case  $(ax)(y)$ ,
- $F(UD)^{a-x}UU$  for the case  $(x)(ya)$ , and
- $F(UD)^{y-x}F$  for the case  $(x)(y)$ .

Finally, in any  $n$ -length word we have  $n - 1$  occurrences of 2-length patterns, thus  $\hat{\mathbf{n}} - \hat{\mathbf{1}} = \mathbf{11} + \mathbf{21} + \mathbf{12}$ . Applying the bijection  $f$  to both parts of the equation, we obtain  $f(\mathbf{12}) = \hat{\mathbf{n}} - \hat{\mathbf{1}} - f(\mathbf{11}) - f(\mathbf{21}) = \hat{\mathbf{n}} - \hat{\mathbf{1}} - \mathbf{UU} - \mathbf{FF}$ .  $\square$

**Theorem 2.9.** For  $p \in \{11, 12, 21\}$ , the trivariate generating functions  $F_p(x, y, z)$  where the coefficient at  $x^n y^k z^t$  is the number of  $k$ -ary faro words of length  $n$  containing exactly  $t$  occurrences of the pattern  $p$  are:

$$F_{11}(x, y, z) = \frac{2y(xz - x - 1)}{-xyz + xy + x^3z - x^3 + y - x^2 + xz + x - 1 + (xz - x - 1)A_1},$$

$$F_{21}(x, y, z) = \frac{2y}{-y + x^2z - 2x + 1 + A_2},$$

$$F_{12}(x, y, z) = \frac{y(x^3z^2 - x^3z + x^2z + xyz - xy - 3xz + x + y - 1 + (xz - x + 1)A_2)}{(x^3z^2 - x^3z + x^2z - xyz + xy - xz - x - y + 1 + (xz - x + 1)A_2)(-1 + y)z} + \frac{y}{1 - y},$$

where  $A_1 = \sqrt{x^4 - 2x^2y - 2x^2 + y^2 - 2y + 1}$  and  $A_2 = \sqrt{x^4z^2 - 2x^2yz - 2x^2z + y^2 - 2y + 1}$ .

**Proof.** We have  $f(\mathcal{S}_{n,k}) = \mathcal{B}_{n+2(k-1),k-1}$ . Thus, for any pattern  $p$ , the trivariate generating function  $F_p(x, y, z)$  is given by  $y \cdot B_p(x, \frac{y}{xz}, z)$  where  $B_p(x, y, z)$  is the trivariate generating function whose coefficient at  $x^n y^k z^t$  is equal to the number of dispersed Dyck paths  $P \in \mathcal{B}_{n,k}$  such that  $\mathbf{q}(P) = t$ , where  $\mathbf{q} = f(\mathbf{p})$ .

For  $p = 21$ , Theorem 2.8 has  $f(\mathbf{21}) = \mathbf{UU}$ . Therefore, we decompose the set  $\mathcal{D}$  of Dyck paths as follows:

$$\mathcal{D} = \epsilon \uplus UDD \uplus U(\mathcal{D} \setminus \epsilon)DD.$$

We also decompose the set  $\mathcal{B}$  of dispersed paths as follows:

$$\mathcal{B} = \epsilon \uplus FB \uplus UDB \uplus U(\mathcal{D} \setminus \epsilon)DB.$$

If  $D(x, y, z)$  is the generating function where  $x^n y^k z^t$  is the number of Dyck paths of length  $n$  with  $k$  peaks and  $t$  occurrences of  $UU$ , then the above algebraic equation yields  $D(x, y, z) = 1 + x^2yD(x, y, z) + x^2z(D(x, y, z) - 1)D(x, y, z)$ . If  $B_{21}(x, y, z)$  is the generating function whose coefficient at  $x^n y^k z^t$  is the number of dispersed Dyck paths of length  $n$  with  $k$  peaks and  $t$  occurrences of  $UU$ , then the above decomposition of  $\mathcal{B}$  yields the functional equation

$$B_{21}(x, y, z) = 1 + xB_{21}(x, y, z) + x^2yB_{21}(x, y, z) + x^2z(D(x, y, z) - 1)B_{21}(x, y, z),$$

which, in turn, yields the desired result.

For  $p = 11$ , Theorem 2.8 has  $f(\mathbf{11}) = \mathbf{FF}$ . Therefore, we decompose the set  $\mathcal{D}$  of Dyck paths as follows:

$$\mathcal{D} = \epsilon \uplus UDD \uplus U(\mathcal{D} \setminus \epsilon)DD.$$

We also decompose the set  $\mathcal{B}$  of dispersed Dyck paths as follows:

$$\mathcal{B} = \epsilon \uplus \mathcal{F} \uplus \mathcal{F}UDB \uplus \mathcal{F}U(\mathcal{D} \setminus \epsilon)DB \uplus UDB \uplus U(\mathcal{D} \setminus \epsilon)DB,$$

where  $\mathcal{F}$  is the infinite set of paths  $F^k$  for  $k \geq 1$ . Denote by  $F(x, y, z)$  the generating function for  $\mathcal{F}$ , where its coefficient at  $x^n y^k z^t$  is the number of  $n$ -length paths from  $\mathcal{F}$  having  $k$  peaks and  $t$  occurrences of a pattern  $FF$ . Notice that  $F(x, y, z) = \frac{x}{1-xz}$ . If  $D(x, y)$  is the generating function where the coefficient at  $x^n y^k$  is the number of Dyck paths of length  $n$  with  $k$  peaks, then the above set decomposition yields  $D(x, y) = 1 + x^2 y D(x, y) + x^2 (D(x, y) - 1) D(x, y)$ . Using the second set decomposition of  $\mathcal{B}$ , we obtain a functional equation

$$B_{11}(x, y, z) = 1 + x(1 + z(F(x, y, z) - 1)) + x^3(1 + z(F(x, y, z) - 1))B_{11}(x, y, z) + x^3(1 + z(F(x, y, z) - 1))D(x, y)B_{11}(x, y, z) + x^2 y B_{11}(x, y, z) + x^2(D(x, y) - 1)B_{11}(x, y, z),$$

which provides the result.

For  $p = 12$ , we have, for any  $P \in \mathcal{B}_{n,k}$ , that  $\mathbf{12}(P) = n - 1 - \mathbf{11}(P) - \mathbf{21}(P)$  (that is  $\mathbf{12} = \hat{\mathbf{n}} - \hat{\mathbf{1}} - \mathbf{11} - \mathbf{21}$ ), and thus

$$F_{12}(x, y, z) = \frac{1}{z} \left( F_{11+21} \left( xz, y, \frac{1}{z} \right) - \frac{y}{1-y} \right) + \frac{y}{1-y}.$$

According to Theorem 2.8, we have  $f(\mathbf{11} + \mathbf{21}) = \mathbf{FF} + \mathbf{UU}$ . Therefore, we decompose the set  $\mathcal{B}$  as before for the case of pattern 11, and construct a functional equation by taking into account the different occurrences of  $FF$  and  $UU$ , which yields the claimed result.  $\square$

**Corollary 2.10.** For  $n \geq 0$ , the popularity of pattern  $p \in \{11, 12, 21\}$  in  $S_{n,k}$  is given by the bivariate generating function  $G_p(x, y)$ :

$$G_{11}(x, y) = \frac{4x^2 y}{(1 - y - 2x + x^2 + A_1)^2},$$

$$G_{21}(x, y) = \frac{2x^2 y (1 + y - x^2 - A_1)}{(1 - y - 2x + x^2 + A_1)^2 A_1},$$

$$G_{12}(x, y) = \frac{2xy (A_3 + (x^3 - 2x^2 + 2xy - 2x - 2y + 2)A_1)}{(1 - y - 2x + x^2 + A_1)^2 (1 - y) A_1},$$

where  $A_1 = \sqrt{x^4 - 2x^2 y - 2x^2 + y^2 - 2y + 1}$  and  $A_3 = x^5 - 2x^4 - x^3 y - 3x^3 + 4x^2 y + 4x^2 - 2xy - 2y^2 + 2x + 4y - 2$ .

