Two combinatorial statistics on Dyck paths

Alain DENISE* Rodica SIMION†

October 29, 1992

Abstract

Two combinatorial statistics, the pyramid weight and the number of exterior pairs, are investigated on the set of Dyck paths. Explicit formulae are given for the generating functions of Dyck paths of prescribed pyramid weight and prescribed number of exterior pairs. The proofs are combinatorial and rely on the method of q-grammars as well as on two new q-analogues of the Catalan numbers derived from statistics on non-crossing partitions. Connections with the combinatorics of Motzkin paths are pointed out.

1 Introduction

The subject matter of this paper falls in the area of enumeration according to *combinatorial statistics* and q-analogues. In this setting, given a set X of combinatorial objects, one considers a function $s: X \to \mathbf{N}$ and the series $F(X;q) := \sum_{x\in X} q^{s(x)}$. $F(X;q)$ is the generating function of the objects in X enumerated according to the value of the statistic s. When X is finite, $F(X;q)$ is referred to as a q-analogue of the integer $|X| = F(X;1)$. In an analogous manner one may consider several statistics on X simultaneously and their joint distribution, thus obtaining a generating function in several variables and a "multi-q"-analogue.

One motivation for the study of q-analogues (or multi-q-analogues) is that they provide a refined enumeration, according to natural statistics of combinatorial interest. The relationship between different statistics – revealed either by the combinatorics of the objects involved, or by the generating functions – and relations of combinatorial statistics with problems in other fields (such as analysis or physics) contribute to the general interest in q-enumeration.

*LaBRI, Unité associée au CNRS n^o 1304, Université Bordeaux I, 33405 Talence Cedex, France, with partial support from PRC Math-Info.

†Department of Mathematics, The George Washington University, Washington DC 20052, USA. This work was carried out in part during this author's visit at LaBRI, Université Bordeaux I, and with partial support from NSF grant DMS–9108749.

There is an extensive literature on q-enumeration of permutations, integer partitions, set partitions, lattice paths, and many other combinatorial objects. This includes a number of q- and multi-q-analogues of the Catalan numbers, $C_n = \frac{1}{n+1}\binom{2n}{n}$, obtained from statistics defined on various sets of the Catalan family (that is, sets whose cardinalities are given by the numbers in the Catalan sequence $1, 2, 5, 14, 42, 132, \ldots$). By way of a limited number of examples of previous work, we mention the enumeration of binary trees [27]; planar maps [7], [8]; skew Ferrers diagrams counted according to their perimeter [19], area [11], area and number of columns [14]; Dyck words counted according to inversions [4], [5], [12], and according to major index [1], [16]; staircase polyominoes counted according to area and number of columns [12].

In this paper we investigate two related statistics on the set of Dyck paths, which belongs to the Catalan family. These statistics are called the *pyramid weight* and the *number of exterior pairs*, and our main results, Theorems 2.3, 3.1 and 4.1, provide explicit formulae for the joint distributions of the path length and each of these two statistics. Our results are obtained through a combination of enumerative methods: Schützenberger's *DSV method* [21], its recent q-analogue called the *method of q-grammars*, due to M. Delest and J.M. Fédou [9], and new q-analogues of the Catalan numbers resulting from statistics on non-crossing partitions.

Our results settle in the affirmative two conjectures formulated in [12]. As might be expected, the present work has connections with a number of others which feature the Catalan numbers or the method of q-grammars. Very briefly, our results can be reformulated in the language of Motzkin paths which play a central role in Viennot's powerful combinatorial theory of moments of orthogonal polynomials [26], and the questions which we address stem from an overview [12] of the applicability of the method of q-grammars to the distribution of a variety of statistics on Dyck paths, some of which coincide with previously studied q-analogues of Catalan numbers [4], [5], [16].

This paper is organized as follows. Section 2 contains the definitions and notation used throughout the paper, together with a brief background on the key tools used in our proofs: the method of q-grammars and the notion of non-crossing partition. This section also includes Theorem 2.3, whose proof illustrates the usefulness of the method of q-grammars and constitutes the point of departure for later results. Sections 3 and 4 focus on the main results, Theorems 3.1 and 4.1, and several combinatorial consequences derived from these.

2 Definitions, notation, and preliminary results

2.1 Dyck paths, pyramid weight, and exterior pairs

A *Dyck path* is a path in the first quadrant, which begins at the origin, ends at $(2n, 0)$, and consists of steps $(+1, +1)$ (North-East) and $(+1, -1)$ (South-East).

