Filomat 31:3 (2017), 543–558
DOI 10.2298/FIL1703543M

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Avoidance of Classical Patterns by Catalan Sequences

Toufik Mansour[a], Mark Shattuck[b]

[a]*Department of Mathematics, University of Haifa, 3498838 Haifa, Israel*
[b]*Department of Mathematics, University of Tennessee, Knoxville, TN 37996*

Abstract. A certain subset of the words of length n over the alphabet of non-negative integers satisfying two restrictions has recently been shown to be enumerated by the Catalan number C_{n-1}. Members of this subset, which we will denote by $W(n)$, have been termed *Catalan words* or *sequences* and are closely associated with the 321-avoiding permutations. Here, we consider the problem of enumerating the members of $W(n)$ satisfying various restrictions concerning the containment of certain prescribed subsequences or patterns. Among our results, we show that the generating function counting the members of $W(n)$ that avoid certain patterns is always rational for four general classes of patterns. Our proofs also provide a general method of computing the generating function for all the patterns in each of the four classes. Closed form expressions in the case of three-letter patterns follow from our general results in several cases. The remaining cases for patterns of length three, which we consider in the final section, may be done by various algebraic and combinatorial methods.

1. Introduction

A *Catalan sequence* $w = w_1w_2\cdots w_n$ is a word in the alphabet of non-negative integers satisfying the following two properties:

(i) $w_{i+1} \geq w_i - 1$ for $1 \leq i < n$, and

(ii) if $w_i = k > 0$ with i minimal, then there exist $i_1 < i < i_2$ such that $w_{i_1} = w_{i_2} = k - 1$.

Let $W(n)$ denote the set of Catalan sequences of length n. For example, there are five members of $W(4)$, namely, 0000, 0100, 0010, 0110 and 0101. The first property states that within members of $W(n)$, there are no descents of size greater than one (with no restriction on ascent sizes), while the second property states that the left-most occurrence of each letter $k > 0$ has $k - 1$ somewhere to its left and somewhere to its right. The set was first considered by Albert et al. [1] in conjunction with 321-avoiding permutations and the question was raised in [12] of finding a one-to-one correspondence between $W(n)$ and any family of objects enumerated by the Catalan numbers. Stump [11] was successful in finding such a correspondence between $W(n)$ and the set of Dyck paths of length $2n-2$. See also [6] for an algebraic generalization of Stump's result as well as the related paper [5].

2010 *Mathematics Subject Classification*. Primary 05A15; Secondary 05A05
Keywords. pattern avoidance, Catalan sequence, generating function, Fibonacci numbers
Received: 06 December 2014; Accepted: 08 July 2015
Communicated by Francesco Belardo
Email addresses: tmansour@univ.haifa.ac.il (Toufik Mansour), shattuck@math.utk.edu (Mark Shattuck)

If $k \geq 1$, then let $[k] = \{1, 2, \ldots, k\}$, with $[0] = \emptyset$. Given $w \in [k]^n$, we define the *reduction* of w, denoted $\mathrm{red}(w)$, to be the word obtained by replacing all occurrences of the i-th smallest letter of w with i for each i. For example, $\mathrm{red}(6841846) = 3421423$. The words v and w are *order-isomorphic* if $\mathrm{red}(v) = \mathrm{red}(w)$, and we will denote this by $v \sim w$. A *pattern* will refer to a word that contains every letter in $[j]$ for some $j \geq 1$. A word $\pi = \pi_1\pi_2 \cdots \pi_n$ *contains* $\sigma = \sigma_1\sigma_2 \cdots \sigma_\ell$ as a *classical* pattern if there exists a subsequence $\pi_{i_1}\pi_{i_2} \cdots \pi_{i_\ell}$ for $1 \leq i_1 < i_2 < \cdots < i_\ell \leq n$ such that $\pi_{i_1}\pi_{i_2} \cdots \pi_{i_\ell} \sim \sigma$. Otherwise, π is said to *avoid* σ. For example, the word $w = 521352324565 \in [6]^{12}$ contains three occurrences of the pattern 1212, as witnessed by the subsequences 2323 and 3535 (twice), but avoids the pattern 4321. Note that equal letters within a pattern must be equal within an occurrence of the pattern.

In this paper, we consider the problem of pattern avoidance on Catalan sequences and enumerate members of $W(n)$ avoiding a single classical pattern in several cases. In prior studies of pattern avoidance, analogous questions have been considered on such structures as permutations, compositions, and set partitions. We refer the reader to the texts of Kitaev [4] and Heubach and Mansour [3] and the references contained therein. Given $n \geq 1$ and a pattern τ, let $W_\tau(n)$ denote the subset of $W(n)$ whose members avoid τ and let $a_\tau(n) = |W_\tau(n)|$. In the case when τ is of one of the following four general forms, namely, $11\cdots1$, $12\cdots k$, $1\cdots12$ or $2\cdots21$, it is shown that the generating function $\sum_{n\geq1} a_\tau(n)x^n$ is always rational. We note that general rationality results of this kind have been found before for particular classes of permutations; see, for example, [7, Theorem 3.1] and [8, Theorem 2.1].

In the case $12\cdots k$, we solve the functional equation satisfied by the generating function by expressing its solution in terms of Chebyshev polynomials, from which one can deduce its rationality for all k. For the other three cases, we first find systems of linear equations satisfied by certain generating functions which refine $\sum_{n\geq1} a_\tau(n)x^n$ (obtained by considering the statistic on $W(n)$ which records the number of zeros). By applying Cramer's rule to these systems, not only is the rationality of $\sum_{n\geq1} a_\tau(n)x^n$ apparent, but one also obtains general expressions involving determinants for these generating functions which can be computed explicitly (with the aid of programming) for a given pattern within one of the classes.

The organization of the paper is as follows. In the next section, we present several general rationality results and obtain as corollaries explicit expressions for $a_\tau(n)$ in the case when τ has length three. In the third section, we determine $a_\tau(n)$ for the remaining cases when τ has length three. In the particular cases 321 and 211, we provide combinatorial proofs of our results, while for 312 and 213, we make use of generating function methods. Our results show that the generating function $\sum_{n\geq1} a_\tau(n)x^n$ is rational for all patterns τ of length three.

We will make use of the following notation throughout. Given positive integers m and n, let $[m, n] = \{m, m+1, \ldots, n\}$ if $m \leq n$, with $[m, n] = \emptyset$ if $m > n$. If P is a statement, then $\chi(P)$ equals one or zero depending on the truth or falsity of P. If $n \geq 1$ and $1 \leq m \leq n$, then let $W_\tau(n, m)$ denote the subset of $W_\tau(n)$ whose members contain m distinct letters, with $a_\tau(n, m) = |W_\tau(n, m)|$. Let

$$a_\tau(n; y) = \sum_{\pi \in W_\tau(n)} y^{\nu(\pi)}, \qquad n \geq 1,$$

where $\nu(\pi)$ records the number of distinct letters appearing in π. Finally, if $n \geq 1$, then let F_n denote the Fibonacci number defined by $F_n = F_{n-1} + F_{n-2}$ if $n \geq 3$, with $F_1 = F_2 = 1$ (see [10, A000045]).

Table 1 below gives explicit formulas for $a_\tau(n)$ in all cases when τ has length three.

τ	$a_\tau(n)$	Reference
111	$a_{111}(n) = 2^{\lfloor \frac{n-2}{2} \rfloor}$	Corollary 2.3
112	F_n	Corollary 2.8
121	1	Observation 3.1
122	$n-1$	Observation 3.1
123, 132	$2^{n-1} - n + 1$	Corollary 2.5
211	$\sum_{i=1}^{3} \frac{c_i^2 - c_i - 1}{2c_i^2 + 2c_i - 3} c_i^n$, where c_1, c_2, c_3 are the roots of $x^3 - 2x^2 - x + 1 = 0$	Theorem 3.5
212, 221	2^{n-2}	Corollary 2.11
213	$\frac{(1+\sqrt{2})^n - (1-\sqrt{2})^n}{4\sqrt{2}} - \frac{1}{2}n + 1$	Theorem 3.8
231, 321	F_{2n-3}	Theorem 3.3
312	$\frac{3^{n-2}+1}{2}$	Theorem 3.7

Table 1: $a_\tau(n)$ for $n \geq 2$ when τ has length three.