**Proof.** Using Theorem 2.9, we obtain the result by calculating  $(\frac{\partial}{\partial z} F_p(x, y, z))|_{z=1}$  for  $p \in \{11, 21, 12\}$ .  $\square$

**Corollary 2.11.** For  $p \in \{11, 12, 21\}$ , the bivariate generating functions  $H_p(x, y)$  whose coefficient at  $x^n y^k$  is the number of  $k$ -ary faro words of length  $n$  avoiding the pattern  $p$  are:

$$H_{11}(x, y) = \frac{2y(x+1)}{1 - x - y - xy + x^2 + x^3 + (1+x)\sqrt{(x^2 - 2x - y + 1)(x^2 + 2x - y + 1)}},$$

$$H_{21}(x, y) = \frac{y}{1 - x - y},$$

$$H_{12}(x, y) = \frac{y(-x^3 y + x^2 y - xy^2 + xy + y^2 - 2y + 1)}{xy^3 - 3xy^2 - y^3 + 3xy + 3y^2 - x - 3y + 1}.$$

**Proof.** Note that  $H_p(x, y) = F_p(x, y, 0)$ , where  $F_p(x, y, z)$  is as in Theorem 2.9.  $\square$

Now we discuss the two special cases of  $k = 2$  and  $k = n$ , which correspond respectively to binary words and  $n$ -ary words of length  $n$  (see Table 1 for numerical values).

Case  $k = 2$ : using Corollary 2.10, we can easily prove that the popularity of the pattern 11 in  $S_{n,2}$  generates a shift of the sequence A212964 in [26], which also counts the number of 3-element subsets  $A$  of  $\{1, \dots, n + 1\}$  such that all the sums  $a_1 + a_2$  with  $a_1 \leq a_2$  and  $a_1, a_2 \in A$  are distinct. The popularity of 21 generates a shift of the sequence A006918 where the general term is given by  $\binom{n+3}{3}/4$  if  $n$  is odd, and  $n(n+2)(n+4)/24$  if  $n$  is even. The other patterns do not provide known sequences in [26].

Case  $k = n$ : the sequences of popularity of  $p \in \{11, 21, 12\}$  are not listed in [26], and we have not succeeded in finding a closed form for the diagonal of  $G_p(x, y)$ . However, using the Maple package *gfun* [24], we conjecture that the popularity sequence for 11 satisfies a recurrence equation  $Q_1(n)u_n + Q_2(n)u_{n+1} + Q_3(n)u_{n+2} + Q_4(n)u_{n+3} = 0$ , where  $Q_1, Q_2, Q_3, Q_4$



**Table 1**  
Popularity of patterns  $p$  of length two in  $\mathcal{S}_{n,2}$  and  $\mathcal{S}_{n,n}$ .

$k$	Pattern $p$	Popularity of $p$ in $\mathcal{S}_{n,k}$ for $1 \leq n \leq 9$
2	11	0, 2, 6, 14, 26, 44, 68, 100, 140, ...
	21	0, 1, 2, 5, 8, 14, 20, 30, 40, ...
	12	0,1,4,8,14,22,32,45,60, ...
$n$	11	0,2,12,80,490, 3192, 20076,13094,83655, ...
	21	0,1,8,85, 574,4788,31800, 24489,162305, ...
	12	0,1,16,135,1036, 7700,53964,38646,2636920, ...

are some polynomial functions of degree at most 10, which suggests that the generating function of the diagonal is D-finite when  $p = 11$ . However, we have not succeeded in obtaining a closed form of the diagonal of  $H_{11}(x, y)$ . In contrast, a simple study of the residues (see [27] Section 6.3) of  $H_{21}(x/y, y)$  at the pole  $y_0 = (1 - \sqrt{1 - 4x})/2$  yields the generating function  $(1 - \sqrt{1 - 4x})/(2\sqrt{1 - 4x})$  of the diagonal of  $H_{21}(x, y)$ , and its general term is, therefore,  $\binom{2n-1}{n}$  (see sequence A001700). A similar study for the pattern 12 yields the diagonal  $x(x^3 - 2x^2 + x + 1)/(1 - x)^2$  (here, the pole is  $y_0 = x$ ), which generates the sequence  $u_1 = 1, u_2 = 3, u_n = n$  for  $n \geq 3$ .

Statistic correspondences for other patterns can be obtained using a method similar to that of Theorem 2.8. Therefore, we list directly (without proof) in Theorem 2.12 the  $f$ -images of all statistics associated to a pattern of length three. It is worth noting that the reverse-complement  $\chi$  is a bijection on  $\mathcal{S}_{n,k}$ , which proves that the statistics 112 and 122 (resp. 121 and 212, resp. 132 and 213) have the same distribution on  $\mathcal{S}_{n,k}$ .

**Theorem 2.12.** For  $n \geq 0$ , the bijection  $f$  from  $\mathcal{S}_{n,k}$  to  $\mathcal{B}_{n+2(k-1),k-1}$  translates statistics associated to patterns of length three as follows:

$$\begin{aligned}
 f(111) &= \text{FFF}, \\
 f(112) &= \text{FF(UD)}^+ \text{F} + \text{FF(UD)}^* \text{UU}, \\
 f(122) &= \text{F(UD)}^+ \text{FF} + \text{DD(UD)}^* \text{FF}, \\
 f(121) &= \text{FUU} + \text{UUU}, \\
 f(212) &= \text{DDF} + \text{DDD}, \\
 f(132) &= \text{F(UD)}^+ \text{UU} + \text{U(UD)}^+ \text{UU} + \text{DD(UD)}^* \text{UU}, \\
 f(213) &= \text{DD(UD)}^+ \text{F} + \text{DD(UD)}^+ \text{D} + \text{DD(UD)}^* \text{UU}, \\
 f(123) &= \text{DD(UD)}^* \text{F(UD)}^* \text{UU} + \text{DD(UD)}^* \text{F(UD)}^+ \text{F} + \text{F(UD)}^+ \text{F(UD)}^* \text{UU} + \text{F(UD)}^+ \text{F(UD)}^+ \text{F}, \\
 f(211) &= f(221) = f(231) = f(312) = f(321) = \hat{0}.
 \end{aligned}$$

It would be interesting to see how the method developed in [1,2] could be applied to obtain more pattern distributions in dispersed Dyck paths, but this is beyond the scope of the present paper. The multivariate generating functions for patterns of length three are quite technical, not particularly interesting and laborious to obtain. So, we decide to leave them as an exercise for the reader.

### 3. Patterns in faro permutations

We say that a  $k$ -ary faro word  $w$  of length  $n$  is *injective* (resp. *surjective*) if and only if any value in  $w$  appears only once in  $w$  (resp. any value  $x \in [1, k]$  appears in  $w$ ). A *faro permutation* of length  $n$  is an  $n$ -ary faro word that is both injective and surjective. Let  $\mathcal{P}_n$  be the set of length  $n$  faro permutations. For instance, we have  $\mathcal{P}_3 = \{123, 132, 213\}$ . Since faro permutations are entirely determined by the choice of their values on the odd indices, the cardinality of  $\mathcal{P}_n$  is  $\binom{n}{\lfloor n/2 \rfloor}$ . Note that faro permutations are permutations avoiding the three consecutive patterns 231, 321 and 312 (see Remark 1.3).

**Theorem 3.1.** The bijection  $f$  maps surjective  $k$ -ary faro words of length  $n$  onto dispersed Dyck paths in  $\mathcal{B}_{n+2(k-1),k-1}$  avoiding  $UDUD$  that neither start nor end with  $UD$ .