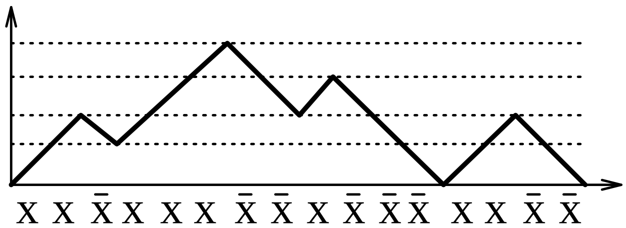

Figure 2.1: Example of a Dyck path and its associated Dyck word

We will refer to n as the *length* of the path, denoted $l(p)$.

Let D be the set of all Dyck paths and D_l the set of Dyck paths whose length is l. It is well-known (see, for example, [6]) that $|D_l| = C_l = \frac{1}{l+1}\binom{2l}{l}$, the l^{th} Catalan number, and thus

$$\sum_{l \geq 0} |D_l| t^l = \frac{1-\sqrt{1-4t}}{2t}.$$

Among the many sets of combinatorial objects whose cardinalities are also given by the Catalan numbers, the most classical example is that of complete parenthesis systems [6]. North-East and South-East steps in a Dyck path correspond to left and right parentheses, respectively. Equivalently, we may encode each North-East step by a letter x and each South-East step by a letter $\bar{x}$, thus obtaining the frequently used encoding of Dyck paths by *Dyck words*. Figure 2.1 shows an example of a Dyck path and its associated Dyck word. In the sequel, we will at times use interchangeably a Dyck path p and its Dyck word w. If w is a word on the alphabet $\{x, \bar{x}\}$, the symbol $|w|_x$ ($|w|_{\bar{x}}$, respectively) stands for the number of occurrences of the letter x ($\bar{x}$, respectively) in w. Obviously, if p is a Dyck path with associated Dyck word w, then $|w|_x = |w|_{\bar{x}} = l(p)$.

A non-empty Dyck word w is called *primitive* if $|w'|_x > |w'|_{\bar{x}}$ for every factorization $w = w'w''$ with w' and w'' non-empty. In terms of paths, primitivity means that the only intersections of the path with the x-axis are the initial and final points of the path. The notion of primitive Dyck path is helpful in establishing grammar rules for the formal language of Dyck words, which permit – in turn – the derivation of generating functions. More on this will appear in our discussion of the DSV method and subsequent sections.

We will be interested in two statistics on Dyck paths, whose definitions require the notion of *pyramid*. This was introduced and used by Vauchaussade de Chaumont and Viennot [25] to solve an enumeration problem in molecular biology. It will be obvious that this is a geometrically natural statistic for Dyck paths, though its definition is most conveniently stated in terms of Dyck words.

Definitions 2.1 *A* pyramid *in a Dyck word is a factor of the form $x^h\bar{x}^h$. We*

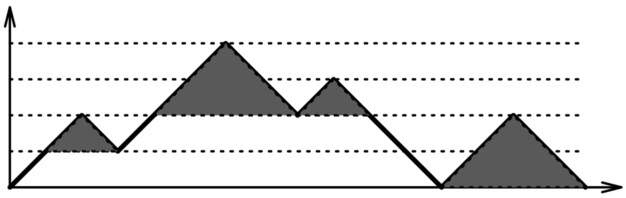

Figure 2.2: Maximal pyramids and exterior pairs in a Dyck path

refer to h as the height *of the pyramid.*
A pyramid in a Dyck word w is maximal *if, as a factor in w, it is not immediately preceded by an x and immediately followed by an $\bar{x}$.*
The pyramid weight *of a Dyck path p is the sum of the heights of its maximal pyramids, and is denoted $P(p)$.*

Definitions 2.2 *An* exterior pair *in p is a pair consisting of an x and its matching $\bar{x}$ (when viewed as parentheses) which do not belong in any pyramid. The* number of exterior pairs *of a Dyck path p is denoted by $E(p)$.*

Clearly, the sum of $P(p)$ and $E(p)$ equals the length of the path p. Figure 2.2 illustrates the maximal pyramids and exterior pairs in a Dyck path for which $l(p) = 8$, $P(p) = 6$ and $E(p) = 2$.

