Pattern statistics in faro words and permutations

Jean-Luc Baril[1], Alexander Burstein[2], and Sergey Kirgizov[1]

[1]LIB, Université de Bourgogne Franche-Comté
B.P. 47 870, 21078 Dijon Cedex France
E-mails: {barjl, sergey.kirgizov}@u-bourgogne.fr

[2]Howard University, Department of Mathematics
Washington, DC 20059, USA
E-mail: aburstein@howard.edu

October 13, 2020

Abstract

We study the distribution and the popularity of some patterns in words obtained by interlacing the letters of the two nondecreasing k-ary words of lengths differing by at most one. We present a bijection between these words and dispersed Dyck paths with a given number of peaks. We show how the bijection maps statistics of consecutive patterns into linear combinations of other pattern statistics on paths. 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.

1 Introduction

The faro shuffle is a well known technique to shuffle a deck of cards [13, 20]. 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. For two k-ary words u and v such that $0 \leqslant |u| - |v| \leqslant 1$, 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: $u_1v_1u_2v_2u_3v_3\ldots$. A *k-ary faro word* is a faro shuffle of two nondecreasing k-ary words. 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 $|\mathcal{S}_{4,2}| = 9$. A *faro permutation* of length n is a n-ary faro word of length n that contains every letter in the interval $[1, n]$ 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) $p = p_1 p_2 \dots p_k$ if there does not exist a strictly increasing sequence of indices $i = i_1 i_2 \dots i_k$ with (resp. one with $i_{j+1} = i_j + 1$ for $1 \leqslant j \leqslant k-1$) such that $w_{i_1} w_{i_2} \dots w_{i_k}$ is order-isomorphic to p (see [16] for instance). Obviously, any faro word avoids the classical (and thus consecutive) pattern 321. Let $Av_n(\sigma)$ denote the set of permutations avoiding a classical pattern σ, then we have $\mathcal{P}_n \subset Av_n(321)$ for $n \geqslant 0$ and $\mathcal{P}_n \neq Av(321)$ for $n \geqslant 3$. In the following, we almost always consider consecutive patterns. Therefore, for the sake of brevity we distinguish the two kinds of pattern avoidance by omitting *consecutive* and by adding *classical* when necessary.

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. *Dispersed Dyck paths* (see [15]) 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 such that all level steps are on the x-axis. The empty path is denoted by ε. Let $\mathcal{B}_n$ be the set of dispersed Dyck paths of length n (or, equivalently, consisting of n steps) and set $\mathcal{B} = \cup_{n\geqslant 0}\mathcal{B}_n$. A *Dyck path* of semilength $n \geqslant 0$ 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 $\mathcal{D} = \bigcup_{n\geqslant 0} \mathcal{D}_n$. Dispersed Dyck paths of length n are in straightforward bijection with prefixes of Dyck paths of length n, also known as ballot paths [7, 25]. 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 [23], 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 [12, 19]). 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 pattern X, an occurrence of the pattern X^+ in a path P is a maximal sequence of consecutive repetitions of X, which is also denoted X^k for some $k \geqslant 1$. More generally, for two possibly empty patterns Y and Z, the pattern YX^+Z corresponds to a pattern of the form YX^kZ obtained by concatenation of Y, X^k and Z for $k \geqslant 1$. Similarly, for any not simultaneously empty patterns Y and Z, an occurrence of YX^*Z corresponds to an occurrence YX^kZ with $k \geqslant 0$. For instance, the path $FUDUDFFUDF$ contains two occurrences of the pattern $F(UD)^+F$ and three occurrences of $F(UD)^*F$.

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 $\mathbf{p} : \mathcal{A} \to \mathbb{N}$ 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 returning the constant value n is denoted by $\mathbf{n}$, and the statistic giving the number of occurrences of the pattern 123 (resp. $UDUD$) in a word (resp. a lattice path) is denoted by $\mathbf{123}$ (resp. $\mathbf{UDUD}$). The *popularity* of a pattern p in $\mathcal{A}$ is the total number of occurrences of p over all objects of A, that is $\mathbf{p}(A) = \sum_{a\in\mathcal{A}} \mathbf{p}(a)$ (see [5, 9, 16]). For instance, for a dispersed Dyck path $P = FFUDFUUDUUUDDDD$ we have $\mathbf{FF}(P) = 1$, $\mathbf{DDD}(P) = 2$, $\mathbf{UD}(P) = 3$ and $\mathbf{UUUU}(P) = 0$. 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. Then, the statistic $\mathbf{X}^+$ on the lattice paths will be equal to $\sum_{k\geqslant 1} \mathbf{X}^k$ where $\mathbf{X}^k$ is the statistic giving the number of occurrences of maximal X^k, $k \geqslant 1$ and $\mathbf{YX^+Z} = \sum_{k\geqslant 1} \mathbf{YX^kZ}$. Let $\mathcal{A}$ and $\mathcal{B}$ be two sets of combinatorial objects, and let $\mathbf{T}_\mathcal{A}$ and $\mathbf{T}_\mathcal{B}$ be the corresponding statistic sets. 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 $f : \mathcal{A} \to \mathcal{B}$ 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.