**Proof.** Using the definition of the bijection  $f$  and in particular the definition of the sequence  $T$ , surjective faro words are those that have a sequence  $T$  satisfying (i)  $T_0 + T_1 > 1$ , (ii)  $T_{3(x-1)-1} + T_{3(x-1)} + T_{3(x-1)+1} > 2$  for  $x \in [2, k - 1]$ , and (iii)  $T_{3(k-1)-1} + T_{3(k-1)} > 1$ . Since  $T_1 \geq 1$ , the condition (i) is equivalent to  $T_0 \neq 0$ , or  $T_0 = 0$  and  $T_1 > 1$ , which means that  $f(w)$  does not start with  $UD$ . Similarly, the condition (iii) is equivalent to the fact that  $f(w)$  does not end with  $UD$ . Since

$T_{3(x-1)-1} \geq 1$  and  $T_{3(x-1)+1} \geq 1$ , the condition (ii) is equivalent to  $T_{3(x-1)} > 0$ , or  $T_{3(x-1)} = 0$  and  $T_{3(x-1)-1} + T_{3(x-1)+1} > 2$ , which means that  $f(w)$  does not contain any occurrence of  $UDUD$ .  $\square$

**Theorem 3.2.** *The bijection  $f$  maps injective  $k$ -ary faro words of length  $n$  into dispersed Dyck paths in  $\mathcal{B}_{n+2(k-1),k-1}$  avoiding the patterns  $FF, DDD, UUU, DDF, FUU$ , and  $DDUU$ .*

**Proof.** Using the definition of  $f$ , injective faro words are those that have a sequence  $T$  satisfying (i)  $T_{3(x-1)} < 2, x \in [1, k]$ ; (ii)  $T_{3(x-1)-1} < 3, x \in [2, k]$ ; (iii)  $T_{3(x-1)+1} < 3, x \in [1, k-1]$ ; (iv)  $T_{3(x-1)-1} + T_{3(x-1)} < 3, x \in [2, k]$ ; (v)  $T_{3(x-1)} + T_{3(x-1)+1} < 3$ , for  $x \in [1, k-1]$ ; and (vi)  $T_{3(x-1)-1} + T_{3(x-1)+1} < 3$ , for  $x \in [2, k-1]$ . It means that  $f(w)$  avoids, respectively, the patterns  $FF, DDD, UUU, DDF, FUU$  and  $DDUU$ .  $\square$

**Theorem 3.3.** *The image by  $f$  of  $\mathcal{P}_n$  is the subset  $\mathcal{B}'_{3n-2,n-1}$  of dispersed Dyck paths in  $\mathcal{B}_{3n-2,n-1}$  that neither start nor end with  $UD$  and where any two consecutive occurrences of  $UD$  are separated by exactly one step.*

**Proof.** The two previous theorems imply that  $f(\mathcal{P}_n)$  is the set of dispersed Dyck paths in  $\mathcal{B}_{3n-2,n-1}$  that neither start nor end with  $UD$  and that avoid the patterns  $FF, DDD, UUU, DDUU, UDUD, DDF$  and  $FUU$ , which is exactly the dispersed Dyck paths that neither start nor end with  $UD$  and where any two consecutive occurrences of  $UD$  are separated by exactly one step. Indeed, in any dispersed Dyck path, the subpath between two consecutive occurrences of  $UD$  is necessarily of the form  $D^i F^j U^k$  for  $i, j, k \geq 0$ . Then, the avoidance of  $FF, DDD$  and  $UUU$  implies that  $i, j, k \leq 1$ , and the avoidance of the other patterns implies that  $i + j + k = 1$  as claimed.  $\square$

Thus, we deduce a one-to-one correspondence  $g$  between length  $n$  faro permutations and dispersed Dyck paths of length  $n$ , where  $g(p)$  is obtained from  $p \in \mathcal{P}_n$  by removing all occurrences of  $UD$  in  $f(p)$ . For instance, if  $p = 1243576$  then  $f(p) = FUDFUDUDDUDFUDUDD$  and  $g(p) = FFUDFUD$ .

**Theorem 3.4.** *For  $n \geq 0$ , the bijection  $g$  from  $\mathcal{P}_n$  to  $\mathcal{B}_n$  transports the pattern statistics as follows:*

$$\begin{aligned} g(\mathbf{21}) &= \mathbf{U}, \\ g(\mathbf{12}) &= \mathbf{DU} + \mathbf{DD} + \mathbf{DF} + \mathbf{FF} + \mathbf{FU}, \\ &= \hat{\mathbf{n}} - \hat{\mathbf{1}} - \mathbf{U}, \\ g(\mathbf{132}) &= \mathbf{FU} + \mathbf{UU} + \mathbf{DU}, \\ g(\mathbf{213}) &= \mathbf{DF} + \mathbf{DD} + \mathbf{DU}, \\ g(\mathbf{123}) &= \mathbf{DFU} + \mathbf{DFF} + \mathbf{FFU} + \mathbf{FFF}, \\ &= \hat{\mathbf{n}} - \hat{\mathbf{2}} - \mathbf{FU} - \mathbf{UU} - 2 \mathbf{DU} - \mathbf{DF} - \mathbf{DD}, \\ g(\mathbf{231}) &= g(\mathbf{312}) = g(\mathbf{321}) = \hat{\mathbf{0}}. \end{aligned}$$

**Proof.** For a faro permutation  $w$ ,  $g(w)$  is obtained from  $f(w)$  by removing all peaks  $UD$ . In  $f(w)$  consecutive occurrences of  $UD$  are separated by one letter exactly, and no  $U^k, D^k$  or  $F^k$  exists for  $k \geq 3$ . The last  $U$  in  $U^2$  and the first  $D$  in  $D^2$  must be a part of an occurrence of  $UD$ . The claimed statistic equations are obtained from Theorem 2.8 and Theorem 2.12 by deleting peaks  $UD$  and replacing all remaining  $U^2$  (resp.  $D^2$ ) with  $U$  (resp.  $D$ ) in all considered pattern statistics. Only two (resp. three) patterns of length 2 (resp. 3) are possible in a faro permutation, namely 12 and 21 (resp. 123, 132 and 213). In any  $n$ -length word there is  $n - 1$  (resp.  $n - 2$ ) occurrences of patterns of length 2 (resp. 3). It follows that  $g(\mathbf{12}) = \hat{\mathbf{n}} - \hat{\mathbf{1}} - \mathbf{U}$  and  $g(\mathbf{123}) = \hat{\mathbf{n}} - \hat{\mathbf{2}} - \mathbf{FU} - \mathbf{UU} - 2 \mathbf{DU} - \mathbf{DF} - \mathbf{DD}$ .  $\square$

**Theorem 3.5.** *For  $p \in \{21, 12, 132, 213, 123\}$ , the bivariate generating functions  $K_p(x, y, z)$ , where the coefficient at  $x^n y^k$  is the number of faro permutations of length  $n$  containing exactly  $k$  occurrences of the pattern  $p$ , are:*

$$\begin{aligned} K_{21}(x, y) &= \frac{2}{1 - 2x + \sqrt{1 - 4x^2y}}, \\ K_{12}(x, y) &= \frac{1 + y + 2xy - 2xy^2 + (y - 1)\sqrt{1 - 4x^2y}}{y(1 - 2xy + \sqrt{1 - 4x^2y})}, \\ K_{132}(x, y) &= \frac{1 + y + (y - 1)\sqrt{1 - 4x^2y}}{y(1 - 2x + \sqrt{1 - 4x^2y})}, \end{aligned}$$

$$K_{213}(x, y) = K_{132}(x, y),$$

$$K_{123}(x, y) = \frac{2 + 3x - 3xy + 2x^2 - 2x^2y - x(1 - y)\sqrt{1 - 4x^2}}{1 - 2xy + \sqrt{1 - 4x^2}}.$$

**Proof.** For  $p = 21$ , Theorem 3.4 has  $g(\mathbf{21}) = \mathbf{U}$ . So, we decompose the set of Dyck paths as  $\mathcal{D} = \epsilon \uplus UDDDD$ , the set of dispersed Dyck paths as  $\mathcal{B} = \epsilon \uplus FB \uplus UDD\bar{B}$ , and obtain the following system:

$$\begin{cases} D(x, y) = 1 + x^2yD^2(x, y), \\ B = 1 + xB(x, y) + x^2yD(x, y)B(x, y), \end{cases}$$

where  $D(x, y)$  (resp.  $B(x, y)$ ) is the generating function for the set of Dyck paths (resp. dispersed Dyck paths) with respect to the number of occurrences of  $U$ . Solving it, we obtain  $K_{21}(x, y) = B(x, y)$ .

Since only the two length 2 patterns (12 and 21) are possible in a faro permutation, we have  $\mathbf{12} = \hat{\mathbf{n}} - \hat{\mathbf{1}} - \mathbf{21}$ . Hence,  $K_{12}(x, y) = (K_{21}(xy, \frac{1}{y}) - 1)/y + 1$ .