The notion of pyramid can be viewed in a different context. If we regard x and $\bar{x}$ as inverse elements in a group, then a Dyck word w is equal to the identity. Finding and eliminating the maximal pyramids constitutes then one stage in the reduction of w to the identity. The resulting shorter word is Dyck as well, and corresponds to the path obtained by erasing the plateaux created by the elimination of maximal pyramids from the path of w. Vauchaussade and Viennot [25] defined the "order of a Dyck word" w as the number of stages required for its complete simplification, a combinatorial statistic related to the biologists' notion of order of a secondary RNA structure. In [25], Dyck words are enumerated according to the "order" statistic giving rise to Strahler numbers, and their generating function is expressed in terms of Chebyshev polynomials of the second kind. The distribution of the number of exterior pairs is expressible in terms of Chebyshev polynomials as well, in a manner which appears in [12]. We include a brief discussion of this connection with Chebyshev polynomials following the proof of Lemma 4.3.

In yet another setting, pyramids and pyramid weight appear in Zeilberger's article [29]. A maximal pyramid in a Dyck word corresponds bijectively to a maximal hanging branch in an ordered tree. In fact, the set of Dyck words of length l and pyramid weight k is in bijective correspondence with the set of ordered trees on l vertices and in which the sum of the lengths of the maximal hanging branches is k.

Based on the pyramid weight and number of exterior pairs, we have two refinements of the enumeration of Dyck paths:

$$F(q,t){:=}\sum_{p\in D} q^{P(p)}t^{l(p)} \quad \text{and} \quad G(q,t){:=}\sum_{p\in D} q^{E(p)}t^{l(p)}.$$

Our goal is to establish explicit formulae for the generating functions of Dyck paths counted according to length for fixed pyramid weight and, in turn, fixed number of exterior pairs. That is, we aim to find explicit expressions for the coefficients of q^k in $F(q,t)$ and $G(q,t)$,

$$f_k(t){:=}\sum_{l\geq 0} |\{p\in D_l{:}\, P(p)=k\}|t^l$$

and

$$g_k(t){:=}\sum_{l\geq 0} |\{p\in D_l{:}\, E(p)=k\}|t^l.$$

These are the generating functions for the sequences of numbers in the columns of the tables in Figures 2.3 and 2.4.

We will take advantage of explicit formulae for $F(q,t)$ and $G(q,t)$ which were obtained in [12] through an application of the method of q-grammars. In the interest of self-containment, these are presented in the next subsection.

2.2 The DSV and q-grammar methods

The *DSV method*, introduced by M.P. Schützenberger [21], uses theoretical computer science notions related to *algebraic* or *context-free languages*. Let $X=\{x_1,x_2,\ldots,x_n\}$ be an *alphabet* and X^* be the free monoid generated by X; thus, X^* is the set of all finite-length words over X, including the empty word. A *language* over the alphabet X is a subset L of X^*. The *generating function in non-commutative variables* of the language L, denoted $\underline{L}$, is the formal sum of all words in L,

$$\underline{L}=\sum_{w\in L} w.$$

Let α be the homomorphism which permutes the letters in X. Then $\mathcal{L}=\alpha(\underline{L})$ is the *generating function* of the language L, with the words enumerated according to their length,

$$\mathcal{L}(x_1,x_2,\ldots,x_n)=\sum_{i_1,i_2,\ldots,i_n\geq 0} \lambda_{i_1,i_2,\ldots,i_n} x_1^{i_1}x_2^{i_2}\ldots x_n^{i_n},$$

where $\lambda_{i_1,i_2,\ldots,i_n}$ is the number of words in L consisting of i_1 letters x_1, i_2 letters x_2, …, i_n letters x_n. The DSV method allows the derivation of this generating function starting from an *unambiguous grammar* for L, if L is context-free.

Consider the language D of Dyck words, which provides a simple preliminary illustration of the DSV method. The language D is generated by the context-free grammar

$$D \to \epsilon + xD\bar{x}D,$$

where ϵ is the empty word, and which gives in a sense a recursive definition of the language: every word is constructible by successive applications of the grammar rules. For example, the word $x\bar{x}xx\bar{x}\bar{x}$ is generated as follows:

$$D \to xD\bar{x}D \to x\bar{x}D \to x\bar{x}xD\bar{x}D \to x\bar{x}xxD\bar{x}D\bar{x}D$$

$$\to x\bar{x}xx\bar{x}D\bar{x}D \to x\bar{x}xx\bar{x}\bar{x}D \to x\bar{x}xx\bar{x}\bar{x}.$$

Since each word can be derived in only one way, the grammar is unambiguous.