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 classical pattern 321. Finally, we prove that the set of subexcedent words in $S_{n,n}$ is related to ternary trees and Dumont permutations of the second kind [11] avoiding the classical pattern 2143, 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.

An n-long k-ary word $w = w_1 w_2 \dots w_n$ is a faro word if and only if $w_i \leqslant w_{i+2}$ for any $i \in [1, n-2]$. A *pair* in a faro word w, is an occurrence $w_i w_{i+1}$ with $w_i > w_{i+1}$. 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

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

We define an order relation on $\mathcal{L}_k$ as follows: for $p, q \in \mathcal{L}_k$,

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

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

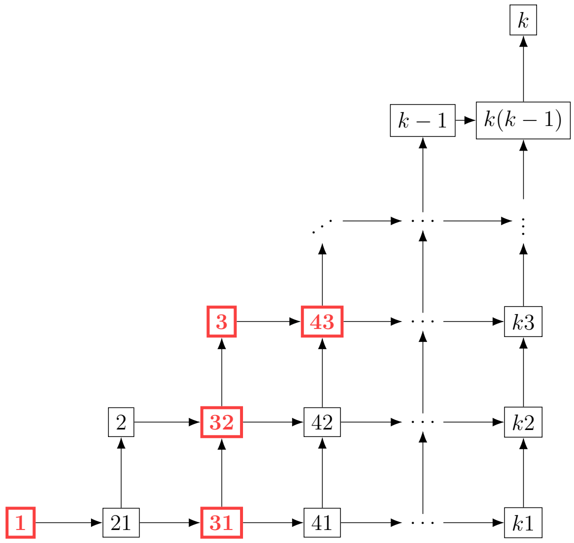

Figure 2.1: The faro lattice $(\mathcal{L}_k, \preceq)$. Red blocks represent the multichain associated to the k-ary faro word $11313232343 = 11^2(31)(32)^2 3(43)$.

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

Remark 2.1. There is a one-to-one correspondence between k-ary faro words and the multichains of $\mathcal{L}_k$. Indeed, 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 $b_1 \preceq b_2 \preceq \dots \preceq b_\ell$ in $\mathcal{L}_k$, and vice versa.

For instance, the faro word $11313232343 = 11(31)(32)(32)3(43)$ corresponds to the multichain $1 \preceq 1 \preceq 31 \preceq 32 \preceq 32 \preceq 3 \preceq 43$ (see Figure 2.1).

Remark 2.2. If a k-ary faro word w contains a singleton x in its decomposition into blocks, then it satisfies the following property: the number of pairs of the form ab, $b < a \leqslant x$, equals the number of pairs of the form cd, $d \leqslant x - 1$.

2.1 A bijection to the set of dispersed Dyck paths

As mentioned by E. Deutsch in [23] (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 n-long k-ary faro words to the set $B_{n+2(k-1),k-1}$ 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 $0 \leqslant i \leqslant 3(k-1)$ as follows:

– if $i = 3(x-1)$ then T_i is the number of occurrences of 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.

For instance, the images by f of the 5-ary words $\epsilon, 12345, 3141, 111111212222$ are, respectively, $UDUDUDUD$, $FUDFUDFUDFUDF$, $UUDUDDUDDUD$ and $FFFFFFUUDDFFFFUDUDUD$. 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 ϵ is UD when $k = 2$, while $f(\epsilon) = UDUD$ for $k = 3$. We refer to Figure 2.2 for an example of this bijection.

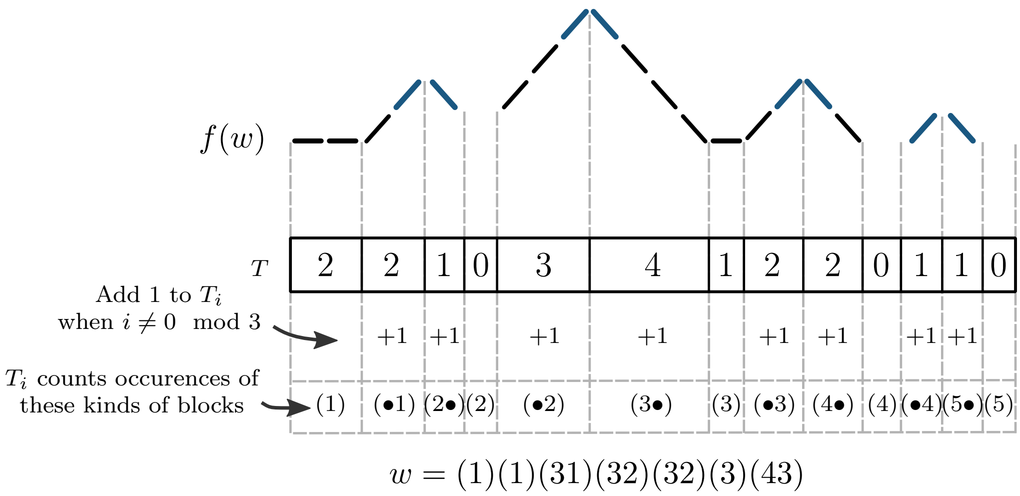

Figure 2.2: The image by f of the 5-ary faro word $w = 11313232343$ is $f(w) = FFUUDUUUDDDDFUUDDUD$.