Only tree patterns of length 3 are possible in a faro permutation, 123, 132 and 213, so we have  $\mathbf{123} = \hat{\mathbf{n}} - \hat{\mathbf{2}} - \mathbf{132} - \mathbf{213}$ . By Theorem 3.4,  $g(\mathbf{132} + \mathbf{213}) = \mathbf{FU} + \mathbf{UU} + 2\mathbf{DU} + \mathbf{DF} + \mathbf{DD}$ . We decompose the sets of Dyck and dispersed Dyck paths as follows:

$$\begin{cases} \mathcal{D} = \epsilon \uplus UD \uplus U(\mathcal{D} \setminus \epsilon)D \uplus UD(\mathcal{D} \setminus \epsilon) \uplus U(\mathcal{D} \setminus \epsilon)D(\mathcal{D} \setminus \epsilon), \\ \mathcal{B} = \epsilon \uplus \bar{B} \uplus U(\mathcal{D} \setminus \epsilon)D \uplus U(\mathcal{D} \setminus \epsilon)D\bar{B} \uplus U(\mathcal{D} \setminus \epsilon)D(\mathcal{B} \setminus (\epsilon \uplus \bar{B})) \\ \quad \uplus UD \uplus UD\bar{B} \uplus UD(\mathcal{B} \setminus (\epsilon \uplus \bar{B})), \\ \bar{B} = \mathcal{F} \uplus \mathcal{F}(\mathcal{B} \setminus (\epsilon \uplus \bar{B})), \end{cases}$$

where  $\mathcal{F}$  is the set of paths  $F^k$ ,  $k \geq 1$  and  $\bar{B}$  is the set of dispersed Dyck paths starting with a level step. From this decomposition we obtain the following system of functional equations:

$$\begin{cases} D(x, y) = 1 + x^2 + 2x^2y^2(D(x, y) - 1) + x^2y^4(D(x, y) - 1)^2, \\ B(x, y) = 1 + \bar{B}(x, y) + x^2y^2(D(x, y) - 1) + x^2y^3(D(x, y) - 1)\bar{B}(x, y) + \\ \quad + x^2y^4(D(x, y) - 1)(B(x, y) - \bar{B}(x, y) - 1) + x^2 + x^2y\bar{B}(x, y) \\ \quad + x^2y^2(B(x, y) - \bar{B}(x, y) - 1), \\ \bar{B}(x, y) = \frac{x}{1-x} + \frac{xy}{1-x}(B(x, y) - \bar{B}(x, y) - 1), \end{cases}$$

where  $D(x, y)$  (resp.  $B(x, y)$ , resp.  $\bar{B}(x, y)$ ) is the generating function for the set of Dyck paths (resp. dispersed Dyck paths, resp. dispersed Dyck paths starting with  $F$ ) with respect to the statistics  $\mathbf{FU} + \mathbf{UU} + 2\mathbf{DU} + \mathbf{DF} + \mathbf{DD}$ . After solving this system, we obtain the result by evaluating  $K_{123}(x, y) = 1 + x + (B(xy, \frac{1}{y}) - 1 - xy)/y^2$ . Note that it is possible to look directly at  $g(\mathbf{123}) = \mathbf{DFU} + \mathbf{DFF} + \mathbf{FFU} + \mathbf{FFF}$  rather than at  $g(\mathbf{132} + \mathbf{213})$  as we did, but the decomposition will be more complicated.

Note that  $K_{132}(x, y) = K_{213}(x, y)$ , by taking the reverse-complement of faro permutations. We remark that  $\mathbf{DF} + \mathbf{DD} + \mathbf{DU}$  corresponds to the number of occurrences of  $D$  except the last symbol if it is a  $D$ . So, if  $\mathcal{D}$  is the set of Dyck paths and  $\mathcal{B}$  the set of dispersed paths, then the classical decompositions  $\mathcal{D} = \epsilon \uplus DUDD$ ,  $\mathcal{B} = \epsilon \uplus BF \uplus BUDD$  provide the following system of functional equations:

$$\begin{cases} D(x, y) = 1 + x^2yD(x, y)^2, \\ B(x, y) = 1 + xB(x, y) + x^2yB(x, y)D(x, y), \\ \bar{B}(x, y) = 1 + xB(x, y) + x^2B(x, y)D(x, y), \end{cases}$$

where  $D(x, y)$  and  $B(x, y)$  (resp.  $\bar{B}(x, y)$ ) are the generating functions for the sets of Dyck paths and dispersed Dyck paths with respect to the number of  $D$  (resp. number of  $D$  except the last symbol if it is a  $D$ ). Solving the system, we obtain  $K_{132}(x, y) = K_{213}(x, y) = \bar{B}(x, y)$ .  $\square$

**Corollary 3.6.** For  $n \geq 0$ , the popularity of pattern  $p \in \{21, 12, 132, 213, 123\}$  in  $\mathcal{P}_n$  is given by the generating function  $L_p(x)$ :

$$L_{21}(x) = \frac{1 - \sqrt{1 - 4x^2}}{2(1 - 2x)\sqrt{1 - 4x^2}},$$

$$L_{12}(x) = \frac{2x(-1 + 4x^2 + x + \sqrt{1 - 4x^2})}{(1 - 2x)(1 + \sqrt{1 - 4x^2})\sqrt{1 - 4x^2}},$$

**Table 2**  
Popularity of patterns  $p$  of length at most three in faro permutations.

Pattern $p$	Popularity of $p$ in $\mathcal{P}_n$ for $1 \leq n \leq 11$	OEIS
21	0, 1, 2, 7, 14, 38, 76, 187, 374, 874, 1748, ...	A107373
12	0, 1, 4, 11, 26, 62, 134, 303, 634, 1394, 2872, ...	A340567
132, 213	0, 0, 1, 4, 10, 28, 61, 152, 318, 748, 1538, ...	A340568
123	0, 0, 1, 4, 10, 24, 53, 116, 246, 520, 1082, ...	A340569
231,312,321	0, 0, 0, 0, ...	

$$L_{132}(x) = \frac{x(-1 + 4x^2 + 2x + (1 - 2x)\sqrt{1 - 4x^2})}{(1 - 2x)(1 + \sqrt{1 - 4x^2})\sqrt{1 - 4x^2}},$$

$$L_{213}(x) = L_{132}(x),$$

$$L_{123}(x) = \frac{x(1 + 2x)(1 - \sqrt{1 - 4x^2})}{(1 - 2x)(1 + \sqrt{1 - 4x^2})}.$$

**Proof.** We evaluate  $\frac{\partial K_p(x,y)}{\partial y} \Big|_{y=1}$ .  $\square$

**Corollary 3.7.** For  $n \geq 2$ , the popularity of pattern  $p \in \{21, 12, 132, 213, 123\}$  in  $\mathcal{P}_n$  is given by  $\mathbf{p}(\mathcal{P}_n)$ :

$$\mathbf{21}(\mathcal{P}_n) = \frac{n+1}{2} \binom{n}{\lfloor \frac{n}{2} \rfloor} - 2^{n-1} \sim \sqrt{\frac{n}{2\pi}} \cdot 2^n,$$

$$\mathbf{12}(\mathcal{P}_n) = (n-1) \binom{n}{\lfloor \frac{n}{2} \rfloor} - \mathbf{21}(\mathcal{P}_n) \sim \sqrt{\frac{n}{2\pi}} \cdot 2^n,$$

$$\mathbf{123}(\mathcal{P}_n) = 2^n - 2 \binom{n-1}{\lfloor \frac{n-1}{2} \rfloor} - \binom{n}{\lfloor \frac{n}{2} \rfloor} \sim 2^n,$$

$$\mathbf{132}(\mathcal{P}_n) = \mathbf{213}(\mathcal{P}_n) = \frac{1}{2} \left( (n-2) \binom{n}{\lfloor \frac{n}{2} \rfloor} - \mathbf{123}(\mathcal{P}_n) \right) \sim \sqrt{\frac{n}{2\pi}} \cdot 2^n.$$

**Proof.** Recall that  $|\mathcal{P}_n| = \binom{n}{\lfloor n/2 \rfloor}$ . Then we have  $\mathbf{21}(\mathcal{P}_n) + \mathbf{12}(\mathcal{P}_n) = (n-1) \cdot \binom{n}{\lfloor n/2 \rfloor}$  and  $\mathbf{132}(\mathcal{P}_n) + \mathbf{213}(\mathcal{P}_n) + \mathbf{123}(\mathcal{P}_n) = (n-2) \cdot \binom{n}{\lfloor n/2 \rfloor}$ , and considering  $\mathbf{132}(\mathcal{P}_n) = \mathbf{213}(\mathcal{P}_n)$ , it suffices to prove the result for  $\mathbf{21}(\mathcal{P}_n)$  and  $\mathbf{123}(\mathcal{P}_n)$ . Due to Corollary 3.6, we have

$$L_{21}(x) = \frac{W(x)}{2x} - \frac{1}{2(1-2x)}$$

where  $W(x) = \frac{x}{(1-2x)\sqrt{1-4x^2}}$  is the generating function for the sequence A100071 in [26] which has the general term  $\frac{n}{2} \cdot \binom{n-1}{\lfloor (n-1)/2 \rfloor}$ . This induces directly  $\mathbf{21}(\mathcal{P}_n) = \frac{n+1}{2} \cdot \binom{n}{\lfloor n/2 \rfloor} - 2^{n-1}$ .