From the grammar we deduce an equation (in general, a system of equations) for the generating function in noncommutative variables, $\underline{D}$, of the language D:

$$\underline{D} = 1 + x\underline{D}\bar{x}\underline{D}.$$

Hence, the generating function $\mathcal{D}$ of the language is a solution of

$$\mathcal{D}(x, \bar{x}) = 1 + x\bar{x}\mathcal{D}^2.$$

If we let $d(t) := \mathcal{D}(t, 1)$, then

$$d(t) = 1 + td^2$$

and we obtain

$$d(t) = \frac{1 - \sqrt{1 - 4t}}{2t}$$

whose expansion around $t = 0$ gives the Catalan numbers as coefficients.

The DSV method featured in Cori's enumeration of planar maps [7], and served to solve a number of combinatorial problems (see, e.g., [25] [11]). Viennot presents in [27] an overview of applications of the DSV method as of 1985, and a survey of more recent results on the enumeration of polyominoes in [28].

The method of *q-grammars*, formalized by Delest et Fédou [9] [10], comes into play when one wants to enumerate the words of a context-free language L according to length *and* a second combinatorial statistic.

A q-grammar is defined in terms of the notion of *attribute*, introduced by Knuth [17]. To each word w in a context-free language L, we associate a monomial $\varphi(w) \in (X \cup \{q\})^*$ in which the number of occurrences of q encodes the value of a combinatorial statistic of interest on L. For example, let $L = \{a, b\}^*$, where we consider $a < b$, and the following problem: enumerate the words in L according to length and the number of inversions. An inversion is a subword ba and, for instance, the word $w = ababbab$ has 4 inversions.

We can proceed by defining the unambiguous grammar $L \to \epsilon + aL + bL$ which generates L, and

$$\begin{aligned}&\varphi(\epsilon) = 1\\ &\varphi(aw) = a\varphi(w)\\ &\varphi(bw) = bq^{|w|_a}\varphi(w).\end{aligned}$$

Thus, $\varphi(w) = wq^{inv(w)}$. With a and b playing the role of variables or alphabet elements clear from the context, we define the q-generating function

$$l(q, a, b) := \sum_{w \in L} a^{|w|_a} b^{|w|_b} q^{inv(w)}.$$

It is shown in [10] that from the grammar for L and φ, we can obtain the equation

$$l(q, a, b) = 1 + al(q, a, b) + bl(q, aq, b),$$

whence

$$\begin{aligned} l(q, a, b) &= \frac{1}{1-a} + \frac{b}{1-a} l(q, aq, b) \\ &= \sum_{n \geq 0} \frac{b^n}{(1-a)(1-aq)(1-aq^2)\cdots(1-aq^n)}. \end{aligned}$$

We can then obtain $f_L(q, t) := l(q, t, t)$, the generating function of the words in L counted by length and number of inversions. One easily recognizes that $f_L(q, t)$ is the generating function for Ferrers diagrams counted by number of rows and number of cells (details appear in [10]).

With the aid of the method of q-grammars, Fédou [14], [15] carried out the enumeration of certain polyominoes in terms of q-analogues of Bessel functions, while Bousquet-Mélou [2], [3] solved a previously open problem in the enumeration of convex polyominoes.

We now turn to the following Theorem which is the starting point of our subsequent results. Its proof uses the method of q-grammars.

Theorem 2.3 *Let $F = F(q, t)$ be the generating function for Dyck paths counted according to the pyramid weight and length, $F(q, t) = \sum_{p \in D} q^{P(p)} t^{l(p)}$. Then $F(q, t)$ satisfies the equation*

$$t(1 - tq)F^2 - (1 + t - 2tq)F + (1 - tq) = 0,$$

hence,

$$F(q, t) = \frac{1 + t - 2qt - \sqrt{(1-4t)(1-qt)^2 + t(1-q)(2+t-3qt)}}{2t(1-qt)}.$$

If $G(q, t)$ is the generating function for Dyck paths counted according to the number of exterior pairs and length, then we also have

$$G(q, t) = \frac{1 + qt - 2t - \sqrt{(1-4qt)(1-t)^2 + t(q-1)(2+qt-3t)}}{2qt(1-t)}.$$