2. General Results

In this section, we consider four general families of patterns and determine formulas for the generating function that counts members of the corresponding avoidance classes. In each of the four cases, this generating function works out to a rational function of x. Explicit formulas for the cardinality of an avoidance class are determined in several particular cases.

2.1. *The Pattern* $11\cdots1$

In this subsection, we consider the case of avoiding the pattern $1\cdots1$ of length $k \geq 2$. By definition, note that $a_{1\cdots1}(n)$ counts the Catalan words of length n in which no letter appears k or more times. We have the following rationality result.

Theorem 2.1. *The generating function* $\sum_{n\geq1} a_{1\cdots1}(n)x^n$ *is always rational.*

Proof. To show this, we first refine the numbers $a_{1\cdots1}(n)$ as follows. Let $a_{n,m} = a_{n,m}^{(k)}$ denote the number of Catalan words of length n containing m zeros and avoiding the pattern $1\cdots1$ of length k. By definition, $a_{n,m} = 0$ if $m \geq k$. Furthermore, assume $a_{n,m} = 0$ if it is not the case that $n \geq m \geq 1$. The numbers $a_{n,m}$ satisfy the recurrence

$$a_{n,m} = \sum_{j=1}^{k-1} \left(\binom{m+j-1}{j} - 1 \right) a_{n-m,j}, \qquad 2 \leq m < k \quad \text{and} \quad n > m, \tag{1}$$

with $a_{n,1} = \delta_{n,1}$ and $a_{n,n} = \chi(n < k)$ for $n \geq 1$. To show (1), observe that there are

$$\left(\binom{m+j-1}{j} - 1 \right) a_{n-m,j}$$

words of the form enumerated by $a_{n,m}$ and containing exactly j ones for $1 \leq j \leq k-1$. This follows from adding a single zero at the beginning of a Catalan word of length $n-m$, expressed using positive instead of non-negative letters, and then inserting $m-1$ additional zeros across $j+1$ possible positions so that not all of the zeros occur in the first position.

Define $A_m(x) = \sum_{n\geq m} a_{n,m}x^n$. Then recurrence (1) may be written equivalently as

$$A_m(x) = x^m + x^m \sum_{j=1}^{k-1} \left(\binom{m+j-1}{j} - 1 \right) A_j(x), \qquad 2 \leq m < k. \tag{2}$$

Let $\alpha_{i,j} = \binom{i+j}{i} - 1$. Then recurrence (2) may be expressed as

$$\begin{pmatrix} 1 - x\alpha_{1,0} & -x\alpha_{2,0} & \cdots & -x\alpha_{k-1,0} \\ -x^2\alpha_{1,1} & 1 - x^2\alpha_{2,1} & \cdots & -x^2\alpha_{k-1,1} \\ \vdots & \vdots & \ddots & \vdots \\ -x^{k-1}\alpha_{1,k-2} & -x^{k-1}\alpha_{2,k-2} & \cdots & 1 - x^{k-1}\alpha_{k-1,k-2} \end{pmatrix} \begin{pmatrix} A_1(x) \\ A_2(x) \\ \vdots \\ A_{k-1}(x) \end{pmatrix} = \begin{pmatrix} x \\ x^2 \\ \vdots \\ x^{k-1} \end{pmatrix}.$$

Note that the determinant of the coefficient matrix in the preceding system of linear equations has determinant 1 when $x = 0$ and is thus invertible for all x sufficiently close to zero, by continuity. By Cramer's rule, it follows that each $A_i(x)$, $1 \le i \le k-1$, is a rational function of x. Let $A^{(k)}(x) = \sum_{n\ge1} a_{1\cdots1}(n)x^n$. Then we have, by definition, $A^{(k)}(x) = A_1(x) + A_2(x) + \cdots + A_{k-1}(x)$, which implies $A^{(k)}(x)$ is rational. □

By the prior proof, we have the following expressions for $A^{(k)}(x)$ where $2 \le k \le 5$.

Corollary 2.2. *We have*

$$\begin{aligned} A^{(2)}(x) &= x, \\ A^{(3)}(x) &= \frac{x(1 + x - x^2)}{1 - 2x^2}, \\ A^{(4)}(x) &= \frac{x(1 + x - 7x^3 - 3x^4 + x^5)}{1 - 2x^2 - 9x^3 + 3x^5}, \\ A^{(5)}(x) &= \frac{x(1 + x - 6x^3 - 34x^4 - 22x^5 + 3x^6 + 25x^7 + 6x^8 - x^9)}{1 - 2x^2 - 9x^3 - 34x^4 + 3x^5 + 32x^6 + 40x^7 - 4x^9}. \end{aligned}$$

In the case $k = 3$, we have the following further enumerative result, which can also be obtained by extracting the coefficient of x^n in the expression above for $A^{(3)}(x)$.

Corollary 2.3. *Let $n \ge 2$ and $t = \lfloor \frac{n+1}{2} \rfloor$. Then we have $a_{111}(n, i) = 2^{\lfloor \frac{n-2}{2} \rfloor} \cdot \delta_{i,t}$ for $1 \le i \le t$.*

Proof. We provide a combinatorial proof. If n is even, then all letters within a member of $W_{111}(n)$ must occur twice, while if n is odd, all letters occur twice, except for the largest, which occurs once. Then we only need to show $a_{111}(n) = 2^{\lfloor \frac{n-2}{2} \rfloor}$ for $n \ge 2$. First assume $n = 2m$, in which case we show $a_{111}(2m) = 2^{m-1}$. Then each letter $0, 1, \ldots, m-1$ occurs exactly twice. We first form the word $01 \cdots (m-1)$ and write an $\times$ after the letter $m-1$ and some subset of $[m-2]$, where we are assuming $m \ge 2$. After the smallest i after which an $\times$ is placed, we write all letters in $\{0\} \cup [i-1]$ directly following i in a descending block and erase the $\times$. For each subsequent i followed by an $\times$, we write all letters in $\{0\} \cup [i-1]$ that have not yet appeared twice in a descending block following i. After the descending block of letters that are to follow $m-1$ is written, we write an $m-1$ either at the very end of the current sequence or directly following the $m-1$ already appearing in it.

For example, if $m = 8$ and $\times$'s are to follow $\{1, 2, 5\} \subseteq [6]$ in addition to the 7, then we would have $01{\times}2{\times}345{\times}67{\times}$, which would give rise to the words

$$0102134543267657 \quad \text{and} \quad 0102134543267765.$$

Note that within each descending block, one must include all numbers not yet appearing twice, in particular, the smallest, lest there be an occurrence of 111. Thus, all members of $W_{111}(2m)$ can be formed in this way, which implies $a_{111}(2m) = 2^{m-2} \cdot 2 = 2^{m-1}$. If $n = 2m + 1$, then write $01 \cdots m$ and place an $\times$ after the m and some subset of $[m-1]$. Then write all letters in $\{0\} \cup [i-1]$ that have not yet appeared twice in a descending block after i for each i followed by an $\times$. This implies $a_{111}(2m + 1) = 2^{m-1}$ and completes the proof. □

Remark: While we do not have an explicit expression for $a_{11\cdots1}(n)$ for all k, it is possible to express it in terms of compositions in the following way. Let $d = (d_0, d_1, \ldots)$ denote a composition of n in which each part d_i

satisfies $1 \le d_i < k$. Then by [6, Proposition 3.1], we have

$$a_{11\cdots 1}(n) = \sum_{d} \prod_{i\ge 0} \left(\binom{d_i + d_{i+1} - 1}{d_{i+1}} - 1 \right), \qquad n \ge 1, \tag{3}$$

where the sum is taken over all compositions $d = (d_0, d_1, \ldots)$ of n whose parts are less than k (with no restriction as to the number of parts).

2.2. The Pattern $12\cdots k$

In this subsection, we consider the avoidance of the pattern $12\cdots k$, where $k \ge 2$. Note that by the second defining property of a Catalan sequence, a member of $W(n)$ avoids $12\cdots k$ if and only if it contains fewer than k distinct letters. Thus, avoiding the pattern $12\cdots k$ is seen to be logically equivalent to avoiding the pattern $12\cdots(k-2)k(k-1)$. We have the following result concerning either pattern.