Similarly, if we expand the numerator of  $L_{123}(x)$  given in Corollary 3.6, then we obtain four generating functions having the general terms respectively equal to  $2^{n-1} - \binom{n-1}{\lfloor (n-1)/2 \rfloor}$ ,  $-\binom{n-1}{\lfloor (n-1)/2 \rfloor}$ ,  $2^{n-1} - \frac{1}{2} \cdot \binom{n}{\lfloor n/2 \rfloor}$  and  $-\frac{1}{2} \cdot \binom{n}{\lfloor n/2 \rfloor}$ , which implies the claimed result. Finally, asymptotics are easily obtained using  $\binom{n}{\lfloor n/2 \rfloor} \sim \sqrt{\frac{2}{\pi n}} \cdot 2^n$ .  $\square$

Using formulae from Corollary 3.7 the following remark can be easily verified.

**Remark 3.8.** The expected number of the occurrences of the pattern 21 (respectively 12, 132 and 213) in a randomly selected faro permutation of length  $n$  is asymptotically equivalent to  $n/2$  when  $n \rightarrow \infty$ . In contrast, the expectation of 123 is asymptotically equivalent to  $\sqrt{\pi n/2}$  and thus the probability that a random faro permutation contains an occurrence of 123 at a random position approaches 0 as  $n$  grows.

Table 2 provides the first values of the popularity of each pattern of length at most three in faro permutations.

#### 4. Some particular subsets of $\mathcal{P}_n$ and $\mathcal{S}_{n,k}$

In this part, we study particular subsets of faro permutations and faro words which are in one-to-one correspondence with other sets of well-known combinatorial objects. Let us recall the definition of a *standard cycle notation* (s.c.n.) and Foata's first fundamental transformation  $\phi$  (see [21]).

In the standard cycle notation (s.c.n.) of a permutation  $w$  each cycle starts with its largest element, and cycles are ordered from left to right in increasing order of their largest elements.

Foata's first fundamental transformation (see [21]) acts on a permutation  $w$  as follows. Write a permutation  $w$  in s.c.n., and then cyclically rearrange every cycle so that it ends with its largest element. Then, reverse each cycle and delete all parentheses. For instance, if  $w = 7321564$ , then the s.c.n. for  $w$  is  $(32)(5)(6)(741)$ , after rearrangement we have  $(23)(5)(6)(417)$ , and thus  $\phi(w) = 3256714$ . If  $w$  is an involution, it contains only cycles of length one or two, rearrangement and reversion are not needed, we directly obtain the image after dropping parentheses in s.c.n.

**Theorem 4.1.** *Foata's first fundamental transformation bijectively maps the set  $\mathcal{I}_n(3-2-1)$  of involutions of length  $n$  avoiding the classical pattern 3-2-1 onto the set  $\mathcal{P}_n$  of faro permutations.*

**Proof.** Let us prove that the standard cycle notation of  $w \in \mathcal{I}_n(3-2-1)$  cannot contain any of the following consecutive cycles:  $(x)(yz)$  with  $z < x$ ,  $(xy)(z)$  with  $z < x$ ,  $(x)(y)(z)$  with  $z < x$ , or  $(xy)(zt)$  with  $t < y$  or  $z < x$ . Assume that  $w \in \mathcal{I}_n(3-2-1)$  and assume towards contradiction that the standard cycle notation (s.c.n.) of  $w$  contains  $(x)(yz)$  with  $z < x$ . Then we have  $x < y$  and thus  $z < x < y$ , which means that the subsequence  $yxz$  (occurring at indices  $z, x, y$ ) is an occurrence of 3-2-1 in  $w$ , a contradiction. Due to the definition of the s.c.n. of  $w$ , the case  $(x)(y)(z)$  with  $z < x$ , the case  $(xy)(z)$  with  $z < x$  and the case  $(xy)(zt)$  with  $z < x$  do not occur since the cycles are arranged in increasing order of their first elements. If the s.c.n. of  $w$  contains  $(xy)(zt)$  with  $t < y$ , then we have  $t < y < x < z$ , which implies that  $w$  contains an occurrence  $zxy$  (at indices  $t, y, x$ ) of 3-2-1, a contradiction. Thus,  $\phi(\mathcal{I}_n(3-2-1)) \subset \mathcal{P}_n$ . Since  $\phi$  is injective, and  $\mathcal{I}_n(3-2-1)$  is also enumerated by ballot numbers  $b_n$  (see for instance [6,25]), we have  $\phi(\mathcal{I}_n(3-2-1)) = \mathcal{P}_n$ .  $\square$

**Remark 4.2.** It is known that Foata's first transformation  $\phi$  maps the statistic of the number of excedances (values  $w_i$  such that  $w_i > i$ ) to the statistic **21** (number of descents  $w_i > w_{i+1}$ ). Therefore, the generating functions  $K_{21}(x, y)$  and  $L_{21}(x)$  in Corollary 3.6 also give the distribution and the popularity of excedances in  $\mathcal{I}_n(3-2-1)$ .

**Remark 4.3.** We could easily check that  $g(\phi(w)) = \Phi(w)$  for  $w \in \mathcal{I}_n(3-2-1)$ , where  $\Phi$  is a bijection in [6] between involutions and labeled Motzkin paths, which also is a restriction of Biane's bijection [9], which in turn is closely related to Françon-Viennot bijection [16].

The next theorem deals with alternating faro permutations, i.e. permutations  $w$  satisfying  $w_1 > w_2 < w_3 > \dots$ . Let  $\mathcal{A}_n$  be the set of alternating faro permutations of length  $n$ .

**Theorem 4.4.** *There is a bijection between  $\mathcal{A}_{2n}$  and the set of Dyck paths of length  $2n$ .*

**Proof.** Let  $w$  be a faro permutation of length  $2n$ . Then,  $w$  is alternating if and only if  $w$  does not contain any singleton in its block decomposition. Due to the definition of  $f$ , this means that  $f(w)$  does not contain any  $F$ -steps and thus,  $g(w)$  is a Dyck path of length  $2n$ , and vice versa.  $\square$

**Theorem 4.5.** *The set  $\mathcal{A}_{2n}$  is exactly the set of length  $2n$  faro derangements, i.e. faro permutations with no fixed point  $w_i = i$  for  $i \in [1, 2n]$ .*

**Proof.** Let  $w = w_1 w_2 \dots w_{2n-1} w_{2n}$  be a faro permutation of length  $2n$ , that is  $w_i < w_{i+2}$  for  $1 \leq i \leq 2n - 2$ . Then,  $w$  is alternating if and only if  $w$  does not contain any singleton in its block decomposition, or equivalently,  $w$  satisfies  $w_i > w_{i+1}$  if  $i$  is odd, and  $w_i < w_{i+1}$  otherwise. This is equivalent to  $w_i$  is greater than  $w_1, w_2, \dots, w_{i-1}$  and  $w_{i+1}$  if  $i$  is odd, and  $w_i$  is smaller than  $w_{i-1}, w_{i+1}, w_{i+2}, \dots, w_{2n}$  if  $i$  is even, which means that  $w_i > i$  if  $i$  is odd and  $w_i < i$  otherwise. Thus, we have  $w_i \neq i$ , and  $w$  is a derangement. This last implication also is an equivalence because it cannot occur  $w_i < i$  with  $i$  odd, or  $w_i > i$  with  $i$  even in a faro derangement  $w$ .  $\square$

**Theorem 4.6.** *Let  $\mathcal{B}_n$  (resp.  $\mathcal{B}'_n$ ) be the set of length  $n$  faro permutations avoiding the classical pattern 2-3-1 (resp. the pattern 3-1-2), then*

- The cardinality of  $\mathcal{B}_n$  is given by the Fibonacci sequence  $f_n$  defined by  $f_n = f_{n-1} + f_{n-2}$  with  $f_1 = 1, f_2 = 2$ .
- We have  $\mathcal{B}_n = \mathcal{B}'_n$ .
- $\mathcal{B}_n$  is exactly the set of length  $n$  faro involutions.