$P(w)$ / $l(w)$	0	1	2	3	4	5	6	7	8	9	10
0	1										
1	0	1									
2	0	0	2								
3	0	0	1	4							
4	0	0	1	5	8						
5	0	0	1	7	18	16					
6	0	0	1	9	34	56	32				
7	0	0	1	11	55	138	160	64			
8	0	0	1	13	81	275	500	432	128		
9	0	0	1	15	112	481	1205	1672	1120	256	
10	0	0	1	17	148	770	2471	4797	5264	2816	512

Figure 2.3: Enumeration of Dyck paths according to length and pyramid weight

$E(w)$ / $l(w)$	0	1	2	3	4	5	6	7	8
0	1								
1	1								
2	2								
3	4	1							
4	8	5	1						
5	16	18	7	1					
6	32	56	34	9	1				
7	64	160	138	55	11	1			
8	128	432	500	275	81	13	1		
9	256	1120	1672	1205	481	112	15	1	
10	512	2816	5264	4797	2471	770	148	17	1

Figure 2.4: Enumeration of Dyck paths according to length and number of exterior pairs

Proof. To prove this result, we begin with a q-grammar which will lead to the generating function $F(q,t) = \sum_{p\in D} q^{P(p)} t^{l(p)}$ of Dyck paths counted by pyramid weight and length.

To this end we consider four interrelated formal languages over the alphabet $\{x, \bar{x}\}$: D, the language of all Dyck words; $D^+ = D - \{\epsilon\}$, the language of all nonempty Dyck words; $\Pi := \{x^h \bar{x}^h : h > 0\}$, the language of nonempty pyramids; and $N := D^+ - \Pi$, the language of all nonempty Dyck words which are not pyramids.

With this notation, we have the following unambiguous algebraic grammar which generates the Dyck words:

$$\begin{aligned}
&D \to \epsilon + D^+\\
&D^+ \to \Pi + N\\
&N \to xN\bar{x}D + \Pi D^+\\
&\Pi \to x\bar{x} + x\Pi\bar{x}.
\end{aligned}$$

We now define the mapping φ on Dyck words as $\varphi(w) = q^k$ if w has pyramid weight $P(w) = k$. Thus,

$$\begin{array}{lll}
\varphi(\epsilon) = 1 & & \\
\varphi(xu\bar{x}v) = \varphi(u)\varphi(v) & \text{if} & u \in N, v \in D\\
\varphi(uv) = \varphi(u)\varphi(v) & \text{if} & u \in \Pi, v \in D^+\\
\varphi(x\bar{x}) = q & & \\
\varphi(xu\bar{x}) = q\varphi(u) & \text{if} & u \in \Pi.
\end{array}$$

With the aid of the resulting q-grammar which generates the Dyck language *with* the pyramid weight as an attribute, we obtain relations among the generating functions F_D, F_{D^+}, F_Π, and F_N of the words in the languages D, D^+, Π, and N, respectively. Each of these generating functions is a function of three variables, $x, \bar{x}, q$, and a Dyck word w contributes the monomial $wq^{P(w)}$. We have:

$$\begin{aligned}
F_D &= 1 + F_{D^+} & &\\
F_{D^+} &= F_\Pi + F_N & &\\
F_\Pi &= x\bar{x}q + x\bar{x}qF_\Pi & &= \tfrac{x\bar{x}q}{1 - x\bar{x}q}\\
F_N &= x\bar{x}F_N F_D + F_\Pi F_{D^+} & &= \tfrac{x\bar{x}qF_{D^+}}{(1 - x\bar{x}q)(1 - x\bar{x} - x\bar{x}F_{D^+})}.
\end{aligned}$$

Consequently,

$$F_{D^+} = \frac{x\bar{x}q}{1 - x\bar{x}q} + \frac{x\bar{x}qF_{D^+}}{(1 - x\bar{x}q)(1 - x\bar{x} - x\bar{x}F_{D^+})}$$

and the desired formula follows from

$$x\bar{x}(1 - x\bar{x}q)F_{D^+}^2 - (1 - x\bar{x})(1 - 2x\bar{x}q)F_{D^+} + x\bar{x}q(1 - x\bar{x}) = 0$$

by using $F_{D^+} = F_D - 1$ and setting $x = t$, $\bar{x} = 1$ in $F_D(x, \bar{x})$ to obtain $F(q,t)$ as claimed.