Theorem 2.4. *The generating function* $\sum_{n\ge 1} a_{12\cdots k}(n)x^n$ *is always rational.*

Proof. To prove this, we again refine the avoidance class in question by considering the statistic that records the number of zeros. Let $a^{(k)}_{n,m}$ denote the number of Catalan words of length n containing m zeros and avoiding the pattern $12\cdots k$, where $k \ge 2$. Then we have the recurrence

$$a^{(k)}_{n,m} = \sum_{i=1}^{n-m} \left(\binom{i+m-1}{i} - 1 \right) a^{(k-1)}_{n-m,i}, \qquad 2 \le m < n \quad \text{and} \quad k \ge 3, \tag{4}$$

with $a^{(2)}_{n,m} = \delta_{n,m}$ for $1 \le m \le n$ and $a^{(k)}_{n,1} = \delta_{n,1}$ and $a^{(k)}_{n,n} = 1$ for all $n \ge 1$ and $k \ge 3$.

To show (4), note that words enumerated by $a^{(k)}_{n,m}$ may be obtained from words enumerated by $a^{(k-1)}_{n-m,i}$ for some i, expressed using positive letters, by inserting zeros at the beginning and after ones. Given any word of the latter type, there are $\binom{i+m-1}{i} - 1$ words of the former type that are obtained from it in the manner described, with the operation seen to be reversible.

Define $A^{(k)}_m(x) = \sum_{n\ge m} a^{(k)}_{n,m}x^n$. Then recurrence (4) can be written as

$$A^{(k)}_m(x) = x^m + x^m \sum_{i\ge 1} \left(\binom{m-1+i}{i} - 1 \right) A^{(k-1)}_i(x), \qquad m \ge 2, \quad k \ge 3,$$

with $A^{(k)}_1(x) = x$ and $A^{(2)}_m(x) = x^m$.

Define $A^{(k)}(x,t) = \sum_{m\ge 1} A^{(k)}_m(x)t^m$. Then

$$A^{(k)}(x,t) = \frac{xt}{1-xt}(1 - A^{(k-1)}(x,1)) + \frac{xt}{1-xt}A^{(k-1)}(x, 1/(1-xt)), \qquad k \ge 3,$$

with $A^{(2)}(x,t) = \frac{xt}{1-xt}$.

Define $A(x,t,v) = \sum_{k\ge 2} A^{(k)}(x,t)v^k$. Then

$$A(x,t,v) = \frac{xtv}{1-xt}\left(\frac{v}{1-v} - A(x,1,v)\right) + \frac{xtv}{1-xt}A(x, 1/(1-xt), v).$$

Let $\rho_i = \frac{1}{1-x\rho_{i-1}}$ for $i \ge 1$, with $\rho_0 = t$ Then

$$A(x,\rho_i,v) = \frac{x\rho_i v}{1-x\rho_i}\left(\frac{v}{1-v} - A(x,1,v)\right) + \frac{x\rho_i v}{1-x\rho_i}A(x, 1/(1-x\rho_i), v),$$

which is equivalent to

$$A(x,\rho_i,v) = xv\rho_i\rho_{i+1}\left(\frac{v}{1-v} - A(x,1,v)\right) + xv\rho_i\rho_{i+1}A(x,\rho_{i+1},v). \tag{5}$$

Iteration of (5), assuming x and v are sufficiently small in absolute value, leads to

$$A(x,t,v) = \sum_{j\geq 1}(xv)^j \rho_0\rho_1^2\cdots\rho_{j-1}^2\rho_j\left(\frac{v}{1-v} - A(x,1,v)\right). \tag{6}$$

Recall that the Chebyshev polynomials of the second kind (see [9]) are defined by the recurrence $U_i(y) = 2yU_{i-1}(y) - U_{i-2}(y)$ if $i \geq 2$, with $U_0(y) = 1$ and $U_1(y) = 2y$. By the recurrence for U_i and induction, we have

$$\rho_i = \frac{2zU_{i-1}(z) - tU_{i-2}(z)}{\sqrt{x}(2zU_i(z) - tU_{i-1}(z))}, \qquad i \geq 0,$$

where $z = \frac{1}{2\sqrt{x}}$. Therefore,

$$\rho_0\rho_1^2\cdots\rho_{j-1}^2\rho_j = \frac{2zt}{x^j(2zU_{j-1}(z) - tU_{j-2}(z))(2zU_j(z) - tU_{j-1}(z))}, \qquad j \geq 1.$$

Taking $t = 1$ in (6) then gives

$$A(x,1,v) = \sum_{j\geq 1}\frac{2zv^j}{U_j(z)U_{j+1}(z)}\left(\frac{v}{1-v} - A(x,1,v)\right),$$

which implies

$$A(x,1,v) = \left(\frac{v}{1-v}\right)\frac{\sum_{j\geq 1}\frac{v^j}{\sqrt{x}U_j(z)U_{j+1}(z)}}{1 + \sum_{j\geq 1}\frac{v^j}{\sqrt{x}U_j(z)U_{j+1}(z)}}.$$

Let $A^{(k)}(x) = \sum_{n\geq 1} a_{12\cdots k}(n)x^n$. By the definitions, we have that $A^{(k)}(x)$ is the coefficient of v^k in the generating function $A(x,1,v)$, which implies

$$A^{(k)}(x) = [v^k]\left((v + v^2 + \cdots + v^{k-1})\frac{\sum_{j=1}^k \frac{v^j}{\sqrt{x}U_j(z)U_{j+1}(z)}}{1 + \sum_{j=1}^k \frac{v^j}{\sqrt{x}U_j(z)U_{j+1}(z)}}\right). \tag{7}$$

By induction, one can show that if j is odd, then $U_j(z)$ is a polynomial in $\frac{1}{\sqrt{x}}$ containing only odd powers of $\frac{1}{\sqrt{x}}$, while if j is even, $U_j(z)$ is a polynomial in $\frac{1}{x}$. This implies that the product $\sqrt{x}U_j(z)U_{j+1}(z)$ is a polynomial in $\frac{1}{x}$ for all $j \geq 1$. Hence, by (7), the generating function $A^{(k)}(x)$ is always rational. □

Note that the difference $A^{(k+2)}(x) - A^{(k+1)}(x)$ is the generating function for the number of Catalan words of length n having largest letter k and is thus rational.

We have the following explicit formulas for $a_{12\cdots k}(n)$ when $2 \leq k \leq 5$.

Corollary 2.5. *If $n \geq 1$, then $a_{12}(n) = 1$, $a_{123}(n) = a_{132}(n) = 2^{n-1} - n + 1$, $a_{1234}(n) = F_{2n-1} - 2^n + \binom{n}{2} + 2$, and*

$$a_{12345}(n) = (11-n)2^{n-2} + (3^{n-1} - 1)/2 - (n+1)(n^2 - n + 12)/6 - 2F_{2n-2}.$$

Proof. These formulas follow from computing $A^{(k)}(x)$ using (7) and then extracting the coefficient of x^n. The $k = 2$ and $k = 3$ cases of these formulas also follow directly from the definitions. When $k = 3$, note that a Catalan word avoids 123 or 132 if and only if it is binary and that there are $2^{n-1} - (n-1)$ binary Catalan words of length n, upon taking away from the set of all binary words of length n starting with 0 those having the form 0^i1^j, where $1 \leq i \leq n-1$. □

Remark: By [6, Proposition 3.1], we have

$$a_{12\cdots k}(n) = \sum_d \prod_{i\geq 0}\left(\binom{d_i + d_{i+1} - 1}{d_{i+1}} - 1\right), \qquad n \geq 1, \tag{8}$$

where the sum is taken over all compositions $d = (d_0, d_1, \ldots)$ of n having less than k parts (with no restriction as to the size of each part).

2.3. The Pattern $1\cdots 12$

We consider avoidance of the pattern $1\cdots 12$ of length $k+1$, where $k \geq 2$ is fixed. Let $a_{n,m} = a_{n,m}^{(k)}$ denote the number of Catalan words of length n containing m zeros and avoiding the pattern $1\cdots 12$. Let $\mathcal{A}_{n,m} = \mathcal{A}_{n,m}^{(k)}$ denote the subset of Catalan words enumerated by $a_{n,m}$. Let $\mathcal{B}_{n,m}$ and $C_{n,m}$ denote the subsets of $\mathcal{A}_{n,m}$ whose members end in a zero or non-zero letter, respectively. Define $b_{n,m} = |\mathcal{B}_{n,m}|$ and $c_{n,m} = |C_{n,m}|$. Assume all cardinalities are zero if it is not the case that $n \geq m \geq 1$.