**Proof.** A faro permutation  $w$  avoiding the pattern 2-3-1 is of the form  $1w'$  or  $21w'$ , where  $w'$  also is a faro permutation avoiding 2-3-1. Indeed, if a faro permutation  $w$  starts with  $x > 2$ , then  $w$  starts with  $x1y$  for some  $y > x$ . Then the value 2 is to the right of  $x1y$ , which creates an occurrence  $xy2$  of 2-3-1, a contradiction. Therefore, the cardinality  $f_n$  of length  $n$  faro permutations satisfies  $f_n = f_{n-1} + f_{n-2}$  with  $f_1 = 1, f_2 = 2$ . Using the same argument, faro permutations avoiding 2-3-1 are also faro permutations avoiding 3-1-2.

For the third statement, due to the decomposition of  $w \in \mathcal{B}_n$  (either  $w = 1w'$  or  $w = 21w'$  with  $w' \in \mathcal{B}$ ), we conclude by induction that  $w$  is necessarily a faro involution. Conversely, a faro involution  $w$  avoids the classical pattern 321, which implies that  $w = 1w'$  or  $w = 21w'$  where  $w'$  is also a faro involution (if the first entry  $w_1$  of  $w$  satisfies  $w_1 \geq 3$ , then we have  $w_{w_1} = 1$ ; since  $w$  avoids 321, we necessarily have  $w_i > w_1 \geq 3$  for  $1 \leq i \leq w_1 - 1$ , and thus  $w_{w_1-2} > w_{w_1} = 1$ , which is not possible in a faro permutation). A simple induction implies that  $w \in \mathcal{B}_n$ , which completes the proof.  $\square$

In the following, we consider (for convenience) faro words on the  $n$ -ary alphabet  $[0, n - 1]$ , and we focus on the set of *subexcedent* faro words of length  $n$ , i.e. faro words  $w_1w_2 \dots w_n$  satisfying  $w_i \leq i - 1$  for  $1 \leq i \leq n$ . We make a shift  $[1, n] \rightarrow [0, n - 1]$  on the alphabet in order to apply directly the results presented in [12].

**Theorem 4.7.** *There is a bijection between subexcedent faro words of length  $n$  and 2-1-4-3-avoiding Dumont permutations of the second kind of length  $2n$ .*

We will briefly recall the result given in [12] that enumerated 2-1-4-3-avoiding Dumont permutations of the second kind of length  $2n$ . Dumont permutations of the second kind of length  $2n$  are permutations  $\pi$  that satisfy the following conditions for  $i \in [n]$ :

$$\pi(2i - 1) \geq 2i - 1, \quad \pi(2i) \leq 2i - 1.$$

In other words, the values in the odd positions are weak excedances, whereas the values in the even positions are deficiencies. In addition, if  $\pi$  avoids the pattern 2-1-4-3 (i.e. does not contain a subsequence  $\pi(i_1)\pi(i_2)\pi(i_3)\pi(i_4)$  of length 4 such that  $i_1 < i_2 < i_3 < i_4$  and  $\pi(i_2) < \pi(i_1) < \pi(i_4) < \pi(i_3)$ ), then the values in the even positions of  $\pi$  are exactly  $\{1, 2, \dots, n\}$ , and the values in the odd positions of  $\pi$  are exactly  $\{n + 1, n + 2, \dots, 2n\}$ . Moreover, the subsequence of values of  $\pi$  in the even positions avoids the pattern 2-1-3 while the subsequence of values of  $\pi$  in the odd positions avoids the pattern 1-3-2. This allows [12] to construct a bijection as in Krattenthaler [20] from the even-position subsequence of  $\pi$  to north-east integer lattice paths from  $(0, 0)$  to  $(n, \lfloor n/2 \rfloor)$  staying on or below the line  $y = x/2$ , and from the odd-position subsequence of  $\pi$  to the same paths but ending at  $(n + 1, \lfloor (n + 1)/2 \rfloor)$ . Let  $\{a_n\}_{n \geq 0}$  be the sequence A047749 [26], so that

$$a_{2n} = \frac{1}{2n + 1} \binom{3n}{n}, \quad a_{2n+1} = \frac{1}{n + 1} \binom{3n + 1}{n},$$

then the number of 2-1-4-3-avoiding Dumont permutations of the second kind of length  $2n$  is  $a_n a_{n+1}$ . Thus, to prove Theorem 4.7, we only need to construct a bijection from subexcedent faro words of length  $n$  to ordered pairs of north-east lattice paths on or below the line  $y = x/2$  from  $(0, 0)$  to  $(n, \lfloor n/2 \rfloor)$  and  $(n + 1, \lfloor (n + 1)/2 \rfloor)$ , respectively.

**Proof of Theorem 4.7.** Let  $\pi$  be a subexcedent faro word of length  $n$ . As in [12], let  $\pi_o$  and  $\pi_e$  be the odd-position and even-position subsequences of  $\pi$ . Then  $\pi_o$  and  $\pi_e$  are nondecreasing subsequences such that

$$\begin{aligned} \pi_o(i) &= \pi(2i - 1) \in [0, 2i - 2], \quad i \leq \left\lfloor \frac{n + 1}{2} \right\rfloor, \\ \pi_e(i) &= \pi(2i) \in [0, 2i - 1], \quad i \leq \left\lfloor \frac{n}{2} \right\rfloor. \end{aligned} \tag{4.1}$$

Conversely, any word  $\pi$  whose odd-position and even-position subsequences  $\pi_o$  and  $\pi_e$  satisfy the above properties is a subexcedent faro word of length  $n$ . Given sequences  $\pi_o$  and  $\pi_e$  as in (4.1), associate to them a pair of north-east lattice paths as follows. If  $\pi_o$  or  $\pi_e$  has a letter  $a_i$  in position  $i$ , map such an entry to the point  $(i - 1, a_i)$  in the integer lattice. Let  $k = \lfloor \frac{n+1}{2} \rfloor$  for  $\pi_o$  and  $k = \lfloor \frac{n}{2} \rfloor$  for  $\pi_e$ , and let  $a_{k+1} = 2k$  for  $\pi_o$  and  $a_{k+1} = 2k + 1$  for  $\pi_e$ .

Now consider a north-east lattice path (as in Fig. 4.1 from  $(0, 0)$  to  $(k, a_{k+1})$  through vertices  $(0, a_1), (1, a_2), \dots, (k - 1, a_k)$  in that order so that each vertex is joined to the next one by a (possibly empty) sequence of east steps followed by a (possibly empty) sequence of north steps. In other words, consider the path

$$N^{a_1}, E, N^{a_2 - a_1}, E, N^{a_3 - a_2}, E, \dots, E, N^{a_{k+1} - a_k} \tag{4.2}$$

from  $(0, 0)$  to  $(k, a_{k+1})$ , where  $E = (1, 0)$  is the unit east step and  $N = (0, 1)$  is the unit north step. Then this path lies on or below the line  $y = 2x$  for  $\pi_o$  and on or below the line  $y = 2x + 1$  for  $\pi_e$ , and each such path corresponds to a unique  $\pi_o$  or a unique  $\pi_e$ .