The sequence $T = T_0T_1 \dots T_{3(k-1)}$ is a run-length-like encoding of the path $f(w)$, and $f(w)$ can be constructed from w using a linear time algorithm. Values of T at indices $i = 0 \mod 3$ correspond to the lengths of maximal runs of consecutive level steps. Values at indices $i = 1 \mod 3$ (resp. $i = 2 \mod 3$) correspond to the lengths of maximal runs of consecutive up (resp. down) steps. Since for any $i \neq 0 \mod 3$, $1 \leqslant i \leqslant 3(k-1) - 1$ we have $T_i \geqslant 1$, the path w contains exactly $k - 1$ peaks UD. Interpreting Remark 2.2 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, $1 \leqslant j < i \leqslant x + 2$, in w, and u_x equals the number of pairs ij, $1 \leqslant j \leqslant x + 1$, $i \geqslant j + 1$, which implies that $d_x \leqslant u_x$. Also by definition, the total number of up steps (resp. down steps) in $f(w)$ equals the total number of pairs in w. Combining all these observations, $f(w)$ is necessarily a dispersed Dyck path of length $n + 2(k-1)$ with exactly $k - 1$ peaks.

Theorem 2.3. *The map f is a bijection from $\mathcal{S}_{n,k}$ 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 $i \geqslant 1$ 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 position 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 2.1 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 2.1 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 [23]), we conclude that f is a bijection from $\mathcal{S}_{n,k}$ to $\mathcal{B}_{n+2(k-1),k-1}$. □

Note that from a given dispersed Dyck path $P \in \mathcal{B}_{n+2(k-1),k-1}$, we can obtain $f^{-1}(P)$ after applying the following procedure. 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 a F-step is encountered, then write the singleton s. If the next step is not a level step, then update $s = s + 1$.
- If a U-step is encountered in the ith 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 $\mathcal{S}_{n,k}$ 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.

Theorem 2.4. *For $n \geqslant 0$, 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:*

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

Proof. By Remark 2.1, 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 $f(w)$, which is also the number of occurrences of DD in $f(w)$. Thus, $f(\mathbf{21}) = \mathbf{UU} = \mathbf{DD}$.

An instance 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 $b \geqslant a$. This implies that neither a singleton $x \in [a, b]$ nor a pair pq with $p \in]a, b]$ or $q \in [a, b[$ can appear in w. Therefore, $T_{3(x-1)} = 0$ for $x \in [a, b]$, $T_{3(p-1)-1} = 1$ 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 $T_{3(a-1)-1} \geqslant 2$ and the run of U-steps associated to $T_{3(c-1)+1} \geqslant 2$, there are no level steps, and the runs of D-steps and U-steps are of length one, which creates $m = b - a \geqslant 0$ peaks UD. Hence, the occurrence xy is associated to an occurrence of the pattern $DD(UD)^m UU$.

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 $T_{3(a-1)-1} \geqslant 2$ and the run of D-steps associated to $T_{3(y-1)+1} \geqslant 2$, there are no level steps, and the runs of D-steps and U-steps are of length one, which creates $m = y - a \geqslant 0$ peaks UD. Hence, the occurrence xy is associated to an occurrence of the pattern $DD(UD)^m D$.

Subcase 3. The occurrence xy appears in a factor of the form $(ax)(yb)$ with $y = a$. So, we have $T_{3(a-1)-1} \geqslant 3$, which counts all consecutive pairs $az, a > z$ in w. Due to Remark 2.1, all these pairs appear consecutively in w. Thus, the number of occurrences of the form

$(ax)(ab)$, for x, b such that $x \leqslant b < a$ 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.

For the remaining cases, (ii) 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 $(x)(y)$. □

Theorem 2.5. *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:*

$$F_{11}(x,y,z) = \frac{2y\,(xz-x-1)}{-xyz+xy+x^3z-x^3+y-x^2+xz+x-1+(xz-x-1)A_1},$$
$$F_{21}(x,y,z) = \frac{2y}{-y+x^2z-2\,x+1+A_2},$$
$$F_{12}(x,y,z) = \frac{y\left(x^3z^2-x^3z+x^2z+xyz-xy-3\,xz+x+y-1+(xz-x+1)A_2\right)}{\left(x^3z^2-x^3z+x^2z-xyz+xy-xz-x-y+1+(xz-x+1)A_2\right)(-1+y)\,z}+\frac{y}{1-y},$$

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

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 $y \cdot B_p(x, \frac{y}{x^2}, z)$ 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 2.4 has $f(\mathbf{21}) = \mathbf{UU}$. Therefore, we decompose the set $\mathcal{D}$ of Dyck paths as follows:

$$\mathcal{D} = \epsilon + UD\mathcal{D} + U(\mathcal{D}\setminus\epsilon)D\mathcal{D}.$$

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

$$\mathcal{B} = \epsilon + F\mathcal{B} + U(\mathcal{D}\setminus\epsilon)D\mathcal{B}.$$

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

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

which, in turn, yields the desired result.