The arrays $b_{n,m}$ and $c_{n,m}$ are determined recursively as follows.

Lemma 2.6. *If* $n \geq 3$ *and* $2 \leq m < n$*, then*

$$b_{n,m} = \sum_{i=1}^{n-m} \binom{i+k-2}{i} b_{n-m,i}, \qquad m \geq k, \tag{9}$$

and

$$b_{n,m} = \sum_{i=1}^{n-m} \binom{i+m-2}{i} b_{n-m,i}, \qquad m < k, \tag{10}$$

with $b_{n,1} = \delta_{n,1}$ *and* $b_{n,n} = 1$ *for all* $n \geq 1$*. If* $2 \leq m < n$*, then*

$$c_{n,m} = \sum_{i=1}^{n-m} \left(\binom{i+m-2}{i-1} - 1 \right) b_{n-m,i} + \sum_{i=1}^{n-m} \left(\binom{i+m-1}{i} - 1 \right) c_{n-m,i}, \qquad m < k, \tag{11}$$

with $c_{n,m} = 0$ *if* $m \geq k$ *and* $c_{n,1} = c_{n,n} = 0$ *for all* $n \geq 1$*.*

Proof. The initial values follow easily from the definitions. Note also that $c_{n,m} = 0$ if $m \geq k$ since a word not ending in zero and containing k or more zeros would have a subsequence of $1\cdots 12$ in which the 2 corresponds to the final letter of the word. To show the recurrences for $b_{n,m}$, we insert m zeros into $\pi \in \mathcal{B}_{n-m,i}$ for some i, expressed using positive letters. Once the first zero is written at the beginning of π, we have for $m \geq k$ that there are $\binom{i+k-2}{i}$ ways in which to distribute the next $k-2$ zeros within the word 0π. The final $m-k+1$ zeros are then written at the end of the present word which yields a member of $\mathcal{B}_{n,m}$. Note that members of $\mathcal{B}_{n,m}$ where $m \geq k$ must end in at least $m-k+1$ zeros in order to avoid an occurrence of $1\cdots 12$. If $m < k$, then there are $\binom{i+m-2}{i}$ ways to insert $m-2$ zeros into the word $0\pi 0$. In either case, the operation of inserting zeros is seen to be reversible. Summing over all possible i then gives (9) and (10).

To show (11), we insert m zeros into $\pi \in \mathcal{A}_{n-m,i}$ expressed using positive letters to obtain members of $C_{n,m}$. If π ends in a 1, then there are $\binom{i+m-2}{i-1} - 1$ ways to distribute $m-1$ zeros across i possible positions within 0π since no zero can follow the final 1 and since it cannot be the case that all of the zeros precede the first 1. Thus, the first sum on the right-hand side of (11) counts all members of $C_{n,m}$ ending in 1. If π does not end in 1, then there are $\binom{i+m-1}{i} - 1$ ways of distributing the zeros within 0π. This implies the second sum on the right-hand side of (11) counts all members of $C_{n,m}$ ending in a letter greater than 1, which completes the proof. □

We can now establish the following result.

Theorem 2.7. *The generating function* $\sum_{n\geq 1} a_{1\cdots 12}(n)x^n$ *is always rational.*

Proof. Let $B_m(x) = \sum_{n\geq m} b_{n,m}x^n$. Then recurrences (9) and (10) imply

$$B_m(x) = x^m + x^m \sum_{i\geq 1} \binom{i+k-2}{i} B_i(x) = x^{m-k}B_k(x), \qquad m > k, \tag{12}$$

$$B_m(x) = x^m + x^m \sum_{i\geq 1} \binom{i+m-2}{i} B_i(x), \qquad 2 \leq m \leq k, \tag{13}$$

with $B_1(x) = x$. Define

$$\begin{aligned}\alpha_{m,k} &= \sum_{i\geq k}\binom{i+m-2}{i}B_i(x)/B_k(x) = \sum_{i\geq k}\binom{i+m-2}{i}x^{i-k}\\ &= \frac{1}{x^k}\left(\sum_{i\geq 0}\binom{i+m-2}{i}x^i - \sum_{i=0}^{k-1}\binom{i+m-2}{i}x^i\right)\\ &= \frac{1}{x^k}\left(\frac{1}{(1-x)^{m-1}} - \sum_{i=0}^{k-1}\binom{i+m-2}{i}x^i\right).\end{aligned}$$

Then the recurrences in (12) and (13) can be expressed in matrix form as

$$\begin{pmatrix} 1-x^2\binom{2}{0} & -x^2\binom{3}{0} & \cdots & -x^2\binom{k-1}{0} & -x^2\alpha_{2,k}\\ -x^3\binom{3}{1} & 1-x^3\binom{4}{1} & \cdots & -x^3\binom{k}{1} & -x^3\alpha_{3,k}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ -x^{k-1}\binom{k-1}{k-3} & -x^{k-1}\binom{k}{k-3} & \cdots & 1-x^{k-1}\binom{2k-4}{k-3} & -x^{k-1}\alpha_{k-1,k}\\ -x^k\binom{k}{k-2} & -x^k\binom{k+1}{k-2} & \cdots & -x^k\binom{2k-3}{k-2} & 1-x^k\alpha_{k,k}\end{pmatrix}\begin{pmatrix}B_2(x)\\ B_3(x)\\ \vdots\\ B_{k-1}(x)\\ B_k(x)\end{pmatrix} = \begin{pmatrix}g_2\\ g_3\\ \vdots\\ g_{k-1}\\ g_k\end{pmatrix},$$

where $g_i = x^i(1+(i-1)x)$ for $2 \leq i \leq k$. Since the $\alpha_{m,k}$ are rational functions, it follows from Cramer's rule that $B_m(x)$ is rational for any $m = 2, 3, \ldots, k$.

Let $C_m(x) = \sum_{n\geq m} c_{n,m}x^n$. By (11), we have

$$\begin{aligned}C_m(x) &= x^m\sum_{i=1}^{k-1}\left(\binom{i+m-2}{i-1}-1\right)B_i(x) + x^m\sum_{i=1}^{k-1}\left(\binom{i+m-1}{i}-1\right)C_i(x)\\ &\quad + x^{m-k+1}B_k(x)\left(\frac{1}{(1-x)^m} - \sum_{i=0}^{k-2}\binom{i+m-1}{i}x^i - \frac{x^{k-1}}{1-x}\right), \qquad 1<m<k, \end{aligned} \tag{14}$$

with $C_m(x) = C_1(x) = 0$ if $m \geq k$. Since the $B_i(x)$, $1 \leq i \leq k$, are rational functions, it follows that $C_m(x)$ is also rational for $m = 2, 3, \ldots, k-1$. From the definitions and (12), we have

$$\sum_{n\geq 1} a_{1\cdots 12}(n)x^n = \sum_{m\geq 1}(B_m(x)+C_m(x)) = \frac{B_k(x)}{1-x} + \sum_{m=1}^{k-1}(B_m(x)+C_m(x)),$$

which gives the desired result. □

Taking $k = 2$ in the prior proof implies

$$B_2(x) = \frac{x^2(1+x)}{1-x^2\alpha_{2,2}} = \frac{x^2(1-x^2)}{1-x-x^2}.$$

We then have

$$\sum_{n\geq 1} a_{112}(n)x^n = x + \frac{B_2(x)}{1-x} = \frac{x}{1-x-x^2} = \sum_{n\geq 1} F_n x^n.$$

Thus we have the following apparently new combinatorial interpretation of the Fibonacci sequence.