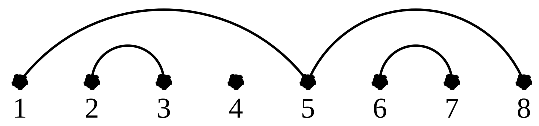

Figure 2.5: The graphical representation of the non-crossing partition 1 5 8 / 2 3 / 4 / 6 7

The formula for $G(q,t)$ follows from the remark that for each Dyck path p, we have $P(p) + E(p) = l(p)$. This implies that $G(q,t) = F(q^{-1}, qt)$, so the formula for $G(q,t)$ follows. □

2.3 Non-crossing partitions

Our formulae for $f_k(t)$ and $g_k(t)$ will be proved combinatorially, using non-crossing partitions.

A partition of the set $[n]:= \{1, 2, \ldots, n\}$ is, as usual, a collection of non-empty, pairwise disjoint subsets B_i called *blocks*, whose union is $[n]$. We can write a partition into k blocks as $B_1/B_2/\ldots/B_k$, with the blocks indexed in increasing order of their minima.

We will be interested in partitions which enjoy an additional property.

Definition 2.4 *A partition of* $[n]$ *is called* non-crossing *if for every four elements* $1 \leq a < b < c < d \leq n$*, the following condition is satisfied: if* a *and* c *lie in the same block, and* b *and* d *lie in the same block, then all four elements lie in the same block.*

For example, the partition 1 5 8 / 2 3 / 4 / 6 7 of [8] is non-crossing, while in the partition 1 5 8 / 2 3 /4 6 7 the blocks B_1 and B_3 cross.

It is often convenient to represent a non-crossing partition by a diagram such as in Figure 2.5, in which the integers 1 through n are placed as on the real axis and two successive elements of the same block are joined by and arc which runs above the axis. The non-crossing property is then reflected in the fact that such a representation is possible with no crossing between arcs.

The set $NC(n)$ of all non-crossing partitions of $[n]$ has cardinality C_n. Let us remark that $NC(n)$ has been investigated as a partially ordered set in [13], [18], [23], and that q-analogues of Catalan and Narayana numbers were obtained from set-partition statistics [22]. Non-crossing partitions arise as well in the context of hypermaps, as maps with one face having genus zero [7].

In the proofs of our main results, we will use non-crossing partitions as a convenient alternate combinatorial model for the Catalan numbers. The non-crossing partitions will be suitably weighted and, on one occasion, colored according to certain rules. To facilitate the exposition, we will refer to blocks

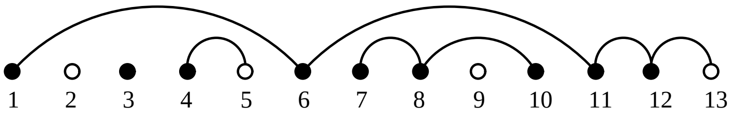

Figure 2.6: A non crossing partition with the “filler” points indicated as white dots

whose cardinality is 1 as *singletons*, and we call *trivial* the partition all of whose blocks are singletons. The following additional terminology will be used in the proof of Lemma 3.3.

Definition 2.5 *Let $\pi \in NC(n)$, $n > 1$. A point $i \in \{2, 3, \ldots, n\}$ is called a* filler *of π if either (i) $i-1$ and i are in the same block and i is the largest element of its block, or (ii) i forms a singleton block and $i-1$ is not the largest element in its block.*

For example, if $\pi = 1\ 6\ 11\ 12\ 13/2/3/4\ 5/7\ 8\ 10/9$, then π has 4 fillers, namely $2, 5, 9, 13$, as shown in Figure 2.6

3 Dyck words counted according to the pyramid weight

This section is devoted to the derivation of an explicit formula for $f_k(t)$, followed by consequences of the result and its method of proof.

Theorem 3.1 *Let $f_k(t)$ be the generating function for Dyck paths of pyramid weight k, counted according to their length. Then $f_0(t) = 1$ and*

$$f_k(t) = \frac{t^k}{(1-t)^{k-1}} \left(\frac{1-(1-t)^k}{t} + \sum_{j \geq 1} (-1)^{j-1} \frac{1-(1-t)^j}{t} \binom{k-j-1}{j} C_{k-j-1} \right)$$

if $k > 1$.

