LAST SYMBOL DISTRIBUTION IN PATTERN AVOIDING CATALAN WORDS

JEAN-LUC BARIL, JAVIER F. GONZÁLEZ, AND JOSÉ L. RAMÍREZ

Abstract. We study the distribution of the last symbol statistics on the sets of Catalan words avoiding a pattern of length at most three. For each pattern p, we provide a bivariate generating function where the coefficient $\boldsymbol{c}_p(n,k)$ of $x^n y^k$ in its series expansion is the number of length n Catalan words avoiding p and ending with the symbol k. We deduce recurrence relations or close forms for $\boldsymbol{c}_p(n,k)$ and we provide asymptotic approximations for the expectation of the last symbol on all Catalan words avoiding p. We end this paper by describing a computational approach using computer algebra.

1. Introduction

A *restricted growth word* $w = w_1 w_2 \cdots w_n$ is a word over the set of non-negative integers such that $w_1 = \mathtt{0}$ and $0 \leq w_i \leq \mathtt{st}(w_1 \cdots w_{i-1}) + 1$, where $\mathtt{st}$ is an integer statistic. These words have been widely studied in the literature since they simultaneously generalize several classes of combinatorial objects as restricted growth functions [13], staircase words [15], ascent sequences [4, 8], and Catalan words [10].

In this paper, we focus on Catalan words which are obtained whenever $\mathtt{st}(w)$ returns the last symbol of w. More formally, a word $w = w_1 w_2 \cdots w_n$ is a *Catalan word* if

$$w_1 = \mathtt{0} \quad \text{and} \quad 0 \leq w_i \leq w_{i-1} + 1 \quad \text{for} \quad i = 2, \ldots, n.$$

For $n \geq 0$, let $\mathcal{C}_n$ denote the set of Catalan words of length n (the Catalan word of length 0 is the empty word λ). For example, we have

$$\mathcal{C}_4 = \{\mathtt{0000}, \mathtt{0001}, \mathtt{0010}, \mathtt{0011}, \mathtt{0012}, \mathtt{0100}, \mathtt{0101}, \\ \mathtt{0110}, \mathtt{0111}, \mathtt{0112}, \mathtt{0120}, \mathtt{0121}, \mathtt{0122}, \mathtt{0123}\}.$$

The cardinality of the set $\mathcal{C}_n$ is given by the Catalan number $c(n) = \frac{1}{n+1}\binom{2n}{n}$, see [16, Exercise 80]. Catalan words have already been studied in the context of exhaustive generation of Gray codes for growth-restricted words [10]. More recently, Baril et al. [1, 2] study the distribution of descents on restricted Catalan words avoiding a pattern or a pair of patterns of length at most three. Ramírez and Rojas [12] also study the distribution of descents for Catalan words avoiding consecutive patterns of length at most three. Additionally, in [5, 9], the authors started a study of combinatorial statistics on the polyominoes associated with words in $\mathcal{C}_n$.

Date: November 13, 2022.

2010 *Mathematics Subject Classification.* 05A15, 05A19.

Key words and phrases. Catalan word; generating function; pattern avoidance.

The goal of this work is to complete all these studies by providing enumerative results for Catalan words avoiding a pattern of length at most three with respect to the length and the last symbol. Using classical methods presented in [7, 11], we also give the exact value or an asymptotic approximation for the expectation of the last symbol for these words. Notice that this work is, in a way, a counterpart for Catalan words of the study of Bernini et al. [3] which enumerate permutations avoiding generalized patterns with respect to the last symbol.

The remaining of this paper is structured as follows. In Section 2, we give some definitions and notations about pattern avoidance and generating functions. As a preliminary result, we yield the bivariate generating function that counts the number of Catalan words (without avoidance of a pattern), with respect to the length and the last symbol, and we deduce the expectation (i.e. the average) of the last symbol on all length n Catalan words. In Section 3, we make a similar study for Catalan words avoiding a pattern of length 2. In Section 4, we focus on the avoidance of patterns p of length three, and we give asymptotic approximations (see [7, 11]) for the expectation of the last symbol. Finally, in Section 5 we show how we can obtain recurrence relations and asymptotic approximations for the coefficients of generating functions and for the expectations of the last symbol, by using *Mathematica* and *Maple* libraries.

2. Notations and preliminaries

For an integer $r \geq 2$, a *pattern* $p = p_1p_2\cdots p_r$ is a word (of length r) over the set $\{0, 1, \dots, r-1\}$ satisfying the condition: if $j > 0$ appears in p, then $j-1$ also appears in p. For instance, there are three patterns of length two $(00, 01, 10)$, and thirteen patterns of length three $(000, 001, 010, 011, 012, 021, 100, 101, 102, 110, 120, 201, 210)$. A Catalan word $w = w_1w_2\cdots w_n$ *contains* the pattern $p = p_1p_2\cdots p_r$ if there exists a subsequence $w_{i_1}w_{i_2}\cdots w_{i_r}$ of w $(i_1 < i_2 < \cdots < i_r)$ which is order-isomorphic to p. We say that w *avoids* p whenever w does not contain the pattern p. For example, the Catalan word 0123455543 avoids the pattern 001 and contains three subsequences isomorphic to the pattern 210.

For $n \geq 0$, let $\mathcal{C}_n(p)$ denote the set of Catalan words of length n avoiding the pattern p. We denote by $\boldsymbol{c}_p(n)$ the cardinality of $\mathcal{C}_n(p)$, $\mathcal{C}(p) := \bigcup_{n\geq 0}\mathcal{C}_n(p)$, and $\mathcal{C}(p)^+ := \mathcal{C}(p) \setminus \{\lambda\}$. We denote by $\mathsf{last}(w)$ the last symbol of w. Let $\mathcal{C}_{n,k}(p)$ denote the set of Catalan words $w \in \mathcal{C}_n(p)$ such that $\mathsf{last}(w) = \mathtt{k}$, and let $\boldsymbol{c}_p(n,k) := |\mathcal{C}_{n,k}(p)|$. Obviously, we have $\boldsymbol{c}_p(n) = \sum_{k=0}^{n-1}\boldsymbol{c}_p(n,k)$.

We introduce the bivariate generating function

$$\boldsymbol{H}_p(x,y) := \sum_{w\in\mathcal{C}(p)^+} x^{|w|}y^{\mathsf{last}(w)} = \sum_{n\geq 1,\ k\geq 0} \boldsymbol{c}_p(n,k)x^ny^k,$$

and we set

$$\boldsymbol{H}_p(x) := \sum_{w\in\mathcal{C}(p)^+} x^{|w|} = \boldsymbol{H}_p(x,1).$$

Notice that these generating functions do not consider the empty word. The expectation of the last symbol on all Catalan words in $\mathcal{C}_n(p)$ is given by (see [7])

$$\boldsymbol{a}_p(n) := \frac{[x^n]\partial_y \boldsymbol{H}_p(x,y)|_{y=1}}{\boldsymbol{c}_p(n)} = \frac{[x^n]\partial_y \boldsymbol{H}_p(x,y)|_{y=1}}{[x^n]\boldsymbol{H}_p(x,1)}.$$

Moreover, if we omit the subscript p in all previous notations, then this means we refer to the set of all Catalan words of length n (without avoidance of a pattern).

For a Catalan word $w = w_1 w_2 \cdots w_n$, $n \geq 0$, and an integer $a > 0$, we denote by $(w+a)$ the word $(w_1+a)(w_2+a)\cdots(w_n+a)$ obtained from w by increasing by a each of its symbols. In the case where w is the empty word λ, we set $(w+a) = \lambda$. Using this notation, any non-empty Catalan word has a *first return decomposition* as follows: $w = \mathtt{0}(w'+1)w''$ where w', w'' are two Catalan words. From this decomposition, a non-empty Catalan word w is either of the three forms (i) $w = \mathtt{0}$, or (ii) $w = \mathtt{0}(w'+1)$ with $w' \in \mathcal{C}^+$, or (iii) $w = \mathtt{0}(w'+1)w''$ with $w' \in \mathcal{C}$ and $w'' \in \mathcal{C}^+$. So, we deduce the following functional equation

$$\boldsymbol{H}(x,y) = x + xy\boldsymbol{H}(x,y) + x(\boldsymbol{H}(x)+1)\boldsymbol{H}(x,y),$$

which provides

$$\boldsymbol{H}(x,y) = \frac{1-2xy-\sqrt{1-4x}}{2(1-y+xy^2)} \quad \text{and} \quad \boldsymbol{H}(x,1) = \boldsymbol{C}(x),$$

where $\boldsymbol{C}(x) := \frac{1-2x-\sqrt{1-4x}}{2x}$ is the well known generating function for the non-empty Catalan words. Thanks to [5, 17], we obtain

$$[x^n y^k]\boldsymbol{H}(x,y) = \frac{k+1}{n}\binom{2n-2-k}{n-1}, \quad n \geq 1, k \geq 0 \text{ and}$$

$$[x^n]\partial_y \boldsymbol{H}(x,y)|_{y=1} = [x^n]\boldsymbol{C}(x)^2 = \frac{4}{n+2}\binom{2n-1}{n-2}.$$

Then, the expectation of the last symbol of a Catalan word in $\mathcal{C}_n$ is

$$\boldsymbol{a}(n) = \frac{4\binom{2n-1}{n-2}(n+1)}{(n+2)\binom{2n}{n}} = \frac{2(n-1)}{n+2},$$

and its limit is 2 when $n \to \infty$. Figure 1 shows the distribution of the last symbol over 10000 random Catalan words of length 17.