Corollary 2.8. *If $n \geq 1$, then $a_{112}(n) = F_n$. Furthermore, we have $a_{112}(n,m) = \binom{n-m}{m-1}$ for $1 \leq m \leq \lfloor\frac{n+1}{2}\rfloor$.*

Proof. We provide a combinatorial proof of the first statement which will also imply the second. Suppose $\lambda \in W_{112}(n)$. Then we have for some $m \geq 1$,

$$\lambda = 01\cdots(m-2)(m-1)^{a_{m-1}}\cdots 0^{a_0},$$

where $\sum_{i=0}^{m-1} a_i = n-m+1$ and each a_i is a positive integer. Recall that F_n counts the number of linear square-and-domino tilings [2, Chapter 1] of length $n-1$, the set of which we will denote by $\mathcal{T}_{n-1}$. Define a mapping $f: W_{112}(n) \to \mathcal{T}_{n-1}$ by letting

$$f(\lambda) = s^{a_0-1}(ds^{a_1-1})(ds^{a_2-1})\cdots(ds^{a_{m-1}-1}),$$

where s and d denote square and domino, respectively. Note that the tiling $f(\lambda)$ has length given by $\sum_{i=0}^{m-1} a_i - m + 2(m-1) = n-1$, and one may verify that f is a bijection. Furthermore, the number of distinct letters of λ is one more than the number of dominos in $f(\lambda)$ for all λ, whence the second statement follows from the fact that there are $\binom{n-m}{m-1}$ members of $\mathcal{T}_{n-1}$ containing $m-1$ dominos. □

2.4. The Pattern $2\cdots21$

We consider the avoidance of the pattern $2\cdots21$ of length $k+1$, where $k \geq 2$. Let $a_{n,m} = a_{n,m}^{(k)}$ denote the number of Catalan words of length n containing m zeros and avoiding the pattern $2\cdots21$. The sequence $a_{n,m}$ is determined recursively as follows.

Lemma 2.9. *If $n \geq 3$, then*

$$a_{n,m} = \sum_{i=1}^{k-1}\left(\binom{i+m-1}{i}-1\right)a_{n-m,i} + \sum_{i=k}^{n-m}\left(\binom{k+m-2}{k-1}-1\right)a_{n-m,i}, \qquad 2 \leq m < n, \tag{15}$$

with $a_{n,1} = \delta_{n,1}$ and $a_{n,n} = 1$ for all $n \geq 1$. Furthermore, any pattern of the same length of the form $2\cdots212\cdots2$ is equivalent to $2\cdots21$ in terms of avoidance by Catalan sequences, with this equivalence respecting the number of occurrences of each letter.

Proof. Let $\rho = 2\cdots21$ and let $\mathcal{A}_{n,m}$ denote the subset of $W_\rho(n)$ whose members are enumerated by $a_{n,m}$. Then members of $\mathcal{A}_{n,m}$ may be obtained by writing zeros at the beginning and after any of the ones within members of $\mathcal{A}_{n-m,i}$, expressed using positive letters, if $i < k$, and after only the first $k-1$ ones if $i \geq k$. There are thus $\binom{i+m-1}{i}-1$ possible ways to insert zeros if $i<k$ and $\binom{k+m-2}{k-1}-1$ ways if $i \geq k$. Furthermore, note that no occurrences of ρ are created in which the role of the 2 is played by some number $\ell > 1$ when zeros are inserted as described, for otherwise the member of $\mathcal{A}_{n-m,i}$ would have already contained an occurrence of ρ, which is impossible. Summing over all i then gives (15).

Now let $\tau = 2^r12^s$, where $r+s=k$ and $r,s \geq 1$. We argue that the corresponding enumerating sequence for τ satisfies (15) as well, from which the second statement will follow (as the initial conditions are the same). Clearly, the first sum on the right-hand side of (15) is the same since there must be at least k ones to produce an occurrence of τ when the zeros are inserted. If $i \geq k$, then one can insert zeros either at the beginning or after the first $r-1$ ones or after the final s ones. Thus, there are a total of $r-1+s=k-1$ ones after which one may insert zeros, which implies the second sum in (15). □

We have the following rationality result which thus applies to each of the patterns in the previous lemma.

Theorem 2.10. *The generating function $\sum_{n\geq1} a_{2\cdots21}(n)x^n$ is always rational.*

Proof. Let $A_m(x) = \sum_{n\geq m} a_{n,m}x^n$. Then recurrence (15) may be written as

$$A_m(x) = x^m + x^m\sum_{i=1}^{k-1}\left(\binom{i+m-1}{i}-1\right)A_i(x) + x^m\left(\binom{k+m-2}{k-1}-1\right)\sum_{i\geq k}A_i(x), \qquad m \geq 2, \tag{16}$$

with $A_1(x) = x$. Let $A(x) = \sum_{m\geq 1} A_m(x)$. Then summing both sides of (16) over m implies

$$A(x) = \frac{x}{1-x} + \sum_{i=1}^{k-1} x\left(\frac{1}{(1-x)^{i+1}} - \frac{1}{1-x}\right)A_i(x) + x\left(\frac{1}{(1-x)^k} - \frac{1}{1-x}\right)\sum_{i\geq k} A_i(x).$$

Since $A(x) = \sum_{i=1}^{k-1} A_i(x) + \sum_{i\geq k} A_i(x)$, this last equation may be rewritten as

$$\sum_{i=1}^{k-1}\left(1 - x\left(\frac{1}{(1-x)^{i+1}} - \frac{1}{1-x}\right)\right)A_i(x) + \left(1 - x\left(\frac{1}{(1-x)^k} - \frac{1}{1-x}\right)\right)\sum_{i\geq k} A_i(x) = \frac{x}{1-x}. \tag{17}$$

Let $\alpha_{i,j} = \binom{i+j}{i} - 1$, $\beta_i(x) = 1 - x\left(\frac{1}{(1-x)^{i+1}} - \frac{1}{1-x}\right)$, $g_i(x) = x^i(1+(i-1)x)$, $g(x) = \frac{x^2}{(1-x)^2}$ and $A^*(x) = \sum_{i\geq k} A_i(x)$. Taking $2 \leq m \leq k-1$ in (16), together with (17), yields the linear system of equations $AX = Y$, where

$$A = \begin{pmatrix} 1 - x^2\alpha_{2,1} & -x^2\alpha_{3,1} & \cdots & -x^2\alpha_{k-1,1} & -x^2\alpha_{k-1,1} \\ -x^3\alpha_{2,2} & 1 - x^3\alpha_{3,2} & \cdots & -x^3\alpha_{k-1,2} & -x^3\alpha_{k-1,2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -x^{k-1}\alpha_{2,k-2} & -x^{k-1}\alpha_{3,k-2} & \cdots & 1 - x^{k-1}\alpha_{k-1,k-2} & -x^{k-1}\alpha_{k-1,k-2} \\ \beta_2(x) & \beta_3(x) & \cdots & \beta_{k-1}(x) & \beta_{k-1}(x) \end{pmatrix}$$

and

$$X = \begin{pmatrix} A_2(x) \\ A_3(x) \\ \vdots \\ A_{k-1}(x) \\ A^*(x) \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} g_2(x) \\ g_3(x) \\ \vdots \\ g_{k-1}(x) \\ g(x) \end{pmatrix}.$$

By Cramer's rule, the $A_i(x)$, $2 \leq i \leq k-1$, and $A^*(x)$ are rational. This implies that

$$\sum_{n\geq 1} a_{2\cdots 21}(n)x^n = A(x) = \sum_{i=1}^{k-1} A_i(x) + A^*(x)$$

is rational, which completes the proof. □

Taking $k = 2$ in the prior proof implies

$$\left(1 - x\left(\frac{1}{(1-x)^2} - \frac{1}{1-x}\right)\right)A^*(x) = \frac{x^2}{(1-x)^2},$$

and thus

$$\sum_{n\geq 1} a_{221}(n)x^n = A(x) = x + A^*(x) = x + \frac{x^2}{1-2x} = \frac{x(1-x)}{1-2x}.$$

Extracting the coefficient of x^n gives the first part of the following result.

Corollary 2.11. *If $n \geq 2$, then $a_{221}(n) = a_{212}(n) = 2^{n-2}$. Furthermore, if $a_n = a_{221}(n; y) = a_{212}(n; y)$, then*

$$a_n = 2a_{n-1} - (1-y)a_{n-2}, \qquad n \geq 3, \tag{18}$$

with $a_1 = a_2 = y$.

Proof. Taking $y = 1$ in (18) gives the first statement. We prove (18) in the case 212. To do so, note that the weight of those members of $W_{212}(n)$ ending in two or more zeros is a_{n-1}, upon removing the terminal zero. The weight of those members ending in a single zero and starting with two or more zeros is $a_{n-1} - a_{n-2}$, by subtraction, upon removing the initial zero. Finally, the weight of all words in $W_{212}(n)$ beginning and ending in a single zero is ya_{n-2}, upon removing both zeros from each word and reducing the remaining letters by one. Adding the contributions from the three previous cases gives (18). By Lemma 2.9, the recurrence (18) also holds for the pattern 221.