For $p = 11$, Theorem 2.4 has $f(\mathbf{11}) = \mathbf{FF}$. Therefore, we decompose the set $\mathcal{D}$ of Dyck paths as follows:

$$\mathcal{D} = \epsilon + UD\mathcal{D} + U(\mathcal{D}\setminus\epsilon)D\mathcal{D}.$$

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

$$\mathcal{B} = \epsilon + \mathcal{F} + \mathcal{F}UD\mathcal{B} + \mathcal{F}U(\mathcal{D} \setminus \epsilon)D\mathcal{B} + UD\mathcal{B} + U(\mathcal{D} \setminus \epsilon)D\mathcal{B},$$

where $\mathcal{F}$ is the infinite set of paths F^k for $k \geqslant 1$. 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

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

which provides the result after noticing that the generating function for $\mathcal{F}$ is given by $F(x,y,z) = \frac{x}{1-xz}$.

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 $\mathbf{12} = \mathbf{n} - \mathbf{1} - \mathbf{11} - \mathbf{21}$), 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 2.4, 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 expected result. □

Corollary 2.6. *For $n \geqslant 0$, the popularity of pattern $p \in \{11, 12, 21\}$ in $\mathcal{S}_{n,k}$ is given by the bivariate generating function $G_p(x,y)$:*

$$\begin{aligned} G_{11}(x,y) &= \frac{4x^2 y}{(1 - y - 2x + x^2 + A_1)^2}, \\ G_{21}(x,y) &= \frac{2x^2 y\,(1 + y - x^2 - A_1)}{(1 - y - 2x + x^2 + A_1)^2\, A_1}, \\ G_{12}(x,y) &= \frac{2xy\,(A_3 + (x^3 - 2x^2 + 2xy - 2x - 2y + 2)A_1)}{(1 - y - 2x + x^2 + A_1)^2\,(1 - y)\, A_1}, \end{aligned}$$

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

Proof. Using Theorem 2.5, we obtain the result by calculating $\left(\frac{\partial}{\partial z} F_p(x,y,z)\right)\Big|_{z=1}$ for $p \in \{11, 21, 12\}$. □

Corollary 2.7. *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:*

$$H_{11}(x,y) = \frac{2y\,(x+1)}{1-x-y-xy+x^2+x^3+(1+x)\sqrt{(x^2-2x-y+1)(x^2+2x-y+1)}},$$
$$H_{21}(x,y) = \frac{y}{1-x-y},$$
$$H_{12}(x,y) = \frac{y\,(-x^3y+x^2y-xy^2+xy+y^2-2y+1)}{xy^3-3xy^2-y^3+3xy+3y^2-x-3y+1}.$$

Proof. Note that $H_p(x,y) = F_p(x,y,0)$, where $F_p(x,y,z)$ is as in Theorem 2.5. □

Now we discuss the two special cases of $k = 2$ and $k = n$.

Case $k = 2$: the popularity of the pattern 11 in $\mathcal{S}_{n,2}$ generates a shift of the sequence A212964 in [23], 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 $a_1 \leqslant a_2$ 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 [23].

Case $k = n$: the sequences of popularity of $p \in \{11, 21, 12\}$ are not listed in [23], and we have not succeeded in finding a closed form for the diagonal of $G_p(x,y)$. However, using the Maple package *gfun* [21], 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 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 [24] 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 $n \geqslant 3$.

Statistic correspondences for other patterns can be obtained using a method similar to that of Theorem 2.4. Therefore, we list directly (without proof) in Theorem 2.8 the f-images of all statistics associated to a pattern of length three. It is worth noting that the reverse-complement χ 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 2.8. *For $n \geqslant 0$, 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:*

$$f(\mathbf{111}) = \mathbf{FFF},$$
$$f(\mathbf{112}) = \mathbf{FF(UD)^+F} + \mathbf{FF(UD)^*UU},$$
$$f(\mathbf{122}) = \mathbf{F(UD)^+FF} + \mathbf{DD(UD)^*FF},$$
$$f(\mathbf{121}) = \mathbf{FUU} + \mathbf{UUU},$$
$$f(\mathbf{212}) = \mathbf{DDF} + \mathbf{DDD},$$
$$f(\mathbf{132}) = \mathbf{F(UD)^+UU} + \mathbf{U(UD)^+UU} + \mathbf{DD(UD)^*UU},$$
$$f(\mathbf{213}) = \mathbf{DD(UD)^+F} + \mathbf{DD(UD)^+D} + \mathbf{DD(UD)^*UU},$$

$$f(\mathbf{123}) = \mathbf{DD(UD)^*F(UD)^*UU} + \mathbf{DD(UD)^*F(UD)^+F} + \mathbf{F(UD)^+F(UD)^*UU} + \mathbf{F(UD)^+F(UD)^+F},$$
$$g(\mathbf{211}) = g(\mathbf{221}) = g(\mathbf{231}) = g(\mathbf{312}) = g(\mathbf{321}) = \mathbf{0}.$$

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.

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 a 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 its values on the odd indices, the cardinality of $\mathcal{P}_n$ is $\binom{n}{\lfloor n/2 \rfloor}$.