3. Avoidance of length 2 patterns

In this section, we provide the generating functions $\boldsymbol{H}_p(x,y)$ for all patterns $p \in \{00, 01, 10\}$, and we give close forms for the coefficients of $x^n y^k$ in their series expansions. So, we deduce the expectation of the last symbol on all Catalan words in $\mathcal{C}_p(n)$.

Theorem 3.1. *We have*

$$\boldsymbol{H}_{00}(x,y) = \frac{x}{1-xy}, \quad \boldsymbol{c}_{00}(n,k) = \begin{cases} 1, & \text{if } k = n-1 \\ 0, & \text{otherwise} \end{cases}, \quad \text{and} \quad \boldsymbol{a}_{00}(n) = n-1.$$

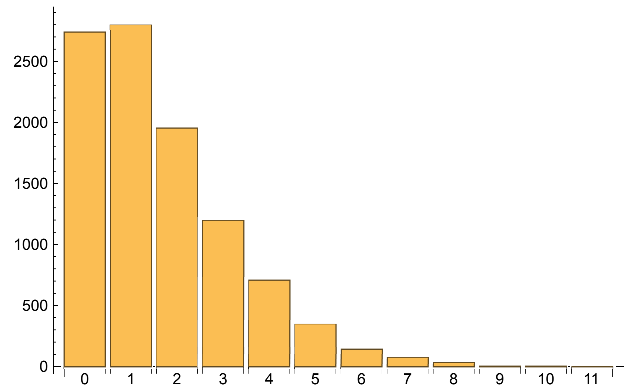

FIGURE 1. Histogram with respect to the last symbol of 10000 random Catalan words of length 17.

Proof. The word `012···n-1` is the unique non-empty word avoiding the pattern 00. □

Theorem 3.2. *We have*

$$\boldsymbol{H}_{01}(x,y)=\frac{x}{1-x},\quad \boldsymbol{c}_{01}(n,k)=\begin{cases}1, & \text{if } k=0\\ 0, & \text{otherwise}\end{cases},\quad \text{and}\quad \boldsymbol{a}_{01}(n)=0.$$

Proof. The word `00···0` is the unique non-empty word avoiding the pattern 01. □

Theorem 3.3. *We have*

$$\boldsymbol{H}_{10}(x,y)=\frac{x}{1-x(1+y)},\quad \boldsymbol{c}_{10}(n,k)=\binom{n-1}{k},\quad \text{and}\quad \boldsymbol{a}_{10}(n)=\frac{n-1}{2}.$$

Proof. Let w be a word in $\mathcal{C}_{n,k}(10)$. If $n\geq 2$, then the symbol immediately before the last one has two options $k-1$ or k. Then $\boldsymbol{c}_{10}(n,k)=\boldsymbol{c}_{10}(n-1,k-1)+\boldsymbol{c}_{10}(n-1,k)$ for $n\geq 2$ and $k\geq 1$. Additionally, $\boldsymbol{c}_{10}(n,0)=1$ for all positive integer n. Therefore, the sequence $\boldsymbol{c}_{10}(n,k)$ is equal to the binomial coefficient $\binom{n-1}{k}$ and the generating function can be deduced easily. □

4. Avoidance of length 3 patterns

In this section, we provide the generating functions $\boldsymbol{H}_p(x,y)$ for all patterns p of length three. We also yield close forms or recurrence relations for the number $\boldsymbol{c}_p(n,k)$ of Catalan words $w\in\mathcal{C}_p(n)$ such that $\mathsf{last}(w)=k$, and we deduce exact values or asymptotic approximations for the expectation of the last symbol on all Catalan words in $\mathcal{C}_p(n)$.

4.1. The pattern 012.

Theorem 4.1. *We have*

$$\boldsymbol{H}_{012}(x,y)=\frac{x(1-x+xy)}{1-2x},$$

and for $n \geq 2$,

$$\boldsymbol{c}_{012}(n,k) = \begin{cases} 2^{n-2}, & \textit{if } k = 0,1 \\ 0, & \textit{otherwise} \end{cases}, \qquad \textit{and} \qquad \boldsymbol{a}_{012}(n) = \frac{1}{2}.$$

Proof. Let w denote a non-empty Catalan word in $\mathcal{C}(012)$. From the definition of Catalan word, w must be made up exclusively of $\mathtt{0}$'s and $\mathtt{1}$'s, and the first symbol is $\mathtt{0}$. Therefore, the Catalan words avoiding the pattern 012 are given by the regular expression

$$\mathtt{0} \cup \mathtt{0}(\mathtt{0} \cup \mathtt{1})^*\mathtt{0} \cup \mathtt{0}(\mathtt{0} \cup \mathtt{1})^*\mathtt{1}.$$

This regular expression translates to the rational generating function

$$\boldsymbol{H}_{012}(x,y) = x + \frac{x^2}{1-2x} + \frac{x^2y}{1-2x} = \frac{x(1-x+xy)}{1-2x}.$$

The closed formula for the sequence $\boldsymbol{c}_{012}(n,k)$ is obtained by comparing coefficients. $\square$

The series expansion of the generating function $\boldsymbol{H}_{012}(x,y)$ is

$$x + (1+y)x^2 + (2+2y)x^3 + (\mathbf{4} + \mathbf{4}\boldsymbol{y})x^4 + (8+8y)x^5 + (16+16y)x^6 + O(x^7).$$

The Catalan words corresponding to the bold coefficients in the above series are

$$\mathcal{C}_{4,0}(012) = \{\mathtt{0000}, \mathtt{0010}, \mathtt{0100}, \mathtt{0110}\} \text{ and } \mathcal{C}_{4,1}(012) = \{\mathtt{0001}, \mathtt{0011}, \mathtt{0101}, \mathtt{0111}\}.$$

4.2. **The pattern 001.**

Theorem 4.2. *We have*

$$\boldsymbol{H}_{001}(x,y) = \frac{x(1-x)}{(1-2x)(1-xy)},$$

and for $n \geq 2$,

$$\boldsymbol{c}_{001}(n,k) = \begin{cases} 2^{n-k-2}, & \textit{if } k < n-1 \\ 1, & \textit{if } k = n-1 \\ 0, & \textit{otherwise} \end{cases}, \qquad \textit{and} \qquad \boldsymbol{a}_{001}(n) = 1 - 2^{1-n}.$$

Proof. Let w denote a non-empty Catalan word in $\mathcal{C}(001)$, and let $w = \mathtt{0}(w'+1)w''$ be the first return decomposition, where $w', w'' \in \mathcal{C}(001)$. If $w'' = \lambda$, then the generating function for this case is $x + xy\boldsymbol{H}_{001}(x,y)$. If w'' is non empty, then $w'' = \mathtt{0}^k$, $k \geq 1$, and the generating function for this case is $x\left(\boldsymbol{H}_{001}(x)+1\right)\frac{x}{1-x}$. Adding both cases we have the functional equation

$$\boldsymbol{H}_{001}(x,y) = x + xy\boldsymbol{H}_{001}(x,y) + \left(\boldsymbol{H}_{001}(x)+1\right)\frac{x^2}{1-x}.$$

Since $\boldsymbol{H}_{001}(x)+1 = \frac{1-x}{1-2x}$ (see Theorem 5 in [1]), we obtain the result. Classical methods are used for the extraction of the coefficients of $[x^ny^k]\boldsymbol{H}_{001}(x,y)$ and $[x^n]\partial_y\boldsymbol{H}_{001}(x,y)|_{y=1}$. $\square$