Alternatively, to complete the proof, one can define a bijection f between $W_{212}(n)$ and $W_{221}(n)$ that preserves the number of distinct letters. Suppose $w \in W_{212}(n)$. Then we can write for some $m \geq 0$,

$$w = 0^{a_0}1^{a_1}\cdots m^{a_m}(m-1)^{b_{m-1}}\cdots 1^{b_1}0^{b_0}, \qquad a_i, b_i \geq 1 \quad \text{for all } i,$$

where $a_m + \sum_{i=0}^{m-1}(a_i + b_i) = n$. Define f by setting

$$f(w) = 0^{a_0}(10^{b_0}1^{a_1-1})(21^{b_1}2^{a_2-1})\cdots(m(m-1)^{b_{m-1}}m^{a_m-1}).$$

One may verify that f is the desired bijection. □

3. Three Letter Patterns

In this section, we consider the remaining patterns of length three that do not follow from the results of the previous section. Our work is reduced by noting the following.

Observation 3.1. *The only 121 avoiding Catalan word of length n is the sequence* 0^n*, while the only members of* $W_{122}(n)$ *are of the form* 0^n *or* 0^i10^j*, where* $1 \leq i \leq n-2$*, whence* $a_{122}(n) = n - 1$ *if* $n \geq 2$*.*

In the first subsection that follows, we consider the cases 321 and 211 and provide combinatorial arguments of our results. In the second, we enumerate the avoidance classes in the cases of 312 and 213. These cases are more readily established by algebraic arguments and we make use of generating function techniques in our proofs.

3.1. The Cases 321 and 211

In this subsection, we consider avoidance of the patterns 321 and 211. We first consider the case 321. Avoiding 321 is equivalent to avoiding 231 as a consequence of the following result.

Proposition 3.2. *If* $k \geq 3$ *and* $1 < i < k$*, then avoiding the pattern* $i\cdots(k-1)k(i-1)\cdots 21$ *by Catalan sequences is logically equivalent to avoiding* $k\cdots 21$*.*

Proof. Let $\rho = k\cdots 21$ and $\tau = i\cdots(k-1)k(i-1)\cdots 21$. First suppose that a Catalan word w contains an occurrence x of ρ. By the second defining property of Catalan words, the subsequence y of initial occurrences of letters is increasing. The part of y corresponding to the letters represented by $i, i+1, \ldots, k$ in x, taken together with the last $i-1$ letters in x, then constitutes an occurrence of τ. Now suppose w contains and occurrence z of τ. By the first property of Catalan words, all drops are of size 1. Thus, between any occurrence of the letter playing the role of k in z and any occurrence of the letter playing the role of $i-1$ to the right of it, all of the letters corresponding to those in $[i, k-1]$ within z must occur within some decreasing subsequence, which implies that w contains an occurrence of ρ. □

We are able to enumerate the avoidance class for the patterns in Proposition 3.2 in the case when $k = 3$.

Theorem 3.3. *If* $n \geq 1$*, then* $a_{321}(n) = a_{231}(n) = F_{2n-3}$*.*

Proof. By Proposition 3.2, avoiding 231 is logically equivalent to avoiding 321, so we only consider the latter. Let $W_n = W_{321}(n)$ and $a_n = |W_n|$. We will show that

$$a_n = 3a_{n-1} - a_{n-2}, \qquad n \geq 3, \tag{19}$$

with $a_1 = a_2 = 1$. We will assume $n \geq 5$, since (19) clearly holds for $n = 3, 4$. Let R denote the subset of W_n whose members start with at least two zeros. Then $|R| = a_{n-1}$, upon adding a zero to the beginning of each member of W_{n-1}. Let S denote the subset of W_n whose members w are of one of the following forms: (i) $w = 010^a1^b0^c\alpha$, where $a, b, c \geq 1$ and α is either empty or is non-empty and starts with 1 or 2, or (ii) $w = 01^{a+1}0^b\alpha$, where $a, b \geq 1$ and α is as in (i). Then writing a 1 directly after the first zero of each member of W_{n-1} defines a bijection between W_{n-1} and the set $S \cup \{010^{n-2}\}$. Thus it remains to show that the cardinality of the set

$$T := W_n - R - (S \cup \{010^{n-2}\})$$

is $a_{n-1} - a_{n-2}$.

Note that $w \in T$ implies that it must be of the form $w = 010^i1^j\beta$, where either β is empty and $i, j \geq 1$ or β is non-empty starting with 2 and $i \geq 1$, $j \geq 0$. Note that $j \geq 1$ in the first case since we have already excluded the word 010^{n-2}. Furthermore, if β is non-empty, then it must contain at least two letters (in particular, a 2 and a 1). Moreover, β can contain no 0's, lest there be an occurrence of the pattern 321 in w.

We now define a mapping f from T to W_{n-1}. Suppose $w \in T$. By the preceding observations, we have that w is of one of the following three forms: (a) $w = 010^i1^j$, where $i, j \geq 1$, (b) $w = 010^i1^j2\gamma$, where $i \geq 1$, $j \geq 0$ and γ is non-empty and starts with 1, or (c) $w = 010^i1^j2\gamma$, where $i \geq 1$, $j \geq 0$ and γ starts with 2. Note that γ contains no zeros in forms (b) and (c). Define f by letting

$$f(w) = \begin{cases} 010^i1^{j-1}, & \text{if } w \text{ is of form (a);} \\ 0^{j+1}1^{i+1}\text{red}(\gamma), & \text{if } w \text{ is of form (b);} \\ 0^{j+1}10^i\text{red}(\gamma), & \text{if } w \text{ is of form (c),} \end{cases}$$

where red(γ) is obtained by reducing each letter of γ by one. One may verify that f is one-to-one, and thus $|T| = |f(T)|$.

Note that $f(T)$ comprises all members x of W_{n-1} in which either (I) x is binary and starts 010 and contains only two runs of 0's, (II) the first run of 1's in x has length at least two, or (III) the first run of 1's has length one, with at least three runs of 0's in x. To complete the proof, it suffices to show $|U| = a_{n-2}$, where $U := W_{n-1} - f(T)$. Observe that U consists of all words u having one of the following three forms, where $i \geq 1$ and $k \geq 0$:

(i) $u = 0^{j+1}10^i1^k\delta$, where $j \geq 0$ and $\delta \neq \varnothing$ starts with 2,
(ii) $u = 0^{j+1}10^i1^k$, where $j \geq 1$, or
(iii) $u = 0^{n-1}$.

Given $u \in U - \{0^{n-1}\}$, let $u' = 0^{j+1}1^{k+1}0^i\delta$ if u is of the form (i) above and let $u' = 0^{j+1}1^{k+1}0^i$ if u is of form (ii). Then the mapping $u \mapsto u'$ is a bijection. Let V denote the subset of words in W_{n-1} of the form

$$v = 0^{j+1}1^{k+1}0^i\delta, \qquad i \geq 1,\ k \geq 0, \tag{20}$$

together with 0^{n-1}, where δ is either empty or is non-empty and starts with 2 and $j \geq 1$ if δ is empty and $j \geq 0$ if δ is non-empty. Then $|V| = |U|$, so to complete the proof, we define a bijection between V and W_{n-2}.