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) $T_{3(x-1)-1} + T_{3(x-1)} + T_{3(x-1)+1} > 2$, and (iii) $T_{3(k-1)-1} + T_{3(k-1)} > 1$. Since $T_1 \geqslant 1$, 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. The condition (iii) is similar equivalent to the fact that $f(w)$ does not end with UD. Since $T_{3(x-1)-1} \geqslant 1$ and $T_{3(x-1)+1} \geqslant 1$, 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 $UDUD$. □

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) $T_{3(x-1)} < 2$, (ii) $T_{3(x-1)-1} < 3$, (iii) $T_{3(x-1)+1} < 3$, (iv) $T_{3(x-1)-1} + T_{3(x-1)} < 3$, and (v) $T_{3(x-1)} + T_{3(x-1)+1} < 3$, which means that $f(w)$ avoids, respectively, the patterns FF, DDD, UUU, DDF and FUU. □

Theorem 3.3. *The image by f of $\mathcal{P}_n$ is the subset $B'_{3n-2,n-1}$ of dispersed Dyck paths in $B_{3n-2,n-1}$ 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 $B_{3n-2,n-1}$ 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. □

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 $f(p) = FUDFUDUUDDUDFUDUUDD$ and $g(p) = FFUDFUD$.

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

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

Proof. For a faro permutation w, $g(w)$ is obtained from $f(w)$ by removing all peaks UD. Thus, the statistic equations are obtained from Theorem 2.4 and Theorem 2.8 by deleting peaks and replacing UU with U and DD with D. □

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

$$\begin{aligned}
K_{21}(x,y) &= \frac{2}{1-2x+\sqrt{1-4\,x^2y}},\\
K_{12}(x,y) &= \frac{1+y+2xy-2\,xy^2+(y-1)\sqrt{1-4\,x^2y}}{y\left(1-2\,xy+\sqrt{1-4\,x^2y}\right)},\\
K_{132}(x,y) &= \frac{1+y+(y-1)\sqrt{1-4\,x^2y}}{y\left(1-2x+\sqrt{1-4\,x^2y}\right)},\\
K_{213}(x,y) &= K_{132}(x,y),\\
K_{123}(x,y) &= \frac{2+3x-3xy+2x^2-2x^2y-x(1-y)\sqrt{1-4x^2}}{1-2xy+\sqrt{1-4x^2}}.
\end{aligned}$$

Proof. For $p = 21$, Theorem 3.4 has $g(\mathbf{21}) = \mathbf{U}$. So, we decompose the set of Dyck paths as $\mathcal{D} = \epsilon + U\mathcal{D}D\mathcal{D}$, the set of dispersed Dyck paths as $\mathcal{B} = \epsilon + F\mathcal{B} + U(\mathcal{D}\setminus\epsilon)D\mathcal{B}$, 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 $\mathbf{12} = \mathbf{n} - \mathbf{1} - \mathbf{21}$. Hence, $K_{12}(x,y) = (K_{21}(xy, \frac{x}{y}) - 1)/y + 1$.

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

$$\begin{cases} \mathcal{D} &= \epsilon + UD + U(\mathcal{D}\setminus\epsilon)D + UD(\mathcal{D}\setminus\epsilon) + U(\mathcal{D}\setminus\epsilon)D(\mathcal{D}\setminus\epsilon), \\ \mathcal{B} &= \epsilon + \overline{\mathcal{B}} + U(\mathcal{D}\setminus\epsilon)D + U(\mathcal{D}\setminus\epsilon)D\overline{\mathcal{B}} + U(\mathcal{D}\setminus\epsilon)D(\mathcal{B}\setminus(\epsilon\cup\overline{\mathcal{B}})) \\ & \quad + UD + UD\overline{\mathcal{B}} + UD(\mathcal{B}\setminus(\epsilon\cup\overline{\mathcal{B}})), \\ \overline{\mathcal{B}} &= \mathcal{F} + \mathcal{F}(\mathcal{B}\setminus(\epsilon\cup\overline{\mathcal{B}})), \end{cases}$$

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

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

where $D(x,y)$ (resp. $B(x,y)$, resp. $\overline{B}(x,y)$) 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 $K_{132}(x,y) = K_{213}(x,y)$, by taking the reverse-complement of faro permutations. Using the same decomposition as for the previous case, we write the following system for the pattern 213:

$$\begin{cases} D(x,y) &= 1 + x^2 + 2x^2y\left(D(x,y)-1\right) + x^2y^2\left(D(x,y)-1\right)^2, \\ B(x,y) &= 1 + \overline{B}(x,y) + x^2y\left(D(x,y)-1\right) + x^2y\left(D(x,y)-1\right)\overline{B}(x,y) + \\ & \quad + x^2y^2\left(D(x,y)-1\right)\Big(B(x,y) - \overline{B}(x,y) - 1\Big) + x^2 + x^2\overline{B}(x,y) \\ & \quad + x^2y\Big(B(x,y) - \overline{B}(x,y) - 1\Big), \\ \overline{B}(x,y) &= \frac{x}{1-x} + \frac{xy}{1-x}\Big(B(x,y) - \overline{B}(x,y) - 1\Big), \end{cases}$$

where $D(x,y)$ (resp. $B(x,y)$, resp. $\overline{B}(x,y)$) 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 $\mathbf{DF} + \mathbf{DD} + \mathbf{DU}$. Solving the system, we obtain $K_{132}(x,y) = K_{213}(x,y) = B(x,y)$. $\square$