The series expansion of the generating function $\boldsymbol{H}_{001}(x,y)$ is

$$x+(1+y)x^2+(2+y+y^2)x^3+(\mathbf{4+2y+y^2+y^3})x^4+(8+4y+2y^2+y^3+y^4)x^5 \\ +(16+8y+4y^2+2y^3+y^4+y^5)x^6+O(x^7).$$

The Catalan words corresponding to the bold coefficients in the above series are

$$\mathcal{C}_{4,0}(001)=\{\texttt{0000},\texttt{0100},\texttt{0110},\texttt{0120}\},\quad \mathcal{C}_{4,1}(001)=\{\texttt{0111},\texttt{0121}\},$$
$$\mathcal{C}_{4,2}(001)=\{\texttt{0122}\},\ \text{ and }\ \mathcal{C}_{4,3}(001)=\{\texttt{0123}\}.$$

The array $[x^ny^k]\boldsymbol{H}_{001}(x,y)=\boldsymbol{c}_{001}(n,k)$ corresponds with the array A155038 in [14] which enumerates the set $\mathcal{P}_{n,k}$ of integer compositions of n with first part k. A constructive bijection ψ between $\mathcal{C}_{n,k}(001)$ and $\mathcal{P}_{n,k}$ can be defined recursively as follows: if $n=0$ then $\psi(\lambda)=0$; if $n\geq 1$, we distinguish two cases: if $w=\texttt{0}(w'+1)\texttt{0}^k, k\geq 1$ then we set $\psi(w)=1^k\psi(\texttt{0}(w'+1))$, and if $w=\texttt{0}(w'+1)$ then we set $\psi(w)=(c_1+1)c_2\cdots c_r$ with $\psi(w')=c_1c_2\cdots c_r$. For instance, we have $\psi(012200)=1,1,3,1$.

4.3. **The pattern 010.** Notice that a Catalan word w is in set $\mathcal{C}_{n,k}(010)$ if and only if w is in the set $\mathcal{C}_{n,k}(10)$. Therefore, from Theorem 3.3 we conclude the following result.

Theorem 4.3. *We have*

$$\boldsymbol{H}_{010}(x,y)=\frac{x}{1-x-xy},$$

and for $n\geq 1$, $0\leq k\leq n-1$, $\boldsymbol{c}_{010}(n,k)=\binom{n-1}{k}$, and $\boldsymbol{a}_{010}(n)=(n-1)/2$.

The series expansion of the generating function $\boldsymbol{H}_{010}(x,y)$ is

$$x+(1+y)x^2+(1+2y+y^2)x^3+(\mathbf{1+3y+3y^2+y^3})x^4+(1+4y+6y^2+4y^3+y^4)x^5 \\ +(1+5y+10y^2+10y^3+5y^4+y^5)x^6+O(x^7).$$

The Catalan words corresponding to the bold coefficients in the above series are

$$\mathcal{C}_{4,0}(010)=\{\texttt{0000}\},\quad \mathcal{C}_{4,1}(010)=\{\texttt{0001},\texttt{0011},\texttt{0111}\},$$
$$\mathcal{C}_{4,2}(010)=\{\texttt{0012},\texttt{0112},\texttt{0122}\},\ \text{ and }\ \mathcal{C}_{4,3}(001)=\{\texttt{0123}\}.$$

The array $[x^ny^k]\boldsymbol{H}_{010}(x,y)=\boldsymbol{c}_{010}(n,k)$ corresponds with the Pascal matrix A007318 in [14]. A constructive bijection ϕ between $\mathcal{C}_{n,k}(010)$ and the set $\mathcal{S}_{n-1,k}$ of k-subsets of $\{1,\dots,n-1\}$ can be defined as follows: Let $w\in\mathcal{C}_{n,k}(010)$, we write it $w=\texttt{0}^{i_0}\texttt{1}^{i_1}\cdots\texttt{k}^{i_k}$, with $i_j\geq 1$, $0\leq j\leq k$, and we set $\phi(w)=\{i_0,i_0+i_1,\dots,i_0+i_1+\cdots+i_{k-1}\}$. For instance, we have $\phi(\texttt{011233344})=\{1,3,4,7\}\in\mathcal{S}_{8,4}$.

4.4. **The pattern 021.**

Theorem 4.4. *We have*

$$\boldsymbol{H}_{021}(x,y)=\frac{x-5x^2+9x^3-6x^4+x^5(1-y)}{(1-2x)^2(1-x)(1-x-xy)}.$$

Moreover, for $n \geq 2$, $\boldsymbol{c}_{021}(n,0) = (n-2)\cdot 2^{n-3} + 1$, and for $k \geq 1$,

$$\boldsymbol{c}_{021}(n,k) = \sum_{i=k-1}^{n-2} \binom{n-2}{i}, \qquad \textit{and} \qquad \boldsymbol{a}_{021}(n) = \frac{2^{n-5}(n^2+3n-2)}{(n-1)2^{n-2}+1} \sim \frac{n+4}{8}.$$

Proof. Let w denote a non-empty Catalan word in $\mathcal{C}(021)$, and let $w = \mathtt{0}(w'+1)w''$ be the first return decomposition, where $w', w'' \in \mathcal{C}(021)$. If $w'' = \lambda$, then $w = \mathtt{0}(w'+1)$ where w' is a possibly empty Catalan word in $\mathcal{C}(10)$. From Theorem 3.3 the generating function for this case is

$$x + xy\boldsymbol{H}_{10}(x,y) = x + \frac{x^2y}{1-x(1+y)}.$$

If $w' = \lambda$ and $w'' \neq \lambda$, then the contribution to the generating function is $x\boldsymbol{H}_{021}(x,y)$. Finally, if w' and w'' are non empty, then we have two cases.

- Case 1: $w' = \mathtt{0}^k$ with $k \geq 1$. Since $w'' \in \mathcal{C}(021)$, the generating function is $x\,(x/(1-x))\,\boldsymbol{H}_{021}(x,y)$.
- Case 2: $w' \in \mathcal{C}(10) \setminus \mathcal{C}(01)$. In this case $w'' \in \mathcal{C}(01)$, that is, $w'' = \mathtt{0}^k$ $(k \geq 1)$. From Theorems 3.3 and 3.2 the generating function for this case is

$$x\left(\boldsymbol{H}_{01}(x) - \boldsymbol{H}_{10}(x)\right)\frac{x}{1-x} = \frac{x^4}{(1-x)^2(1-2x)}.$$

Therefore, we have the functional equation

$$\boldsymbol{H}_{021}(x,y) = x + \frac{x^2y}{1-x(1+y)} + x\boldsymbol{H}_{021}(x,y) + x\frac{x}{1-x}\boldsymbol{H}_{021}(x,y) + \frac{x^4}{(1-x)^2(1-2x)}.$$

Solving this equation we obtain the desired result. $\square$

The series expansion of the generating function $\boldsymbol{H}_{021}(x,y)$ is

$$x + (1+y)x^2 + (2+2y+y^2)x^3 + (\mathbf{5+4y+3y^2+y^3})x^4 \\ + (13+8y+7y^2+4y^3+y^4)x^5 + (33+16y+15y^2+11y^3+5y^4+y^5)x^6 + O(x^7).$$

The Catalan words corresponding to the bold coefficients in the above series are

$$\mathcal{C}_{4,0}(021) = \{\mathtt{0000}, \mathtt{0010}, \mathtt{0100}, \mathtt{0110}, \mathtt{0120}\}, \quad \mathcal{C}_{4,1}(021) = \{\mathtt{0001}, \mathtt{0011}, \mathtt{0101}, \mathtt{0111}\}$$
$$\mathcal{C}_{4,2}(021) = \{\mathtt{0012}, \mathtt{0112}, \mathtt{0122}\}, \quad \text{and} \quad \mathcal{C}_{4,3}(021) = \{\mathtt{0123}\}.$$

Notice that the array $[\boldsymbol{c}_{021}(n,k)]_{n,k\geq 1}$ coincides with the array A055248 in [14].