Suppose $v \in V$, with $v \neq 0^{n-1}$. Note that if $\delta \neq \varnothing$ in (20), then it may be decomposed as $\delta = \delta_1\delta_2$, where δ_1 starts with 2 and contains letters in $\{1, 2\}$ and δ_2 is either empty or is non-empty and starts with 3. Note that δ_1 must contain at least one letter 1 and that δ_2, if non-empty, must contain at least one 2. Given a non-empty word w, let $\widetilde{w}$ denote the reverse of w and let w^* be obtained from w by deleting the final letter. We define a mapping $g : V \to W_{n-2}$ by letting

$$g(v) = \begin{cases} 0^i1^{k+1}\text{red}(\widetilde{\delta_1})0^j\text{red}(\delta_2), & \text{if } \delta \neq \varnothing \text{ and } \delta_1 \text{ ends in a 1;} \\ 0^{j+1}1^{k+1}0^i\delta_1^*\delta_2, & \text{if } \delta \neq \varnothing \text{ and } \delta_1 \text{ ends in a 2;} \\ 0^j1^{k+1}0^i, & \text{if } \delta = \varnothing; \\ 0^{n-2}, & \text{if } v = 0^{n-1}. \end{cases}$$

One may verify that g is a bijection between V and W_{n-2} and thus $|U| = a_{n-2}$, as desired. □

Remark 3.4. *If $a_n = a_{321}(n; y)$, then the preceding combinatorial argument can be generalized to show*

$$a_n = (2+y)a_{n-1} - ya_{n-2} - y(1-y) + y^2(1-y)(n-3+\sum_{i=1}^{n-5}\sum_{j=1}^{n-4-i}(2^j-1)a_{n-3-i-j}), \qquad n \geq 3,$$

with $a_1 = a_2 = y$. One can show (using generating functions, for example) that this is equivalent to the recurrence

$$a_n = 3a_{n-1} - (2-y)a_{n-2}, \qquad n \geq 3, \tag{21}$$

with the same initial conditions. It would be interesting to have a direct combinatorial proof of (21).

We now consider the case of avoiding 211.

Theorem 3.5. *The sequence $a_n = a_{211}(n; y)$ satisfies the recurrence*

$$a_n = 2a_{n-1} + (2y-1)a_{n-2} - ya_{n-3}, \qquad n \geq 4, \tag{22}$$

with $a_1 = a_2 = y$ and $a_3 = y + y^2$.

Proof. Let $W_n = W_{211}(n)$. To show (22), first note that the weight of all members of W_n starting with two or more 0's is a_{n-1}, upon writing a 0 directly before each member of W_{n-1}. Given a word $w = w_1w_2\cdots$ whose letters are integers, let inc(w) denote the word obtained by increasing each letter of w by one. Then writing a 0 before the initial 1 and after the right-most 1 of inc(w) for each $w \in W_{n-2}$ implies that the weight of all members of W_n starting with a single zero and containing a zero after the right-most 1 is ya_{n-2}. Now observe that within a member of W_n, not all zeros, that there is a single zero either after the first run of ones or the last, with no other zeros outside of the initial run of zeros. Furthermore, there are at most three runs of ones altogether, with three runs possible only when there is a (single) zero after the first run.

Let W'_n denote the subset of W_n whose members start with a single zero, with a zero following the first run of ones, excluding the case $01^{n-2}0$. To complete the proof of (22), we must show that the weight of W'_n is given by

$$w(W'_n) = a_{n-1} + (y-1)a_{n-2} - ya_{n-3}, \qquad n \geq 4. \tag{23}$$

First note that the weight of all binary members of W'_n is given by $(n-3)y^2$, since they are of the form 01^a01^b, where $a, b \geq 1$ and $a+b = n-2$. Next observe that the weight of all members of W'_n in which there is a single 0 after the first run of 1's and containing only two runs of 1's is $y(a_{n-2} - y)$. To see this, write a single zero at the beginning and after the first run of 1's within inc(w) for each $w \in W_{n-2} - \{0^{n-2}\}$. To finish the proof of (23), we must then show that the total weight of all members of W'_n of the form

$$w = 01^a01^b\alpha 1\beta, \qquad a, b \geq 1, \tag{24}$$

where α is non-empty and contains no letters less than two and β is possibly empty, is given by $a_{n-1} - a_{n-2} - ya_{n-3} - (n-4)y^2$. (Note that any letters of β are greater than or equal max(α), lest there be an occurrence of 211.)

To do so, we remove a 1 from the second run of 1's and count members of W_{n-1} of the form in (24) except now $b \geq 0$. Note that members $v \in W_{n-1}$ *not* of this form can be divided into three disjoint classes:

(i) those starting with two or more 0's,

(ii) those starting with one zero and containing a (single) zero after the right-most 1, or

(iii) those of the form 01^c01^d, where $c, d \geq 1$.

The weight of the words in classes (i), (ii) and (iii) is seen to be a_{n-2}, ya_{n-3} and $(n-4)y^2$, respectively. Subtracting from a_{n-1} then gives a weight of $a_{n-1} - a_{n-2} - ya_{n-3} - (n-4)y^2$ for all words in W'_n of the form (24), which completes the proof. □

3.2. The Cases 312 and 213

Here, we consider avoidance of the patterns 312 and 213. We first consider the case 312. By a *primitive* Catalan word, we will mean one having no two consecutive letters the same. It will be more convenient at first to study the primitive members of $W_{312}(n,m)$, the subset of which we will denote by $\mathcal{B}_{n,m}$. If $m \geq 2$, then let $C_{n,m} \subseteq \mathcal{B}_{n,m}$ consist of the members ending in zero. Let $b_{n,m} = |\mathcal{B}_{n,m}|$ if $m \geq 1$ and $c_{n,m} = |C_{n,m}|$ if $m \geq 2$, with $c_{n,1} = 0$ for all n. Note that $b_{n,m}$ and $c_{n,m}$ can assume non-zero values only when $n \geq 1$ and $1 \leq m \leq t$, where $t = \lfloor \frac{n+1}{2} \rfloor$. The arrays $b_{n,m}$ and $c_{n,m}$ are determined by the following intertwined recurrences.

Lemma 3.6. *If $n \geq 3$ and $2 \leq m \leq t$, then*

$$b_{n,m} = c_{n,m} + \sum_{i=2}^{m} \sum_{j=2i-1}^{n-1} c_{j,i}(b_{n-j+1,m-i+1} + b_{n-j,m-i+1}), \tag{25}$$

where $b_{n,1} = \delta_{n,1}$ for all $n \geq 1$, with

$$c_{n,m} = \sum_{i=0}^{n-5} c_{n-2-i,m-1}, \qquad n \geq 5 \quad \text{and} \quad 3 \leq m \leq t, \tag{26}$$

where $c_{n,1} = 0$ for all $n \geq 1$ and $c_{n,2} = \chi(n \text{ is odd})$ for $n \geq 3$.

Proof. The boundary conditions follow easily from the definitions. If $m = 2$, note that $C_{n,2}$ consists of only the word $(01)^{(n-1)/2}0$ if n is odd and is empty if n is even. To show (26), we argue that there are $c_{n-2-i,m-1}$ members of $C_{n,m}$ in which there are exactly $i+2$ letters coming prior to the occurrence of the first 2, whence the result follows from summing over i. Note that $i \leq n-5$ since $m \geq 3$. We define a bijection between $C_{n-2-i,m-1}$ and the subset of $C_{n,m}$ whose members have $i+2$ letters prior to the first 2. Let $\rho \in C_{n-2-i,m-1}$. We consider inserting $i+2$ letters into $\text{inc}(\rho) = p_1p_2\cdots p_{n-2-i}$ so as to produce a member of $C_{n,m}$. Note that $p_1 = p_{n-2-i} = 1$ and $p_2 = 2$. We insert a sequence of the form $0101\cdots$ of length $i+1$ directly before p_1 and a single zero directly after p_{n-2-i} if i is even and insert a sequence of the form $0101\cdots$ of length i directly before p_1 and single zeros directly after p_1 and p_{n-2-i} if i is odd. Observe that no zeros can be inserted in any other positions within $\text{inc}(\rho)$ without introducing an occurrence of 312 and, conversely, given any member of $C_{n,m}$, the positions of the zeros relative to the other letters are as described. Since this operation is reversible, it is the requested bijection, which completes the proof of (26).

For (25), we argue that the sum on the right-hand side gives the cardinality of $\mathcal{B}_{n,m} - C_{n,m}$. Given $2 \leq i \leq m$ and $2i-1 \leq j \leq n-1$, we show that there are $c_{j,i}(b_{n-j,m-i+1} + b_{n-j+1,m-i+1})$ members $\lambda \in \mathcal{B}_{n,m} - C_{n,m}$ in which the right-most 0 of λ occurs at position j, where $i-1$ is the largest letter occurring between the first and last 0. Note that such members are of the form $\lambda = \alpha\beta$, where $\alpha \in C_{j,i}$ and β is non-empty and contains no letters less than $i-1$ (for otherwise, there would be an occurrence of 312 if $i \geq 3$ or the maximality of j would be contradicted if $i = 2$). Furthermore, if β starts with $i-1$, then it constitutes a member of $\mathcal{B}_{n-j,m-i+1}$ on the letters $\{i-1, i, \ldots, m-1\}$, while if β starts with i, then $(i-1)\beta$ constitutes a member of $\mathcal{B}_{n-j+1,m-i+1}$ on the same letters. Thus, there are $c_{j,i}(b_{n-j,m-i+1} + b_{n-j+1,m-i+1})$ members λ, as claimed. Summing over all possible i and j gives (25). $\square$

We can now find explicit formulas for the number of 312-avoiding Catalan words of a given length.