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

$$\begin{aligned}
L_{21}(x) &= \frac{1-\sqrt{1-4x^2}}{2(1-2x)\sqrt{1-4x^2}},\\
L_{12}(x) &= \frac{2x\left(-1+4\,x^2+x+\sqrt{1-4\,x^2}\right)}{(1-2x)(1+\sqrt{1-4x^2})\sqrt{1-4\,x^2}},\\
L_{132}(x) &= \frac{x\left(-1+4x^2+2x+(1-2x)\sqrt{1-4x^2}\right)}{(1-2x)(1+\sqrt{1-4x^2})\sqrt{1-4x^2}},\\
L_{213}(x) &= L_{132}(x),\\
L_{123}(x) &= \frac{x(1+2x)(1-\sqrt{1-4x^2})}{(1-2x)(1+\sqrt{1-4x^2})}.
\end{aligned}$$

Table 1 provides the first values of the popularity of each pattern of length at most three in faro permutations.

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

Table 1: Popularity of patterns p 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.

Theorem 4.1. *Foata's first fundamental transformation bijectively maps the set $\mathcal{I}(321)$ of involutions avoiding the classical pattern 321 onto the set $\mathcal{P}_n$ of faro permutations.*

Proof. Recall how Foata's first fundamental transformation ϕ acts on permutations (see [18]). Given a permutation w, write it in the standard form for cycle notation (i.e. each cycle starts with its largest element, and cycles are ordered from left to right in increasing order of their first elements), then delete all parentheses. For instance, if $w = 31254$, then the standard cycle notation for w is $(321)(54)$, and thus, $\phi(w) = 32154$. If w is an involution then its standard decomposition contains only cycles of length one or two. Now, let us prove that the standard cycle notation of $w \in \mathcal{I}(321)$ 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 $t < y$. Assume that $w \in \mathcal{I}(321)$ 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 zxy is an occurrence of 321 in w, a contradiction. A similar argument holds for the case $(xy)(z)$ with $z < x$. Due to the definition of the s.c.n. of w, the case $(x)(y)(z)$ with $z < x$ does 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 of zxy of 321, a contradiction. Thus, $\phi(\mathcal{I}(321)) \subset \mathcal{P}_n$. Since ϕ is injective, and $\mathcal{I}(321)$ is also enumerated by b_n (see for instance [6, 22]), we have $\phi(\mathcal{I}(321)) = \mathcal{P}_n$. □

Note that it is known that Foata's first transformation ϕ 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 $\mathcal{I}(321)$. We also have $g(\phi(w)) = \Phi(w)$ for $w \in \mathcal{I}(321)$, where Φ is a bijection in [6] between involutions and labeled Motzkin paths, which also is a restriction of Biane's bijection [8], which in turn is closely related to Françon-Viennot bijection [14].

The next Corollary deals with alternating faro permutations, i.e. permutations w satisfying $w_1 > w_2 < w_3 > \ldots$.

Theorem 4.2. *There is a bijection between alternating faro permutations of length $2n$ and the set of Dyck paths of semilength n.*

Proof. Alternating faro permutations w are those that have a block decomposition with no singleton. Due to the definition of f, this means that $f(w)$ does not contain any flat steps and thus, $g(w)$ is a Dyck path of length $2n$, and vice versa. □

Similarly, we can easily prove that length n alternating faro permutations are exactly length n faro derangements, i.e. faro permutations with no fixed point $w_i = i$.

Theorem 4.3. *Length n faro permutations avoiding the classical pattern 231 (resp. the pattern 312) are enumerated by the Fibonacci sequence f_n defined by $f_n = f_{n-1} + f_{n-2}$ with $f_1 = 1$, $f_2 = 2$.*

Proof. A faro permutation w avoiding the pattern 231 is of the form $1w'$ or $21w'$, where w' also is a faro permutation avoiding 231. 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 231, a contradiction. Therefore, the cardinality f_n of length n faro permutations satisfies $f_n = f_{n-1} + f_{n-2}$ with $f_1 = 1$, $f_2 = 2$. The argument is similar for the avoidance of 312. □

Similarly, we can easily prove that length n faro permutations avoiding the classical pattern 231 are exactly length n faro involutions.

In the following, we consider (for convenience) faro words on the n-ary alphabet $\{0, 1, \ldots, n-1\}$, and we focus on the set of *subexcedent* faro words of length n, i.e. faro words $w_1 w_2 \ldots w_n$ satisfying $w_i \leqslant i - 1$ for $1 \leqslant i \leqslant n$.