4.5. **The pattern 102.**

Theorem 4.5. *We have*

$$\boldsymbol{H}_{102}(x,y) = \frac{x(1-4x+4x^2-x^3(1-y))}{(1-x)(1-3x)(1-x(1+y))}.$$

Moreover, $\boldsymbol{c}_{102}(1,0)=1$, *and for* $n\geq 2$

$$\boldsymbol{c}_{102}(n,k)=\begin{cases}\sum_{k=0}^{n-2}\boldsymbol{c}_{102}(n-1,k), & \textit{if } k=0,1\\ \boldsymbol{c}_{102}(n-1,k-1)+\boldsymbol{c}_{102}(n-1,k), & \textit{if } 2\leq k\leq n-1\\ 0, & \textit{otherwise}\end{cases}, \quad \textit{and}$$

$$\boldsymbol{a}_{102}(n)=\frac{4\cdot 3^{n-1}-2^n}{3^{n-1}+1}\sim 4.$$

Proof. Let w denote a non-empty Catalan word in $\mathcal{C}(102)$, and let $w=\mathtt{0}(w'+1)w''$ be the first return decomposition, where $w',w''\in\mathcal{C}(102)$. If $w''=\lambda$, then $w=\mathtt{0}(w'+1)$ with w' possibly empty. The generating function for this case is $x+xy\boldsymbol{H}_{102}(x,y)$. If $w''\neq\lambda$, then $w=\mathtt{0}w''$ or $w=\mathtt{0}(w'+1)w''$ with $w''\in\mathcal{C}(012)^+$. Therefore, the contribution to the generating function is

$$x\boldsymbol{H}_{102}(x,y)+x\boldsymbol{H}_{102}(x)\boldsymbol{H}_{012}(x,y),$$

where (see Theorem 8 of [1])

$$\boldsymbol{H}_{102}(x)=\frac{1-3x+x^2}{(1-x)(1-3x)}-1.$$

Therefore, we have the functional equation

$$\boldsymbol{H}_{102}(x,y)=x+xy\boldsymbol{H}_{102}(x,y)+x\boldsymbol{H}_{102}(x,y)+x\boldsymbol{H}_{102}(x)\boldsymbol{H}_{012}(x,y).$$

Solving this equation we obtain the desired result. □

The series expansion of the generating function $\boldsymbol{H}_{102}(x,y)$ is

$$x+(1+y)x^2+(2+2y+y^2)x^3+(\mathbf{5+5y+3y^2+y^3})x^4$$
$$+(14+14y+8y^2+4y^3+y^4)x^5+(41+41y+22y^2+12y^3+5y^4+y^5)x^6+O(x^7).$$

The Catalan words corresponding to the bold coefficients in the above series are

$\mathcal{C}_{4,0}(102)=\{\mathtt{0000},\mathtt{0010},\mathtt{0100},\mathtt{0110},\mathtt{0120}\}$, $\mathcal{C}_{4,1}(102)=\{\mathtt{0001},\mathtt{0011},\mathtt{0101},\mathtt{0111},\mathtt{0121}\}$
$\mathcal{C}_{4,2}(102)=\{\mathtt{0012},\mathtt{0112},\mathtt{0122}\}$, and $\mathcal{C}_{4,3}(102)=\{\mathtt{0123}\}$.

The array $[\boldsymbol{c}_{102}(n,k)]_{n,k\geq 0}$ does not appear in [14].

4.6. **The pattern 201.**

Theorem 4.6. *We have*

$$\boldsymbol{H}_{201}(x,y)=\frac{x(1-5x+7x^2-2x^3)}{(1-x)(1-3x)(1+x^2y-x(2+y))},$$

$$\boldsymbol{c}_{201}(n,k)=\begin{cases}\frac{3^{n-2}+1}{2}, & \textit{if } k=0\\ \boldsymbol{c}_{201}(n-1,k-1)+\boldsymbol{c}_{201}(n-1,k) & \\ \qquad +\boldsymbol{c}_{201}(n-2,k)+\boldsymbol{c}_{201}(n-2,k+1), & \textit{if } 1\leq k\leq n-1\\ 0, & \textit{otherwise}\end{cases}, \quad \textit{and}$$

$$\boldsymbol{a}_{201}(n) = 2\frac{3^{n-1} - F(2n-2)}{3^{n-1}+1} \sim 2,$$

where $F(n)$ is the n-th term of the Fibonacci sequence defined by $F(n) = F(n-1)+F(n-2)$ anchored with $F(0) = F(1) = 1$.

Proof. Let w denote a non-empty Catalan word in $\mathcal{C}(201)$, and let $w = \mathtt{0}(w'+1)w''$ be the first return decomposition, where $w', w'' \in \mathcal{C}(201)$. If $w'' = \lambda$, then $w = \mathtt{0}(w'+1)$ with w' possibly empty. The generating function for this case is $x + xy\boldsymbol{H}_{201}(x,y)$. If $w'' \neq \lambda$ and $w' = \lambda$, then $w = \mathtt{0}w''$ with $w'' \in \mathcal{C}(201)$. The contribution to the generating function is $x\boldsymbol{H}_{201}(x,y)$. If w' and w'' are non empty, then we have two cases.

- Case 1: $w' = \mathtt{0}^n$ with $n \geq 1$. Since $w'' \in \mathcal{C}(021)^+$, the generating function is $x\frac{x}{1-x}\boldsymbol{H}_{201}(x,y)$.
- Case 2: $w' \in \mathcal{C}(201)^+ \setminus \{\mathtt{0}^k, k \geq 1\}$. In this case $w'' = \mathtt{0}^n$ for $n \geq 1$. The generating function for this case is
$$x\left(\boldsymbol{H}_{201}(x) - \frac{x}{1-x}\right)\frac{x}{1-x},$$
where (see Theorem 8 of [1])
$$\boldsymbol{H}_{201}(x) = \frac{1-3x+x^2}{(1-x)(1-3x)} - 1.$$

Therefore, we have the functional equation

$$\begin{aligned}\boldsymbol{H}_{201}(x,y) &= x + xy\boldsymbol{H}_{201}(x,y) + x\boldsymbol{H}_{201}(x,y)\\ &\quad + x\frac{x}{1-x}\boldsymbol{H}_{201}(x,y) + x\left(\boldsymbol{H}_{201}(x) - \frac{x}{1-x}\right)\frac{x}{1-x}.\end{aligned}$$

Solving this equation we obtain the desired result. □

The series expansion of the generating function $\boldsymbol{H}_{201}(x,y)$ is

$$\begin{aligned}&x + (1+y)x^2 + (2+2y+y^2)x^3 + (\boldsymbol{5+5y+3y^2+y^3})x^4\\ &\quad + (14+13y+9y^2+4y^3+y^4)x^5 + (41+35y+26y^2+14y^3+5y^4+y^5)x^6 + O(x^7).\end{aligned}$$

The Catalan words corresponding to the bold coefficients in the above series are

$$\begin{aligned}&\mathcal{C}_{4,0}(201) = \{\mathtt{0000}, \mathtt{0100}, \mathtt{0110}, \mathtt{0120}\}, \quad \mathcal{C}_{4,1}(201) = \{\mathtt{0001}, \mathtt{0011}, \mathtt{0101}, \mathtt{0111}, \mathtt{0121}\}\\ &\mathcal{C}_{4,2}(201) = \{\mathtt{0012}, \mathtt{0112}, \mathtt{0122}\}, \text{ and } \mathcal{C}_{4,3}(201) = \{\mathtt{0123}\}.\end{aligned}$$

The array $[\boldsymbol{c}_{201}(n,k)]_{n,k\geq 0}$ does not appear in [14].

4.7. The pattern 120.

Theorem 4.7. *We have*

$$\boldsymbol{H}_{120}(x,y)=\frac{(1-x)x}{1-x(2+y)+yx^2},$$

$$\boldsymbol{c}_{120}(n,k)=\begin{cases}2^{n-2}, & \text{if } k=0\\ \boldsymbol{c}_{120}(n-1,k-1)+\boldsymbol{c}_{120}(n-1,k) & \\ \qquad\qquad +\boldsymbol{c}_{120}(n-2,k)+\boldsymbol{c}_{120}(n-2,k+1), & \text{if } 1\leq k\leq n-1\\ 0, & \text{otherwise}\end{cases}, \text{ and}$$

$$\boldsymbol{a}_{120}(n)=\frac{2(n-1)F(2n-2)-(n-3)F(2n-3)}{5F(2n-2)}\sim\frac{2n}{5+\sqrt{5}},$$

where $F(n)$ is the Fibonacci sequence defined in Theorem 4.6.