The proof will be carried out in two stages, presented in lemmas 3.2 and 3.3. The first lemma establishes, by means of the DSV method, that the series $f_k(t)$ is a rational function and gives the recurrence relation whose solution is the polynomial $P_k(t) := f_k(t)(1-t)^{k-1}/t^k$. The second lemma establishes the desired formula for $P_k(t)$ through a combinatorial proof resorting to an interpretation of $P_k(t)$ as the enumerator of weighted non-crossing partitions.

For simplicity in notation we will suppress, on occasion, the argument t of the functions f_k and P_k.

$P_1(t) =$	1							
$P_2(t) =$	2	$-t$						
$P_3(t) =$	4	$-3t$	$+t^2$					
$P_4(t) =$	8	$-6t$	$+4t^2$	$-t^3$				
$P_5(t) =$	16	$-8t$	$+10t^2$	$-5t^3$	$+t^4$			
$P_6(t) =$	32		$+20t^2$	$-15t^3$	$+6t^4$	$-t^5$		
$P_7(t) =$	64	$+48t$	$+40t^2$	$-35t^3$	$+21t^4$	$-7t^5$	$+t^6$	
$P_8(t) =$	128	$+224t$	$+112t^2$	$-70t^3$	$+56t^4$	$-28t^5$	$+8t^6$	$-t^7$

Figure 3.7: Table of the polynomials $P_k(t)$ for $1 \leq k \leq 8$

Lemma 3.2

$$\begin{cases} f_0 = 1 \\ f_k = \dfrac{t^k}{(1-t)^{k-1}} P_k \quad \text{if} \quad k \geq 1, \end{cases}$$

where P_k is the polynomial given by the recurrence relation

$$\begin{cases} P_1 = 1 \\ P_k = (1-t)^{k-1} + \displaystyle\sum_{i=1}^{k-1} \left((1-t)^i P_{k-i} + tP_iP_{k-i}\right) \quad \text{if} \quad k > 1. \end{cases}$$

Proof. It is easy to formulate rules for an unambiguous context-free grammar generating the formal language consisting of the Dyck words of pyramid weight k. We have already described such a grammar in the Subsection 2.2, but here it is more convenient to formulate grammar rules invoking the primitive Dyck words rather than pyramid words as was done earlier. From this grammar one then obtains, by the DSV method, the following system of equations

$$\begin{cases} f_0 = 1 \\ f_k = E_k + \displaystyle\sum_{i=1}^{k-1} E_i f_{k-i} \quad \text{if} \quad k \geq 1 \\ E_k = t^k + t(f_k - t^k) \quad \text{if} \quad k \geq 1, \end{cases}$$

where $E_k(t)$ is the generating function for primitive Dyck paths of pyramid weight k, counted by length. The reader will recognize the source of these equations: the second equation describes the Dyck paths of pyramid weight $k > 0$ via their unique factorization $p = p'p''$ with p' being the maximal left primitive factor of p, and uses the observation that $P(p) = P(p') + P(p'')$; the third equation comes from separating the path which is one single pyramid from the other primitive paths under considerations and which are in bijective correspondence with arbitrary paths of length one unit less, via the pyramid weight preserving addition of an initial North-East step and a final South-East step.

The substitution of the third equation in the second one leads to

$$\begin{cases} f_0 = 1 \\ \dfrac{(1-t)^{k-1}}{t^k} f_k \;\; = (1-t)^{k-1} \\ \qquad\qquad + \displaystyle\sum_{i=1}^{k-1} \Big[\frac{(1-t)^{k-1}}{t^{k-i}} f_{k-i} + \frac{(1-t)^{k-2}}{t^{k-1}} f_i f_{k-i} \Big] \text{ if } k \geq 1, \end{cases}$$

and the desired conclusion now follows by induction on k. □

The proof of Theorem 3.1 will be completed by the next lemma.

Lemma 3.3 *For every $k > 0$, the polynomial $P_k(t)$ is given by the formula*

$$P_k(t) = \frac{1-(1-t)^k}{t} + \sum_{j \geq 1} (-1)^{j-1} \frac{1-(1-t)^j}{t} \binom{k-j-1}{j} C_{k-j-1}.$$

Proof. The plan of the proof is the following: first we will give a combinatorial interpretation of the claimed formula as the enumerator of non-crossing partitions weighted by a certain weight function; then we will prove, with the aid of this interpretation, that the formula satisfies the same recurrence as $P_k(t)$ was shown to satisfy in Lemma 3.2.