Theorem 3.7. *If $n \geq 3$, then*

$$a_{312}(n,m) = \sum_{i=0}^{n-2m+1} \binom{n-m-i}{m-1}\binom{i+m-2}{m-2} 2^{i+m-2}, \qquad 2 \leq m \leq t, \tag{27}$$

and

$$a_{312}(n) = \frac{3^{n-2}+1}{2}. \tag{28}$$

Proof. Define $C(x,y) = \sum_{m\geq 1} C_m(x)y^m = \sum_{m\geq 1}\sum_{n\geq 2m-1} c_{n,m}x^ny^m$. Multiplying both sides of (26) by x^n, and summing over $n \geq 2m-1$, gives

$$\begin{aligned}C_m(x) &= \sum_{i\geq 0} x^{i+2}\sum_{n\geq i+5} c_{n-2-i,m-1}x^{n-2-i} = \sum_{i\geq 0} x^{i+2}\sum_{n\geq 3} c_{n,m-1}x^n\\ &= \frac{x^2}{1-x}C_{m-1}(x), \qquad m\geq 3,\end{aligned}$$

with $C_1(x) = 0$ and $C_2(x) = \frac{x^3}{1-x^2}$. Multiplying the last recurrence by y^m and summing over $m \geq 3$, we obtain

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

Define $B(x,y) = \sum_{m\geq 1} B_m(x)y^m = \sum_{m\geq 1}\sum_{n\geq 2m-1} b_{n,m}x^ny^m$. Multiplying (25) by x^n and summing over $n \geq 2m-1$, we find

$$B_m(x) = (1+x)C_m(x) + \frac{1+x}{x}\sum_{i=2}^{m-1} C_i(x)B_{m+1-i}(x) = \frac{1+x}{x}\sum_{i=2}^{m} C_i(x)B_{m+1-i}(x), \qquad m\geq 2,$$

where we have used $B_1(x) = x$. Multiplying the last recurrence by y^m and summing over $m \geq 2$, we obtain

$$B(x,y) - xy = \frac{1+x}{xy}C(x,y)B(x,y),$$

which gives

$$B(x,y) = \frac{xy(1-x-x^2y)}{1-x-2x^2y}.$$

Now define $A(x,y) = \sum_{m\geq 1} A_m(x)y^m = \sum_{m\geq 1}\sum_{n\geq 2m-1} a_{312}(n,m)x^ny^m$. Then the generating function $A(x,y)$ is given by

$$\begin{aligned}A(x,y) &= B\left(\frac{x}{1-x}, y\right) = \frac{xy(1-3x+2x^2-x^2y)}{(1-x)(1-3x+2x^2-2x^2y)}\\ &= \frac{xy}{1-x} + \frac{x^3y^2}{(1-x)^2(1-2x)\left(1-\frac{2x^2y}{(1-x)(1-2x)}\right)}\\ &= \frac{xy}{1-x} + \frac{x^3y^2}{(1-x)^2(1-2x)}\sum_{j\geq 0}\left(\frac{2x^2y}{(1-x)(1-2x)}\right)^j.\end{aligned}$$

Extracting the coefficient of y^m yields

$$[y^m]A(x,y) = \frac{2^{m-2}x^{2m-1}}{(1-x)^m(1-2x)^{m-1}}, \qquad m\geq 2,$$

and extracting the coefficient of x^n in this last expression implies (27). Formula (28) follows from the fact that $A(x,1) = \frac{x(1-3x+x^2)}{(1-x)(1-3x)}$. □

We now enumerate members of the avoidance class for the final three letter pattern, namely, 213.

Theorem 3.8. *We have*

$$f(x,y) := \sum_{n\geq 1}\sum_{m=1}^{\lfloor\frac{n+1}{2}\rfloor} a_{213}(n,m)x^ny^m = \frac{xy(1-3x+2x^2+x^3y)}{(1-x)^2(1-2x-x^2y)}. \qquad (29)$$

Thus, if $a_n = a_{213}(n)$*, we have*

$$a_n = 4a_{n-1} - 4a_{n-2} + a_{n-4}, \qquad n \geq 5, \tag{30}$$

with $a_1 = a_2 = 1$, $a_3 = 2$, $a_4 = 5$.

Proof. Note that $\lambda \in W_{213}(n) - \{0^n\}$ if and only if it is of the form for some $i \geq 1$,

$$\lambda = 0^{a_0}1^{a_1} \cdots i^{a_i}\alpha_i\alpha_{i-1}\cdots\alpha_1,$$

where α_j, $j \geq 2$, is either the single letter $j-1$ or has length at least two and is binary in the letters $j-1$ and j, with both the first and last letters $j-1$, and α_1 is a non-empty binary word in 0 and 1 starting with 0. Considering whether or not a member of $W_{213}(n)$ contains at least two distinct letters implies

$$\begin{aligned} f(x,y) &= \frac{xy}{1-x} + \sum_{i\geq1}\left(\frac{xy}{1-x}\right)^{i+1}\left(x+\frac{x^2}{1-2x}\right)^{i-1}\cdot\frac{x}{1-2x} \\ &= \frac{xy}{1-x} + \sum_{i\geq1}\frac{x^{i+2}y^{i+1}}{(1-x)^{i+1}(1-2x)^i}[x(1-x)]^{i-1} = \frac{xy}{1-x} + \frac{xy}{(1-x)^2}\sum_{i\geq1}\left(\frac{x^2y}{1-2x}\right)^i \\ &= \frac{xy}{1-x} + \frac{x^3y^2}{(1-x)^2(1-2x-x^2y)} = \frac{xy(1-3x+2x^2+x^3y)}{(1-x)^2(1-2x-x^2y)}. \end{aligned}$$

Recurrence (30) follows from the $y = 1$ case of formula (29) since

$$f(x,1) = \frac{x(1-3x+2x^2+x^3)}{(1-2x+x^2)(1-2x-x^2)} = \frac{x(1-3x+2x^2+x^3)}{1-4x+4x^2-x^4}. \quad \square$$

References

[1] M. Albert, N. Ruškuc, and V. Vatter, Rational generating functions for 321-avoiding subclasses, in preparation, 2014.

[2] A. T. Benjamin and J. J. Quinn, *Proofs that Really Count: The Art of Combinatorial Proof*, Mathematical Association of America, Washington DC, 2003.

[3] S. Heubach and T. Mansour, *Combinatorics of Compositions and Words*, (Discrete Mathematics and its Applications), Chapman & Hall/CRC, Boca Raton, 2009.

[4] S. Kitaev, *Patterns in Permutations and Words*, Springer-Verlag, New York, 2011.

[5] T. Mansour and M. Shattuck, Avoidance of patterns of type (1,2) by Catalan words, pre-print.

[6] T. Mansour and M. Shattuck, Chebyshev polynomials and statistics on a new collection of words in the Catalan family, *J. Difference Equ. Appl.* **20:11** (2014), 1568–1582.

[7] T. Mansour and A. Vainshtein, Restricted permutations, continued fractions, and Chebyshev polynomials, *Electron. J. Combin.* **7** (2000), #R17.

[8] T. Mansour and A. Vainshtein, Restricted 132-avoiding permutations, *Adv. in Appl. Math.* **26** (2001), 258–269.

[9] T. Rivlin, *Chebyshev Polynomials: From Approximation Theory to Algebra and Number Theory*, Wiley, New York, 1990.

[10] N. J. A. Sloane, On-line Encyclopedia of Integer Sequences, http://oeis.org, 2010.

[11] C. Stump, On a new collection of words in the Catalan family, *J. Integer Seq.* **17** (2014), Art. 14.7.1.

[12] V. Vatter, A family of words enumerated by the Catalan numbers, MO:131585. Available at http://www.mathoverflow.net/questions/131585, 2013.