Proof. Let w denote a non-empty Catalan word in $\mathcal{C}(120)$, and let $w=\mathtt{0}(w'+1)w''$ be the first return decomposition, where $w',w''\in\mathcal{C}(120)$. If $w''=\lambda$, then $w=\mathtt{0}(w'+1)$ with w' possibly empty. The generating function for this case is $x+xy\boldsymbol{H}_{120}(x,y)$. If $w''\neq\lambda$, then $w=\mathtt{0}w''$ or $w=\mathtt{0}(w'+1)w''$ with $w'=\mathtt{0}^k$, $k\geq 1$, and $w''\in\mathcal{C}(120)^+$. Therefore, the contribution to the generating function is

$$x\boldsymbol{H}_{120}(x,y)+x\left(\frac{x}{1-x}\right)\boldsymbol{H}_{012}(x,y).$$

So we deduce the functional equation

$$\boldsymbol{H}_{120}(x,y)=x+xy\boldsymbol{H}_{120}(x,y)+x\boldsymbol{H}_{120}(x,y)+x\left(\frac{x}{1-x}\right)\boldsymbol{H}_{012}(x,y).$$

Solving this equation we obtain the desired result. □

The series expansion of the generating function $\boldsymbol{H}_{120}(x,y)$ is

$$x+(1+y)x^2+(2+2y+y^2)x^3+(\mathbf{4+5\boldsymbol{y}+3\boldsymbol{y}^2+\boldsymbol{y}^3})x^4+ \\ (8+12y+9y^2+4y^3+y^4)x^5+(16+28y+25y^2+14y^3+5y^4+y^5)x^6+O(x^7).$$

The Catalan words corresponding to the bold coefficients in the above series are

$$\mathcal{C}_{4,0}(120)=\{\mathtt{0000},\mathtt{0010},\mathtt{0100},\mathtt{0110}\},\quad \mathcal{C}_{4,1}(120)=\{\mathtt{0001},\mathtt{0011},\mathtt{0101},\mathtt{0111},\mathtt{0121}\}$$
$$\mathcal{C}_{4,2}(120)=\{\mathtt{0012},\mathtt{0112},\mathtt{0122}\},\quad\text{and}\quad \mathcal{C}_{4,3}(120)=\{\mathtt{0123}\}.$$

Notice that the array $[\boldsymbol{c}_{120}(n,k)]_{n\geq 1,k\geq 0}$ coincides with the array A105306 in [14]. Therefore we have the combinatorial identity

$$\boldsymbol{c}_{120}(n,k)=\sum_{i=0}^{n-k-1}\binom{k+1}{i}\binom{n-1-i}{k}(-1)^i 2^{n-1-k-i}.$$

This array is related to the number of non-decreasing Dyck path of semi-length n with k weakly symmetric peaks [6]. We leave open the question of finding a constructive bijection.

4.8. **The pattern 101.**

Theorem 4.8. *We have*

$$\boldsymbol{H}_{101}(x,y) = \frac{x(1-3x+2x^2)}{(1-3x+x^2)(1-x(1+y))},$$

$$\boldsymbol{c}_{101}(n,k) = \begin{cases} F(2n-4), & \text{if } k=0 \\ \boldsymbol{c}_{101}(n-1,k-1) + \boldsymbol{c}_{101}(n-1,k), & \text{if } 1 \le k \le n-1\,, \text{ and} \\ 0, & \text{otherwise} \end{cases}$$

$$\boldsymbol{a}_{101}(n) = \frac{F(2n-1) - 2^{n-1}}{F(2n-2)} \sim \frac{\sqrt{5}+1}{2},$$

where $F(n)$ is the Fibonacci sequence defined in Theorem 4.6.

Proof. Let w denote a non-empty Catalan word in $\mathcal{C}(101)$, and let $w = \mathtt{0}(w'+1)w''$ be the first return decomposition, where $w', w'' \in \mathcal{C}(101)$. If $w'' = \lambda$, then $w = \mathtt{0}(w'+1)$ with w' possibly empty. The generating function for this case is $x + xy\boldsymbol{H}_{101}(x,y)$. If $w'' \neq \lambda$, then $w = \mathtt{0}w''$ or $w = \mathtt{0}(w'+1)\mathtt{0}^+$ with $w' \in \mathcal{C}(101)^+$. Therefore, the contribution to the generating function is

$$x\boldsymbol{H}_{101}(x,y) + x\boldsymbol{H}_{101}(x)\left(\frac{x}{1-x}\right),$$

where (see Theorem 9 of [1])

$$\boldsymbol{H}_{101}(x) = \frac{1-2x}{1-3x+x^2} - 1.$$

Therefore, we have the functional equation

$$\boldsymbol{H}_{101}(x,y) = x + xy\boldsymbol{H}_{101}(x,y) + x\boldsymbol{H}_{101}(x,y) + x\boldsymbol{H}_{101}(x)\left(\frac{x}{1-x}\right).$$

Solving this equation we obtain the desired result. $\square$

The series expansion of the generating function $\boldsymbol{H}_{101}(x,y)$ is

$$x + (1+y)x^2 + (2+2y+y^2)x^3 + (\mathbf{5} + \mathbf{4}\boldsymbol{y} + \mathbf{3}\boldsymbol{y}^{\mathbf{2}} + \boldsymbol{y}^{\mathbf{3}})x^4 \\ + (13+9y+7y^2+4y^3+y^4)x^5 + (34+22y+16y^2+11y^3+5y^4+y^5)x^6 + O(x^7).$$

The Catalan words corresponding to the bold coefficients in the above series are

$$\mathcal{C}_{4,0}(101) = \{\mathtt{0000}, \mathtt{0010}, \mathtt{0100}, \mathtt{0110}, \mathtt{0120}\}, \quad \mathcal{C}_{4,1}(101) = \{\mathtt{0001}, \mathtt{0011}, \mathtt{0111}, \mathtt{0121}\}$$
$$\mathcal{C}_{4,2}(101) = \{\mathtt{0012}, \mathtt{0112}, \mathtt{0122}\}, \quad \text{and} \quad \mathcal{C}_{4,3}(101) = \{\mathtt{0123}\}.$$

The array $[\boldsymbol{c}_{101}(n,k)]_{n,k\ge 0}$ does not appear in [14].

4.9. **The pattern 011.**

Theorem 4.9. *We have*

$$\boldsymbol{H}_{011}(x,y)=\frac{x(1-2x+2x^2-x^3y)}{(1-x)^3\,(1-xy)},$$

$$\boldsymbol{c}_{011}(n,k)=\begin{cases}\frac{(n-1)(n-2)}{2}, & \text{if } k=0\\ 1, & \text{if } 1\leq k\leq n-1\\ 0, & \text{otherwise}\end{cases},\quad \text{and}\quad \boldsymbol{a}_{011}(n)=1+\frac{2}{n(n-1)}\sim 1.$$

Proof. Let w denote a non-empty Catalan word in $\mathcal{C}(011)$, and let $w=\mathtt{0}(w'+1)w''$ be the first return decomposition, where $w',w''\in\mathcal{C}(011)$. If $w''=\lambda$, then $w=\mathtt{0}(w'+1)$ with $w'\in\mathcal{C}(00)$ (possibly empty). The generating function for this case is $x+xy\boldsymbol{H}_{00}(x,y)$. If $w''\neq\lambda$, then $w=\mathtt{0}w''$ or $w=\mathtt{0}(w'+1)w''$. In the first case $w''\in\mathcal{C}(011)^+$. In the second case, with $w'\in\mathcal{C}(00)$ and w'' is a word of the form $\mathtt{0}^j\mathtt{123}\cdots$, for $j\geq 1$. Therefore, we have the functional equation

$$\boldsymbol{H}_{011}(x,y)=x+xy\boldsymbol{H}_{00}(x,y)+x\boldsymbol{H}_{011}(x,y)+x\boldsymbol{H}_{00}(x)\frac{x}{1-x}.$$

Solving this equation we obtain the desired result. □

The series expansion of the generating function $\boldsymbol{H}_{011}(x,y)$ is

$$x+(1+y)x^2+(2+y+y^2)x^3+(\boldsymbol{4+y+y^2+y^3})x^4 \\ +(7+y+y^2+y^3+y^4)x^5+(11+y+y^2+y^3+y^4+y^5)x^6+O(x^7).$$

The Catalan words corresponding to the bold coefficients in the above series are

$$\mathcal{C}_{4,0}(011)=\{\mathtt{0000},\mathtt{0010},\mathtt{0100},\mathtt{0120}\},\quad \mathcal{C}_{4,1}(011)=\{\mathtt{0001}\}$$
$$\mathcal{C}_{4,2}(011)=\{\mathtt{0012}\},\quad\text{and}\quad \mathcal{C}_{4,3}(011)=\{\mathtt{0123}\}.$$

The array $[\boldsymbol{c}_{011}(n,k)]_{n,k\geq 0}$ does not appear in [14].