Moreover, notice that if  $n$  is even, then

$$\begin{aligned} \left( \left\lfloor \frac{n + 1}{2} \right\rfloor, 2 \left\lfloor \frac{n + 1}{2} \right\rfloor \right) &= \left( \left\lfloor \frac{n}{2} \right\rfloor, n \right) \\ \left( \left\lfloor \frac{n}{2} \right\rfloor, 2 \left\lfloor \frac{n}{2} \right\rfloor + 1 \right) &= \left( \left\lfloor \frac{n + 1}{2} \right\rfloor, n + 1 \right), \end{aligned}$$

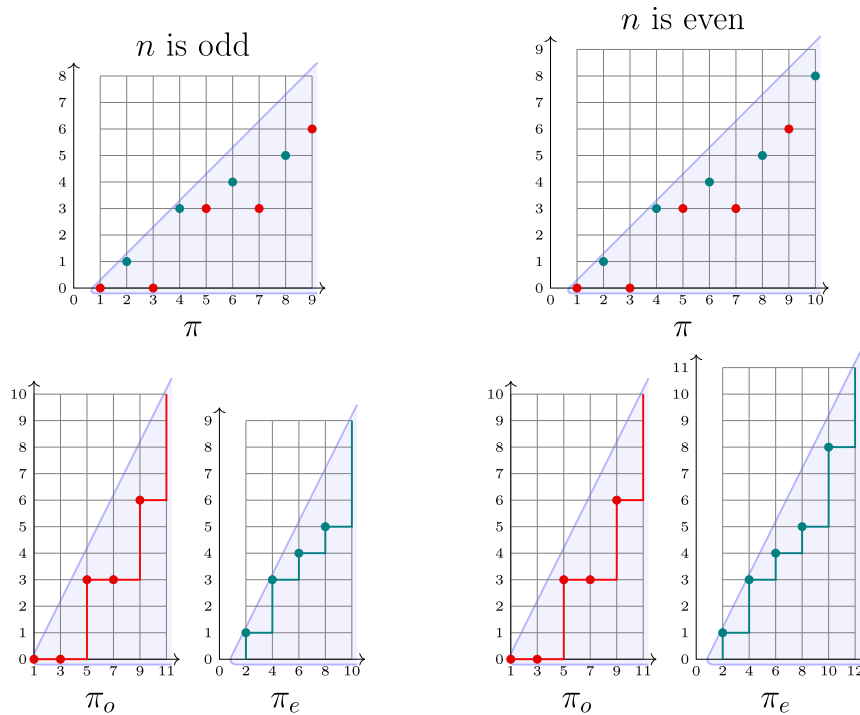


Fig. 4.1. An example of the bijection between a subexcedent faro word and pairs of ternary paths as in the proof of Theorem 4.7.

and if  $n$  is odd, then

$$\left( \left\lfloor \frac{n+1}{2} \right\rfloor, 2 \left\lfloor \frac{n+1}{2} \right\rfloor \right) = \left( \left\lfloor \frac{n+1}{2} \right\rfloor, n+1 \right)$$

$$\left( \left\lfloor \frac{n}{2} \right\rfloor, 2 \left\lfloor \frac{n}{2} \right\rfloor + 1 \right) = \left( \left\lfloor \frac{n}{2} \right\rfloor, n \right).$$

It is easy to see now that the pair of paths thus obtained for  $\pi_o$  and  $\pi_e$  are in bijection with the pair of paths in the proof of the [12, Theorem 3.5] (see also [12, Figure 6]), which yields a bijection between the subexcedent faro words of size  $n$  and 2-1-4-3-avoiding Dumont permutations of the second kind of size  $2n$ .  $\square$

The enumeration of subexcedent faro words may be refined by considering some natural statistics on such words. Together with the bijection of Theorem 4.7 to pairs of ternary paths (or 2-Dyck paths), a recent result [11] lets us find several equidistributed statistics on the odd-position and even-position subsequences of subexcedent faro words.

Recall that a *ternary* (or *2-Dyck*) path is a sequence of unit steps  $u = (1, 1)$  and  $d = (1, -2)$  starting at  $(0, 0)$  and staying in the first quadrant. A *peak* of a 2-Dyck path is an  $ud$ -block in that path, as well as the vertex between the two steps. Likewise, a *double descent* of a 2-Dyck path is a  $dd$ -block in that path, as well as the vertex between the two steps. Define the following statistics on 2-Dyck paths:

- $\mathbf{pk}_0$ , the number of peaks at even height,
- $\mathbf{pk}_1$ , the number of peaks at odd height,
- $\mathbf{dd}$ , the number of double descents.

Then the following results hold.

**Theorem 4.8 ([11]).**

- On 2-Dyck paths ending at height 0, the tristatistic  $(\mathbf{pk}_0 - 1, \mathbf{pk}_1, \mathbf{dd})$  is jointly equidistributed with any of its permutations.
- On 2-Dyck paths ending at height 1, the bistatistics  $(\mathbf{pk}_0, \mathbf{pk}_1)$  and  $(\mathbf{pk}_1, \mathbf{pk}_0)$  are jointly equidistributed.

For a subexcedent faro word  $\pi$  of length  $2n$ , define the following statistics on its odd-position and even-position subsequences  $\pi_o$  and  $\pi_e$ :

- **eOdis**( $\pi$ ), the number of distinct positive even letters in  $\pi_o$  (we exclude 0 since  $\pi_o$  and  $\pi$  always start with 0);
- **oOdis**( $\pi$ ), the number of distinct odd letters in  $\pi_o$ ;
- **aOrpt**( $\pi$ ) =  $\{i \in [n-1] \mid \pi(2i-1) = \pi(2i+1)\}$ , the number of letter repetitions in  $\pi_o$  (the “a” in **aOrpt** stands for “any parity”);
- **eEdis**( $\pi$ ), the number of distinct even letters in  $\pi_e$ ;
- **oEdis**( $\pi$ ), the number of distinct odd letters in  $\pi_e$ .

Then we have the following result.

**Theorem 4.9.** *On subexcedant faro words of length  $n$ ,*

- *the tristatistic (**eOdis**, **oOdis**, **aOrpt**) is jointly equidistributed with any of its permutations.*
- *the bistatistics (**eEdis**, **oEdis**) and (**oEdis**, **eEdis**) are jointly equidistributed.*

**Proof.** For each of  $\pi_o$  and  $\pi_e$ , define  $k$  and  $a_1, a_2, \dots, a_k, a_{k+1}$  as in the proof of Theorem 4.7, and let

$$P = N^{a_1}, E, N^{a_2-a_1}, E, N^{a_3-a_2}, E, \dots, E, N^{a_{k+1}-a_k}$$

be the corresponding north-east path as in (4.2) (when needed, we will distinguish the paths obtained from  $\pi_o$  and  $\pi_e$  as  $P_o$  and  $P_e$ , respectively). Map  $P$  to a lattice path obtained by reversing  $P$  and mapping unit steps  $N \mapsto u = (1, 1)$  and  $E \mapsto d = (1, -2)$ . In other words, consider the map

$$\phi : P \mapsto \phi(P) = u^{a_{k+1}-a_k}, d, u^{a_k-a_{k-1}}, d, \dots, d, u^{a_2-a_1}, d, u^{a_1},$$

where  $\phi(P)$  starts at  $(0, 0)$ . Recall that  $P$  starts at  $(0, 0)$ , stays in the first quadrant on or below  $y = 2x$  for  $\pi_o$  and  $y = 2x + 1$  for  $\pi_e$ , and ends on  $y = 2x$  for  $\pi_o$  and  $y = 2x + 1$  for  $\pi_e$ . Therefore, it is easy to see that  $\phi(P)$  stays in the first quadrant and ends at height 0 for  $\pi_o$  and at height 1 for  $\pi_e$ . Moreover, each distinct letter of  $\pi_o$  or  $\pi_e$  (except for 0 in  $\pi_o$ ) corresponds to a block  $EN$  in the corresponding path  $P$ , which in turn corresponds to a block  $ud$  of  $\phi(P)$ , i.e. to a peak of  $\phi(P)$ .

Furthermore, a repetition of a letter in positions  $i$  and  $i + 1$  of  $\pi_o$  means that  $a_{i+1} = a_i$ , and thus the  $i$ -th and  $(i + 1)$ -st steps  $E$  in  $P$  are adjacent, which in turn corresponds to a block  $dd$  in  $\phi(P)$ . Therefore, **aOrpt**( $\pi$ ) = **dd**( $P_o$ ).