Theorem 4.4. *There is a bijection between subexcedent faro words of length n and 2143-avoiding Dumont permutations of the second kind of length $2n$.*

We will briefly recall the result given in [11] that enumerated 2143-avoiding Dumont permutations of the second kind of length $2n$. Dumont permutations of the second kind of length $2n$ are permutations π that satisfy the following conditions for $i \in [n]$:

$$\pi(2i-1) \geqslant 2i-1, \qquad \pi(2i) \leqslant 2i-1.$$

In other words, the values in the odd positions are weak excedances, whereas the values in the even positions are deficiencies. In addition, if π avoids pattern 2143 (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 π are exactly $\{1, 2, \ldots, n\}$, and the values in the odd positions of π are exactly $\{n+1, n+2, \ldots, 2n\}$. Moreover, the subsequence of values of π in the even positions avoids pattern 213 while the subsequence of values of π in the odd positions avoids pattern 132. This allows [11] to construct a bijection as in Krattenthaler [17] from the even-position subsequence of π 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 π to the same paths but ending at $(n+1, \lfloor (n+1)/2 \rfloor)$. Let $\{a_n\}_{n \geqslant 0}$ be the sequence A047749 [23], so that

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

then the number of 2143-avoiding Dumont permutations of the second kind of length $2n$ is $a_n a_{n+1}$.

Proof of Theorem 4.4. Let π be a subexcedent faro word of length n. As in [11], let π_o and π_e be the odd-position and even-position subsequences of π. Then π_o and π_e are nondecreasing subsequences such that

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

Conversely, any word π whose odd-position and even-position subsequences π_o and π_e satisfy the above properties is a subexcedent faro word of length n. Given sequences π_o and π_e as in (4.1), associate to them a pair of north-east lattice paths as follows. If π_o or π_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 π_o and $k = \lfloor \frac{n}{2} \rfloor$ for π_e, and let $a_{k+1} = 2k$ for π_o and $a_{k+1} = 2k+1$ for π_e.

Now consider a north-east lattice path from $(0,0)$ to (k, a_{k+1}) through vertices $(0, a_1), (1, a_2), \ldots, (k-1, a_k)$ 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

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

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 π_o and on or below the line $y = 2x+1$ for π_e, and each such path corresponds to a unique π_o or a unique π_e.

Moreover, notice that if n is even, then

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

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 π_o and π_e are in bijection with the pair of paths in the proof of the [11, Theorem 3.5] (see also [11, Figure 6]), which yields a bijection between the subexcedent faro words of size n and 2143-avoiding Dumont permutations of the second kind of size $2n$. □

The enumeration of subexcedent faro words may be refined by considering some natural statistics on such words. Together with the bijection of Theorem 4.4 to pairs of ternary paths (or 2-Dyck paths), a recent result [10] 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 a *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 4.5 ([10])**.**

- *On 2-Dyck paths ending at height* 0*, the tristatistic* $(\mathbf{pk}_0 - \mathbf{1}, \mathbf{pk}_1, \mathbf{dd})$ *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 π of length $2n$, define the following statistics on its odd-position and even-position subsequences π_o and π_e:

- $\mathbf{eOdis}(\pi)$, the number of distinct *positive* even letters in π_o (we exclude 0 since π_o and π always start with 0);

– **oOdis**(π), the number of distinct odd letters in π_o;

– **aOrpt**$(\pi) = \{i \in [n-1] \mid \pi(2i-1) = \pi(2i+1)\}$, the number of letter repetitions in π_o (the "a" in **aOrpt** stands for "any parity");

– **eEdis**(π), the number of distinct even letters in π_e;

– **oEdis**(π), the number of distinct odd letters in π_e.

Then we have the following result.

Theorem 4.6. *On subexcedant faro words of length n,*

– *the tristatistic* $(\mathbf{eOdis}, \mathbf{oOdis}, \mathbf{aOrpt})$ *is jointly equidistributed with any of its permutations.*

– *the bistatistics* $(\mathbf{eEdis}, \mathbf{oEdis})$ *and* $(\mathbf{oEdis}, \mathbf{eEdis})$ *are jointly equidistributed.*

Proof. For each of π_o and π_e, define k and $a_1, a_2, \ldots, a_k, a_k+1$ as in the proof of Theorem 4.4, and let

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

be the corresponding north-east path as in (4.2) (when needed, we will distinguish the paths obtained from π_o and π_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

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

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 π_o and $y = 2x+1$ for π_e, and ends on $y = 2x$ for π_o and $y = 2x+1$ for π_e. Therefore, it is easy to see that $\phi(P)$ stays in the first quadrant and ends at height 0 for π_o and at height 1 for π_e. Moreover, each distinct letter of π_o or π_e (except for 0 in π_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 π_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, $\mathbf{aOrpt}(\pi) = \mathbf{dd}(P_o)$.