4.10. **The pattern 000.**

Theorem 4.10. *We have*

$$\boldsymbol{H}_{000}(x,y)=\frac{x(1-3x^2-x^3-(1-x-3x^2+x^3)x^2y)}{(1-x-3x^2+x^3)(1-xy-2x^2y+x^3y^2)},$$

$$\boldsymbol{c}_{000}(n,k)=\begin{cases}1, & \text{if } k=0,\ n=1,2,3\\ \boldsymbol{c}_{000}(n-1,k)+3\boldsymbol{c}_{000}(n-2,k)-\boldsymbol{c}_{000}(n-3,k), & \text{if } k=0,\ n\geq 4\\ \boldsymbol{c}_{000}(n-1,k-1)+2\boldsymbol{c}_{000}(n-2,k-1) & \\ \qquad -\ \boldsymbol{c}_{000}(n-3,k-2), & \text{if } k\geq 1,\ n\geq 4\\ 0, & \text{otherwise}\end{cases}\quad \text{and}$$

$$\boldsymbol{a}_{000}(n)\sim 1.$$

Proof. Let w denote a non-empty Catalan word in $\mathcal{C}(000)$, and let $w = \mathtt{0}(w'+1)w''$ be the first return decomposition, where $w', w'' \in \mathcal{C}(000)$. If $w'' = \lambda$, then $w = \mathtt{0}(w'+1)$ with w' possibly empty. The generating function for this case is $x + xy\boldsymbol{H}_{000}(x, y)$. The second case is when $w'' \neq \lambda$, then we have the following four cases:

- $w = \mathtt{00}$. The generating functions for this case is x^2.
- $w = \mathtt{00}(w'''+1)$, where $w''' \in \mathcal{C}(000)^+$. In this case the generating function is $x^2y\boldsymbol{H}_{000}(x, y)$.
- $w = \mathtt{0}(w'+1)\mathtt{0}$, where $w' \in \mathcal{C}(000)^+$. The generating function for this case is $x^2\boldsymbol{H}_{000}(x)$, where (see Theorem 11 of [1])
$$\boldsymbol{H}_{000}(x) = \frac{(1-2x^2)}{1-x-3x^2+x^3} - 1.$$
- $w = \mathtt{0}(w'+1)\mathtt{0}(w'''+1)$, where w', w'' are words in $\mathcal{C}(000)^+$ and the concatenation $w'w''$ is again a Catalan word of the second case, that is, the generating function of the words $w'w'''$ is the same of the generating function of the words of the form $\mathtt{0}(w'+1)w''$, where w'' is non-empty. Therefore, the contribution to the generating function is $T(x, y) = x^2y(x^2 + x^2y\boldsymbol{H}_{000}(x, y) + T(x, y))$.

Therefore, we have the functional equation
$$\boldsymbol{H}_{000}(x, y) = x + xy\boldsymbol{H}_{000}(x, y) + x^2 + x^2y\boldsymbol{H}_{000}(x, y) + x^2\boldsymbol{H}_{000}(x) + T(x, y).$$
Solving the system of equations we obtain the desired result. □

The series expansion of the generating function $\boldsymbol{H}_{000}(x, y)$ is
$$x + (1+y)x^2 + (1+2y+y^2)x^3 + (\boldsymbol{2+3y+3y^2+y^3})x^4 + (4+4y+6y^2+4y^3+y^4)x^5 + (9+8y+9y^2+10y^3+5y^4+y^5)x^6 + O(x^7).$$
The Catalan words corresponding to the bold coefficients in the above series are
$$\mathcal{C}_{4,0}(000) = \{\mathtt{0110}, \mathtt{0120}\}, \quad \mathcal{C}_{4,1}(000) = \{\mathtt{0011}, \mathtt{0101}, \mathtt{0121}\}$$
$$\mathcal{C}_{4,2}(000) = \{\mathtt{0012}, \mathtt{0112}, \mathtt{0122}\}, \quad \text{and} \quad \mathcal{C}_{4,3}(000) = \{\mathtt{0123}\}.$$
The array $[\boldsymbol{c}_{000}(n, k)]_{n,k\geq 0}$ does not appear in [14].

4.11. **The pattern 100.**

Theorem 4.11. *We have*
$$\boldsymbol{H}_{100}(x, y) = \frac{x(1-3x+x^2(1-y)-x^4(1-2y)+x^3(1+2y))}{(1-2x-2x^2)(1+x^2+x^3y^2-x(2+y))},$$
$$\boldsymbol{c}_{100}(n, k) = \begin{cases} 1, & \text{if } k = 0,\ n = 1, 2 \\ 2, & \text{if } k = 0,\ n = 3 \\ 4, & \text{if } k = 0,\ n = 4 \\ 4\boldsymbol{c}_{100}(n-1, k) - 3\boldsymbol{c}_{100}(n-2, k) - 2\boldsymbol{c}_{100}(n-3, k) + 2\boldsymbol{c}_{100}(n-4, k), & \text{if } k = 0,\ n \geq 5 \\ 2\boldsymbol{c}_{100}(n-1, k) + \boldsymbol{c}_{100}(n-1, k-1) - \boldsymbol{c}_{100}(n-2, k) - \boldsymbol{c}_{100}(n-3, k-2), & \text{if } k \geq 1,\ n \geq 4 \\ 0, & \text{otherwise} \end{cases} \quad \textit{and}$$

$$\boldsymbol{a}_{100}(n) \sim \frac{8\sqrt{3}-12}{9\sqrt{3}-15} \sim 3.154700527.$$

Proof. Let w denote a non-empty Catalan word in $\mathcal{C}(100)$, and let $w = \mathtt{0}(w'+1)w''$ be the first return decomposition, where $w', w'' \in \mathcal{C}(100)$. If $w'' = \lambda$, then $w = \mathtt{0}(w'+1)$ with w' possibly empty. The generating function for this case is $x + xy\boldsymbol{H}_{100}(x,y)$. Now let us assume that $w'' \neq \lambda$. Therefore, w contains at least two symbols $\mathtt{0}$. We distinguish two cases: (i) w contains $k \geq 3$ symbols $\mathtt{0}$, and (ii) w contains exactly two $\mathtt{0}$.

- Case (i): w is necessarily of the form $w = \mathtt{0}^{k-2}u$ where $u \in \mathcal{C}(100)$ contains exactly two symbols $\mathtt{0}$. Thus, the generating function for this case is

$$A_2(x,y) \cdot \sum_{k\geq 3} x^{k-2},$$

  where $A_2(x,y)$ is the generating function for Catalan words in $\mathcal{C}(100)$ having exactly two symbols $\mathtt{0}$.
- Case (ii): w can be written $w = \mathtt{0}(w'+1)\mathtt{0}(w''+1)$ where $w', w'' \in \mathcal{C}(100)$.
  - If $w' = \lambda$ then the generating function is $x^2y\boldsymbol{H}_{100}(x,y)$;
  - If $w'' = \lambda$ then the generating function is $x^2\boldsymbol{H}_{100}(x)$ where (see [1])

$$\boldsymbol{H}_{100}(x) = \frac{1-2x-x^2+x^3}{1-3x+2x^3} - 1$$

  - Otherwise, w' and w'' are non-empty, where the Catalan word $w'w''$ avoids 100 and contains at least two symbols $\mathtt{0}$, that is $w'w'' = \mathtt{0}^{k-2}\mathtt{0}(u+1)\mathtt{0}(v+1)$. So, there are $k-1$ choices possibles for w', namely, $\mathtt{0}, \ldots, \mathtt{0}^{k-2}$ and $\mathtt{0}^{k-2}\mathtt{0}(u+1)$. So, the generating function for these words is $x^2yA_2(x,y)\sum_{k\geq 2}(k-1)x^{k-2}$. So, we obtain

$$A_2(x,y) = x^2y\boldsymbol{H}_{100}(x,y) + x^2\boldsymbol{H}_{100}(x) + x^2yA_2(x,y)\sum_{k\geq 2}(k-1)x^{k-2}.$$

Combining all these cases, we obtain the functional equations:

$$\begin{cases} \boldsymbol{H}_{100}(x,y) & = x + xy\boldsymbol{H}_{100}(x,y) + A_2(x,y)\sum_{k\geq 3} x^{k-2} + A_2(x,y) \\ A_2(x,y) & = x^2y\boldsymbol{H}_{100}(x,y) + x^2\boldsymbol{H}_{100}(x) + x^2yA_2(x,y)\sum_{k\geq 2}(k-1)x^{k-2}. \end{cases}$$