Let  $\ell$  be one of distinct letters of in  $\pi_o$  or  $\pi_e$  (for  $\pi_o$ , also assume  $\ell > 0$ ). Suppose its rightmost occurrence is in position  $j$ . Then there are  $k + 1 - \ell$  east steps and  $a_{k+1} - a_\ell$  north steps in path  $P$  to the right of that point, so the height of the corresponding peak in  $\phi(P)$  is

$$a_{k+1} - a_\ell - 2(k + 1 - \ell) \equiv a_{k+1} - a_\ell \pmod{2} \equiv a_\ell \pmod{2} + a_{k+1} \pmod{2}.$$

It follows that, on  $\pi_o$  (**eOdis**, **oOdis**)( $\pi$ ) = (**pk**<sub>0</sub> - 1, **pk**<sub>1</sub>)( $\phi(P_o)$ ) if  $a_{k+1}$  is even, and (**eOdis**, **oOdis**)( $\pi$ ) = (**pk**<sub>1</sub>, **pk**<sub>0</sub> - 1)( $\phi(P_o)$ ) if  $a_{k+1}$  is odd. Likewise, (**eEdis**, **oEdis**)( $\pi$ ) = (**pk**<sub>0</sub>, **pk**<sub>1</sub>)( $\phi(P_e)$ ) if  $a_{k+1}$  is even, and (**eEdis**, **oEdis**)( $\pi$ ) = (**pk**<sub>1</sub>, **pk**<sub>0</sub>)( $\phi(P_e)$ ) if  $a_{k+1}$  is odd. However, the two statistics on the right-hand side of the equations are jointly equidistributed in each case by Theorem 4.8, and thus the parity of  $a_{k+1}$  is immaterial in each case.  $\square$

From Corollary 1.12 and Equation (2.7) of [11], we can also determine the joint distribution of all the statistics we defined on subexcedent faro words. For this result, we let  $n_o = \lfloor \frac{n+1}{2} \rfloor$  and  $n_e = \lfloor \frac{n}{2} \rfloor$  (so  $n_o + n_e = n$ ). We also let **aErpt**( $\pi$ ) be the number of letter repetitions in  $\pi_e$ , i.e. **aErpt**( $\pi$ ) =  $\{i \in [n-1] \mid \pi(2i) = \pi(2i+2)\}$ .

**Corollary 4.10.** *The number of subexcedent faro words  $\pi$  of length  $n$  such that*

$$(\mathbf{oEdis}, \mathbf{oOdis}, \mathbf{aOrpt}, \mathbf{eEdis}, \mathbf{oEdis}, \mathbf{aErpt})(\pi) = (r_1, r_2, r_2, r_4, r_5, r_6)$$

is

$$\frac{1}{n_o} \binom{n_o}{r_1} \binom{n_o}{r_2} \binom{n_o}{r_3} \frac{r_4 + r_5}{n_e(n_e + 1)} \binom{n_e + 1}{r_4} \binom{n_e + 1}{r_5} \binom{n_e}{r_6}.$$

Note also that  $r_1 + r_2 + r_3 = n_o - 1$  and  $r_4 + r_5 + r_6 = n_e$ .

**Declaration of competing interest**

No conflict of interest.

**Acknowledgements**

We would like to greatly thank anonymous referees for their helpful comments and suggestions.



## References

- [1] A. Asinowski, A. Bacher, C. Banderier, B. Gittenberger, Analytic combinatorics of lattice paths with forbidden patterns, the vectorial kernel method, and generating functions for pushdown automata, *Algorithmica* 82 (3) (2020) 386–428.
- [2] A. Asinowski, C. Banderier, V. Roitner, Generating functions for lattice paths with several forbidden patterns, in: *Séminaire Lotharingien de Combinatoire – Proceedings of the 32nd International Conference on “Formal Power Series and Algebraic Combinatorics”*, 84B, 2020.
- [3] A. Bacher, A. Bernini, L. Ferrari, B. Gunby, R. Pinzani, J. West, The Dyck pattern poset, *Discrete Math.* 321 (2014) 12–23.
- [4] J.-L. Baril, R. Genestier, S. Kirgizov, Pattern distributions in Dyck paths with a first return decomposition constrained by height, *Discrete Math.* 342 (9) (2020) 111995.
- [5] J.-L. Baril, V. Vajnovszki, Popularity of patterns over  $d$ -equivalence classes of words and permutations, *Theor. Comput. Sci.* 814 (April 2020) 249–258.
- [6] M. Barnabei, F. Bonetti, M. Silimbani, Restricted involutions and Motzkin paths, *Adv. Appl. Math.* 47 (1) (2011) 102–115.
- [7] C.S. Barnes, Enumeration of the distinct shuffles of permutations, in: *DMTCS Proceedings, vol. AK, 21st International Conference on Formal Power Series and Algebraic Combinatorics, FPSAC 2009, 2009*, pp. 155–166.
- [8] J. Bertrand, Solution d'un problème, *C. R. Acad. Sci.* 105 (1887) 369.
- [9] P. Biane, Permutations suivant le type d'excédance et le nombre d'inversions et interprétation combinatoire d'une fraction continue de Heine, *Eur. J. Comb.* 14 (4) (July 1993) 277–284.
- [10] M. Bóna, Surprising symmetries in objects counted by Catalan numbers, *Electron. J. Comb.* 19 (1) (March 2012).
- [11] A. Burstein, Distribution of peak heights modulo  $k$  and double descents on  $k$ -Dyck paths, submitted for publication, <https://arxiv.org/pdf/2009.00760.pdf>, 2020.
- [12] A. Burstein, S. Elizalde, T. Mansour, Restricted Dumont permutations, Dyck paths, and noncrossing partitions, *Discrete Math.* 306 (22) (2006) 2851–2869.
- [13] M. Bóna, *A Walk through Combinatorics: An Introduction to Enumeration and Graph Theory*, 2nd ed edition, World Scientific Pub, Hackensack, NJ, 2006.
- [14] E. Deutsch, Dyck path enumeration, *Discrete Math.* 204 (1–3) (1999) 167–202.
- [15] P. Diaconis, R.L. Graham, W.M. Kantor, The mathematics of perfect shuffles, *Adv. Appl. Math.* 4 (2) (June 1983) 175–196.
- [16] J. Françon, G. Viennot, Permutations selon leurs pics, creux, doubles montées et double descentes, nombres d'Euler et nombres de Genocchi, *Discrete Math.* 28 (1) (1979) 21–35.
- [17] B. Hackl, C. Heuberger, H. Prodinger, Counting Ascents in Generalized Dyck Paths, in: *29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms*, vol. 110, AofA 2018, Dagstuhl, Germany, 2018, pp. 26:1–26:15.
- [18] C. Homberger, Expected patterns in permutation classes, *Electron. J. Comb.* 19 (3) (2012), #P43.
- [19] S. Kitaev, *Patterns in Permutations and Words*, Springer Science & Business Media, 2011.
- [20] C. Krattenthaler, Permutations with restricted patterns and Dyck paths, *Adv. Appl. Math.* 27 (2001) 510–530.
- [21] M. Lothaire, *Combinatorics on Words*, *Encyclopedia of Mathematics and Its Applications.*, vol. 17, Addison-Wesley, 1983.
- [22] T. Mansour, Statistics on Dyck paths, *J. Integer Seq.* 9 (1) (2006) 06.
- [23] S.B. Morris, The basic mathematics of the faro shuffle, *Delta Pi Epsilon J.* 6 (2) (1975) 85–92.
- [24] B. Salvy, P. Zimmermann, Gfun, A Maple package for the manipulation of generating and holonomic functions in one variable, *ACM Trans. Math. Softw.* 20 (2) (June 1994) 163–177.
- [25] R. Simion, F. Schmidt, Restricted permutations, *Eur. J. Comb.* 6 (1985) 383–406.
- [26] N.J.A. Sloane, et al., *The On-line Encyclopedia of Integer Sequences*, 2020, <https://oeis.org>.
- [27] R. Stanley, *Enumerative Combinatorics*, vol. 1, Cambridge University Press, 1997.
- [28] W.A. Whitworth, Arrangements of  $m$  things of one sort and  $n$  things of another sort, under certain conditions of priority, *Mess. Math* 8 (1878) 105–114.