Let ℓ be one of distinct letters of in π_o or π_e (for π_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

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

From this, it follows that, on π_o $(\mathbf{eOdis}, \mathbf{oOdis})(\pi) = (\mathbf{pk}_0 - 1, \mathbf{pk}_1)(\phi(P_o))$ if a_{k+1} is even, and $(\mathbf{eOdis}, \mathbf{oOdis})(\pi) = (\mathbf{pk}_1, \mathbf{pk}_0 - 1)(\phi(P_o))$ if a_{k+1} is odd. Likewise, $(\mathbf{eEdis}, \mathbf{oEdis})(\pi) = (\mathbf{pk}_0, \mathbf{pk}_1)(\phi(P_e))$ if a_{k+1} is even, and $(\mathbf{eEdis}, \mathbf{oEdis})(\pi) = (\mathbf{pk}_1, \mathbf{pk}_0)(\phi(P_e))$ 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 4.5, and thus the parity of a_{k+1} is immaterial in each case. □

From Corollary 1.12 and Equation (2.7) of [10], we can also determine the joint distribution of all the statistics we defined on subexcedent faro words. For this result, we let $n_o = \left\lfloor \frac{n+1}{2} \right\rfloor$ and $n_e = \left\lfloor \frac{n}{2} \right\rfloor$ (so $n_o + n_e = n$). We also let $\mathbf{aErpt}(\pi)$ be the number of letter repetitions in π_e, i.e. $\mathbf{aErpt}(\pi) = \{i \in [n-1] \mid \pi(2i) = \pi(2i+2)\}$.

Corollary 4.7. *The number of subexcedent faro words π of length n such that*

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

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 $r_4 + r_5 + r_6 = n_e$.

References

[1] A. Asinowski, A. Bacher, C. Banderier, and B. Gittenberger. Analytic combinatorics of lattice paths with forbidden patterns, the vectorial kernel method, and generating functions for pushdown automata. *Algorithmica*, 82(3):386–428, 2020.

[2] A. Asinowski, C. Banderier, and V. Roitner. Generating functions for lattice paths with several forbidden patterns. *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, and J. West. The Dyck pattern poset. *Discrete Mathematics*, 321:12–23, 2014.

[4] J.-L. Baril, R. Genestier, and S. Kirgizov. Pattern distributions in Dyck paths with a first return decomposition constrained by height. *Discrete Mathematics*, 342(9):111995, 2020.

[5] J.-L. Baril and V. Vajnovszki. Popularity of patterns over d-equivalence classes of words and permutations. *Theoretical Computer Science*, 814:249–258, April 2020.

[6] M. Barnabei, F. Bonetti, and M. Silimbani. Restricted involutions and Motzkin paths. *Advances in Applied Mathematics*, 47(1):102–115, 2011.

[7] J. Bertrand. Solution d'un problème. *Comptes Rendus de l'Académie des Sciences*, 105:369, 1887.

[8] P. Biane. Permutations suivant le type d'excédance et le nombre d'inversions et interprétation combinatoire d'une fraction continue de Heine. *European Journal of Combinatorics*, 14(4):277–284, July 1993.

[9] M. Bóna. Surprising symmetries in objects counted by Catalan numbers. *The Electronic Journal of Combinatorics*, 19(1), March 2012.

[10] A. Burstein. Distribution of peak heights modulo k and double descents on k-Dyck paths. (submitted), 2020.

[11] A. Burstein, S. Elizalde, and T. Mansour. Restricted Dumont permutations, Dyck paths, and noncrossing partitions. *Discrete Mathematics*, 306(22):2851–2869, 2006.

[12] E. Deutsch. Dyck path enumeration. *Discrete Mathematics*, 204(1-3):167–202, 1999.

[13] P. Diaconis, R.L. Graham, and W. M. Kantor. The mathematics of perfect shuffles. *Advances in Applied Mathematics*, 4(2):175–196, June 1983.

[14] J. Françon and G. Viennot. Permutations selon leurs pics, creux, doubles montées et double descentes, nombres d'Euler et nombres de Genocchi. *Discrete Mathematics*, 28(1):21–35, 1979.

[15] B. Hackl, C. Heuberger, and H. Prodinger. Counting Ascents in Generalized Dyck Paths. In *29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018)*, volume 110, pages 26:1–26:15, Dagstuhl, Germany, 2018.

[16] S. Kitaev. *Patterns in permutations and words*. Springer Science & Business Media, 2011.

[17] C. Krattenthaler. Permutations with restricted patterns and Dyck paths. *Advances in Applied Mathematics*, 27:510–530, 2001.

[18] M. Lothaire. *Combinatorics on words*, volume 17 of *Encyclopedia of Mathematics and Its Applications*. Addison-Wesley, 1983.

[19] T. Mansour. Statistics on Dyck paths. *J. Integer Seq*, 9(1):06–1, 2006.

[20] S.B. Morris. The basic mathematics of the faro shuffle. *Pi Mu Epsilon Journal*, 6(2):85–92, 1975.

[21] B. Salvy and P. Zimmermann. Gfun: A maple package for the manipulation of generating and holonomic functions in one variable. *ACM Trans. Math. Softw.*, 20(2):163–177, June 1994.

[22] R. Simion and F. Schmidt. Restricted permutations. *European J. Combin*, 6:383–406, 1985.

[23] N.J.A. Sloane et al. The on-line encyclopedia of integer sequences. https://oeis.org, 2020.

[24] R. Stanley. *Enumerative Combinatorics, Vol. 1.* Cambridge University Press, 1997.

[25] W.A. Whitworth. Arrangements of m things of one sort and n things of another sort, under certain conditions of priority. *Messenger of Math*, 8:105–114, 1878.