A simple calculation provides the result. □

The series expansion of the generating function $\boldsymbol{H}_{100}(x,y)$ is

$$\begin{aligned} &x + (1+y)x^2 + (2+2y+y^2)x^3 + (\boldsymbol{4+5y+3y^2+y^3})x^4 \\ &\quad + (9+12y+9y^2+4y^3+y^4)x^5 + (22+28y+25y^2+14y^3+5y^4+y^5)x^6 + O(x^7). \end{aligned}$$

The Catalan words corresponding to the bold coefficients in the above series are

$$\mathcal{C}_{4,0}(100) = \{\mathtt{0000}, \mathtt{0010}, \mathtt{0110}, \mathtt{0100}\}, \quad \mathcal{C}_{4,1}(100) = \{\mathtt{0001}, \mathtt{0011}, \mathtt{0101}, \mathtt{0111}, \mathtt{0121}\}$$
$$\mathcal{C}_{4,2}(100) = \{\mathtt{0012}, \mathtt{0112}, \mathtt{0122}\}, \quad \text{and} \quad \mathcal{C}_{4,3}(100) = \{\mathtt{0123}\}.$$

4.12. The pattern 110.

Theorem 4.12. *We have*

$$\boldsymbol{H}_{110}(x,y) = \frac{x(1-2x+2x^2-x^2y+x^3y)}{(1-x)(1-2x+x^2-xy+x^3y^2)}.$$

$$\boldsymbol{a}_{110}(n) \sim \frac{4\sqrt{2}-5}{4\sqrt{2}-4} \cdot n \sim 0.3964466089 \cdot n.$$

Proof. Let w denote a non-empty Catalan word in $\mathcal{C}(110)$, and let $w = \mathtt{0}(w'+1)w''$ be the first return decomposition, where $w', w'' \in \mathcal{C}(110)$. If $w'' = \lambda$, then $w = \mathtt{0}(w'+1)$ with w' possibly empty. The generating function for this case is $x + xy\boldsymbol{H}_{110}(x,y)$. If $w'' \neq \lambda$ and $w' = \lambda$, then $w = \mathtt{0}w''$. The contribution is $x\boldsymbol{H}_{110}(x,y)$. If $w = 0(w'+1)w''$ with w' and w'' non-empty, then $w' = \mathtt{01}\cdots\mathtt{k}$ with $k \geq 0$, and w'' is of the form either $w'' = \mathtt{0}^{i_1}\mathtt{1}^{i_2}\cdots\ell^{i_\ell}$ with $0 \leq \ell$, $i_j \geq 1$, that is $w'' \in \mathcal{C}(10)^+$, or $w'' = \mathtt{0}^{i_1}\mathtt{1}^{i_2}\cdots\mathtt{k}^{i_k}(w'''+1)$ with $i_j \geq 1$ and $w''' \in \mathcal{C}(110) \setminus \mathcal{C}(10)$. The contribution for these two cases are respectively $\frac{x}{1-x}\boldsymbol{H}_{10}(x,y)$ and

$$x\left(\sum_{k\geq 1} x^k \cdot \left(\frac{x}{1-x}\right)^k y^k\right)(\boldsymbol{H}_{110}(x,y) - \boldsymbol{H}_{10}(x,y)) = \frac{x^3y}{1-x-x^2y}(\boldsymbol{H}_{110}(x,y) - \boldsymbol{H}_{10}(x,y)).$$

Therefore we have the functional equation

$$\begin{aligned}\boldsymbol{H}_{110}(x,y) = x + xy\boldsymbol{H}_{110}(x,y) + x\boldsymbol{H}_{110}(x,y) + \frac{x}{1-x}\boldsymbol{H}_{10}(x,y)& \\ + \frac{x^3y}{1-x-x^2y}(\boldsymbol{H}_{110}(x,y) - \boldsymbol{H}_{10}(x,y)).&\end{aligned}$$

Solving the system of equations we obtain the desired result. □

The series expansion of the generating function $\boldsymbol{H}_{110}(x,y)$ is

$$\begin{aligned}x + (1+y)x^2 + (2+2y+y^2)x^3 + (\boldsymbol{4+5y+3y^2+y^3})x^4& \\ + (7+12y+9y^2+4y^3+y^4)x^5 + (11+26y+25y^2+14y^3+5y^4+y^5)x^6 + O(x^7).&\end{aligned}$$

The Catalan words corresponding to the bold coefficients in the above series are

$$\begin{aligned}&\mathcal{C}_{4,0}(110) = \{\mathtt{0000}, \mathtt{0010}, \mathtt{0100}, \mathtt{0120}\}, \quad \mathcal{C}_{4,1}(110) = \{\mathtt{0001}, \mathtt{0011}, \mathtt{0101}, \mathtt{0111}, \mathtt{0121}\} \\ &\mathcal{C}_{4,2}(110) = \{\mathtt{0012}, \mathtt{0112}, \mathtt{0122}\}, \quad \text{and} \quad \mathcal{C}_{4,3}(110) = \{\mathtt{0123}\}.\end{aligned}$$

The array $[\boldsymbol{c}_{110}(n,k)]_{n,k\geq 0}$ does not appear in [14]. A recurrence relation for the coefficients $\boldsymbol{c}_{110}(n,k)$ can be obtained using *Mathematica* package (see Section 5), but the relation is very ugly and too involving. So we do not provide it here.

4.13. **The pattern 210.**

Theorem 4.13. *We have*

$$\boldsymbol{H}_{210}(x,y)=\frac{x(1-4x-x^3(1-2y)+x^2(5-y))}{(1-2x)(1+x^3y^2+x^2(2+y)-x(3+y))}.$$

$$\boldsymbol{c}_{210}(n,k)=\begin{cases}1, & \text{if } k=0,\ n=1,2\\ 2, & \text{if } k=0,\ n=3\\ 5\boldsymbol{c}_{210}(n-1,k)-8\boldsymbol{c}_{210}(n-2,k)+4\boldsymbol{c}_{210}(n-3,k), & \text{if } k=0,\ n\geq 4\\ 3\boldsymbol{c}_{210}(n-1,k)+\boldsymbol{c}_{210}(n-1,k-1)-2\boldsymbol{c}_{210}(n-2,k)- & \\ \qquad -\boldsymbol{c}_{210}(n-2,k-1)-\boldsymbol{c}_{210}(n-3,k-2), & \text{if } k\geq 1,\ n\geq 4\\ 0, & \text{otherwise}\end{cases} \quad \text{and}$$

$$\boldsymbol{a}_{210}(n)\sim\frac{(a+1)(-1+2a)}{3a^2+6a-4}\cdot n\sim 0.2630237717\cdot n,$$

where

$$a=-1-\frac{\sqrt{7}\,\sqrt{3}\,\sin\left(\frac{\arctan\left(\frac{\sqrt{3}}{9}\right)}{3}+\frac{\pi}{6}\right)}{3}+\cos\left(\frac{\arctan\left(\frac{\sqrt{3}}{9}\right)}{3}+\frac{\pi}{6}\right)\sqrt{7}.$$

Proof. Let w denote a non-empty Catalan word in $\mathcal{C}(210)$, and let $w=\mathtt{0}(w'+1)w''$ be the first return decomposition, where $w',w''\in\mathcal{C}(210)$. Then, w has one of the following forms:

- $w=\mathtt{0}(w'+1)$ where $w'\in\mathcal{C}(210)$; the generating function for these words is $x+xy\boldsymbol{H}_{210}(x,y)$.
- $w=\mathtt{0}w''$ where $w''\in\mathcal{C}(210)^+$; the generating function for these words is $x\boldsymbol{H}_{210}(x,y)$.
- $w=\mathtt{0}(w'+1)w''$ where $w'\in\mathcal{C}(01)^+$ and $w''\in\mathcal{C}(210)^+$; the generating function for these words is $x\boldsymbol{H}_{01}(x)\boldsymbol{H}_{210}(x,y)$.
- $w=\mathtt{0}\mathtt{1}^{a_1}\mathtt{2}^{a_2}\cdots\mathtt{k}^{a_k}w''$ where $k\geq 2$, $a_i\geq 1$ for $1\leq i\leq k$, and $w''\in\mathcal{C}(10)^+$; the generating function for these words is $\boldsymbol{H}_{10}(x,y)\sum_{k\geq 2}\frac{x^{k+1}}{(1-x)^k}$.
- $w=\mathtt{0}\mathtt{1}^{a_1}\mathtt{2}^{a_2}\cdots\mathtt{k}^{a_k}\mathtt{0}^{b_0}\mathtt{1}^{b_1}\mathtt{2}^{b_2}\cdots(\mathtt{k-2})^{b_{k-2}}(w''+\mathtt{k}-\mathtt{1})$ where $k\geq 2$, $a_i\geq 1$ for $1\leq i\leq k$, $b_i\geq 1$ for $0\leq i\leq k-2$, and $w''\in\mathcal{C}(210)\backslash\mathcal{C}(10)$; the generating function for these words is
$$(\boldsymbol{H}_{210}(x,y)-\boldsymbol{H}_{10}(x,y))\sum_{k\geq 2}\frac{x^{k+1}}{(1-x)^k}\frac{x^{k-1}}{(1-x)^{k-1}}y^{k-1}.$$

Combining these different cases, we deduce the functional equation:

$$\begin{aligned}\boldsymbol{H}_{210}(x,y)&=x+xy\boldsymbol{H}_{210}(x,y)+x\boldsymbol{H}_{210}(x,y)+x\boldsymbol{H}_{01}(x)\boldsymbol{H}_{210}(x,y)\\ &\quad+\boldsymbol{H}_{10}(x,y)\sum_{k\geq 2}\frac{x^{k+1}}{(1-x)^k}+(\boldsymbol{H}_{210}(x,y)-\boldsymbol{H}_{10}(x,y))\sum_{k\geq 2}\frac{x^{2k}}{(1-x)^{2k-1}}y^{k-1},\end{aligned}$$

which gives the results. $\square$

The series expansion of the generating function $\boldsymbol{H}_{210}(x,y)$ is

$$x+(1+y)x^2+(2+2y+y^2)x^3+(\boldsymbol{5+5y+3y^2+y^3})x^4+ (13+14y+9y^2+4y^3+y^4)x^5+(33+40y+28y^2+14y^3+5y^4+y^5)x^6+O(x^7).$$

The Catalan words corresponding to the bold coefficients in the above series are

$\mathcal{C}_{4,0}(210)=\{\texttt{0000},\texttt{0010},\texttt{0100},\texttt{0110},\texttt{0120}\}$, $\mathcal{C}_{4,1}(210)=\{\texttt{0001},\texttt{0011},\texttt{0101},\texttt{0111},\texttt{0121}\}$
$\mathcal{C}_{4,2}(210)=\{\texttt{0012},\texttt{0112},\texttt{0122}\}$, and $\mathcal{C}_{4,3}(210)=\{\texttt{0123}\}$.

5. An experimental approach by using computer algebra

In the previous results we give several recurrence relations and asymptotic expansions for the sequences $\boldsymbol{c}_p(n,k)$ and $\boldsymbol{a}_p(n)$, respectively. To find some of these recurrences, we used the package `Guess` (written by Manuel Kauers). It is a *Mathematica* package for guessing multivariate recurrence equations and it is part of the `RISCErgoSum` bundle, developed by the Research Institute for Symbolic Computation (RISC).

For example, consider the sequence $\boldsymbol{c}_{000}(n,k)$. In Theorem 4.10 we prove that the generating function $\boldsymbol{H}_{000}(x,y)$ is rational, then the term $\boldsymbol{c}_{000}(n,k)$ can be expressed as linear combinations of the previous terms. Once `Guess` installed in a *Mathematica* session we find the recurrence relation given in Theorem 4.10 by executing the following commands.

```
In[1]:= data = CoefficientList[Normal[CoefficientList[Series[
        (x (1 - 3 x^2 - x^3 - (x^2 - x^3 - 3 x^4 + x^5) y)) /
        ((1 - x - 3 x^2 + x^3) (1 - x y - 2 x^2 y + x^3 y^2)),
        {x, 0, 20}],x]], y] // PadRight;
In[2]:= sset = Flatten[Table[f[n - u, k - v], {u, 0, 3}, {v, 0, 3}]]
In[3]:= GuessMultRE[Delete[data, 1], sset, {n, k}, 1] // Simplify
Out[3]= {n (f[-3+n,-2+k] - 2 f[-2+n,-1+k] - f[-1+n,-1+k] + f[n,k]),
         k (f[-3+n,-2+k] - 2 f[-2+n,-1+k] - f[-1+n,-1+k] + f[n,k]),
         f[-3+n,-2+k] - 2f[-2+n,-1+k] - f[-1+n, -1+k] + f[n,k],
         n (f[-3+n,-3+k] - 2 f[-2+n,-2+k] - f[-1+n,-2+k] + f[n,-1+k]),
         k (f[-3+n,-3+k] - 2 f[-2+n,-2+k] - f[-1+n,-2+k] + f[n,-1+k]),
         f[-3+n, -3+k] - 2f[-2+n,-2+k] - f[-1+n,-2+k] + f[n,-1+k]}
```

From this result we guess that for $n\geq 4$ and $k\geq 1$

$$\boldsymbol{c}_{000}(n,k)=\boldsymbol{c}_{000}(n-1,k-1)+2\boldsymbol{c}_{000}(n-2,k-1)-\boldsymbol{c}_{000}(n-3,k-2).$$

From the generating function it is not difficult to prove the above equation. In order to obtain asymptotic approximations for $\boldsymbol{a}_p(n)$, we use two methods: if we can extract the coefficients of $[x^n]\partial_y\boldsymbol{H}_p(x,y)|_{y=1}$ and $[x^n]\boldsymbol{H}_p(x,1)(=\boldsymbol{c}_p(n))$, we do it and we calculate the limit of the quotient; otherwise, we use classical methods from generating functions (see [7, 11]). Notice such a routine is implemented in a Maple function `equivalent` of the `Algolib` and `gfun` libraries (see http://algo.inria.fr/libraries/).

References

[1] J.-L. Baril, S. Kirgizov, and V. Vajnovszki. Descent distribution on Catalan words avoiding a pattern of length at most three. Discrete Math. **341** (2018), 2608–2615.

[2] J.-L. Baril, C. Khalil, and V. Vajnovszki. Catalan words avoiding pairs of length three patterns. Discret. Math. Theor. Comput. Sci. **22** (2) (2021), # 5.

[3] A. Bernini, M. Bouvel, and L. Ferrari. Some statistics on permutations avoiding generalized patterns, Pure Math Appl. (PU.M.A.) **18**, (2007), 223–237.

[4] M. Bousquet-Mélou, A. Claesson, M. Dukes, and S. Kitaev. (2+2)-free posets, ascent sequences and pattern avoiding permutations, J. Comb. Theory Ser. A, **117**(7) (2010), 884–909.

[5] D. Callan, T. Mansour, and J. L. Ramírez. Statistics on bargraphs of Catalan words, J. Autom. Lang. Comb. **26** (2021), 177–196.

[6] S. Elizalde, R. Flólrez, and J. L. Ramírez, Enumerating symmetric peaks in non-decreasing Dyck paths, Ars Math. Contemp. **21** (2021) #P2.04.

[7] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2009.

[8] T. Mansour and M. Shattuck. Some enumerative results related to ascent sequences, Discrete Math. **315** (2014), 29–41.

[9] T. Mansour, J. L. Ramírez, D. A. Toquica. Counting lattice points on bargraphs of Catalan words. Math. Comput. Sci. **15** (2021), 701–713.

[10] T. Mansour and V. Vajnovszki. Efficient generation of restricted growth words. Inform. Process. Lett. **113** (2013), 613–616.

[11] A. G. Orlov. On asymptotic behavior of the Taylor coefficients of algebraic functions, Sib. Math. J. **25**(5) (1994), 1002–1013.

[12] J. L. Ramírez and A. Rojas-Osorio. Consecutive patterns in Catalan words and the descent distribution. Submitted (2022).

[13] F. Ruskey and C. Savage. Gray codes for set partitions and restricted growth tails, Australas. J. Combin. **10** (1994), 85–96.

[14] N. J. A. Sloane, The On-Line Encyclopedia of Integer Sequences. Available at https://oeis.org/.

[15] R. Stanley. *Enumerative Combinatorics*, Vol. 1 Second Ed. Cambridge, 2012.

[16] R. Stanley. *Catalan Numbers*. Cambridge University Press, Cambridge, 2015.

[17] W.-J. Woan, L. Shapiro, and D.G. Rogers. The Catalan numbers, the Lebesgue integral, and 4^{n-2}, Amer. Math. Monthly, **104** (1997), 926–931.

LIB, Université de Bourgogne, B.P. 47 870, 21078 DIJON-Cedex, France
Email address: `barjl@u-bourgogne.fr`

Departamento de Matemáticas, Universidad Nacional de Colombia, Bogotá, Colombia
Email address: `javgonzalezgu@unal.edu.co`

Departamento de Matemáticas, Universidad Nacional de Colombia, Bogotá, Colombia
Email address: `jlramirezr@unal.edu.co`