Let

$$P_k^*(t) := \frac{1-(1-t)^k}{t} + \sum_{j \geq 1} (-1)^{j-1} \frac{1-(1-t)^j}{t} \binom{k-j-1}{j} C_{k-j-1}$$

and define a weight function on $NC(k-1)$ as follows:

$$w(\pi) := \begin{cases} \frac{1-(1-t)^k}{t} & \text{if } \pi = 1/2/\ldots/k-1; \\ t^{m(\pi)-1} & \text{otherwise,} \end{cases}$$

where $m(\pi)$ is the number of "fillers" of π (recall Definition 2.5 from Section 2). For example, if $\pi = 1\ 6\ 11\ 12\ 13/2/3/4\ 5/7\ 8\ 10/9$, then $m(\pi) = 4$, as illustrated after the definition 2.5, and $w(\pi) = t^3$.

We claim that

$$P_k^*(t) = \sum_{\pi \in NC(k-1)} w(\pi).$$

To prove our claim, notice that the sum in the expression defining $P_k^*(t)$ lends itself naturally to the following preliminary combinatorial interpretation: starting from an arbitrary non-crossing partition on $k-j-1$ elements (counted by the Catalan number), we construct a non-crossing partition on $[k-1]$ by inserting j additional points in any of $\binom{k-j-1}{j}$ ways. We interpret this binomial coefficient as the number of insertions of the j points so that each new point follows immediately after a point of the original partition on $k-j-1$ points. Each of the j additional points will become a "filler" according to the following rule: if a new point is added immediately after the largest element of a block, then this new point is adjoined to that block; otherwise, the new point constitutes a singleton block. This produces a non-crossing partition on $[k-1]$ with a positive number of "fillers," in which $j>0$ of the fillers are distinguished. The expression defining $P_k^*(t)$ suggests assigning to such an object the weight $(-1)^{j-1}\frac{1-(1-t)^j}{t}$. Now, a partition $\pi \in NC(k-1)$ having $m(\pi)>0$ "filler" points will arise a total of $2^{m(\pi)}-1$ times, once for each choice of $j>0$ distinguished "fillers." A simple calculation shows that the total weight of the $2^{m(\pi)}-1$ partitions associated with π, namely $\sum_{j=1}^{m(\pi)}\binom{m(\pi)}{j}(-1)^{j-1}\frac{1-(1-t)^j}{t}$, is just $t^{m(\pi)-1}$. A special case is the trivial partition of $[k-1]$. This partition is assigned weight $(1-(1-t)^k)/t$ and it is easy to check that it is the only non-crossing partition for which $m=0$.

Hence, $P_k^*(t)=\sum_{\pi\in NC(k-1)} w(\pi)$ as claimed. From this interpretation of $P_k^*(t)$, it is obvious that $P_1^*(t)=1=P_1(t)$ and we will now prove that

$$P_k^*(t)=(1-t)^{k-1}+P_{k-1}^*(t)+\sum_{i=2}^{k-1}((1-t)^i+tP_i^*(t))P_{k-i}^*(t),$$

that is, $P_k^*(t)$ and $P_k(t)$ satisfy the same recurrence relation.

Among the partitions in $NC(k-1)$, those in which 1 constitutes a singleton block have a total weight of $(1-t)^{k-1}+P_{k-1}^*(t)$, because $w(1/2/3/\ldots/k-1)=(1-t)^{k-1}+w(2/3/\ldots/k-1)$, while $w(\pi)=w(\pi')$ if $\pi=1/\pi'$ and $\pi'\neq 2/3/\ldots/k-1$. Thus, it remains to show that

$$\sum_{\substack{\pi\in NC(k-1)\\ 1 \text{ not singleton}}} w(\pi)=\sum_{i=2}^{k-1}((1-t)^i+tP_i^*)P_{k-i}^*,$$

or, equivalently, that

$$\sum_{\substack{\pi\in NC(k-1)\\ 1 \text{ not singleton}}} w(\pi)=\sum_{i=1}^{k-2}P_i^*((1-t)^{k-i}+tP_{k-i}^*).$$

Observe that $(1-t)^{k-i}+tP_{k-i}^*(t)=\sum_{\sigma\in NC(k-i-1)} t^{m(\sigma)}$. This observation leads us to define a modification, $\overline{w}$, of the weight w on non-crossing partitions, namely $\overline{w}(\pi):=t^{m(\sigma)}$. Thus, we must show that

$$\sum_{\substack{\pi\in NC(k-1)\\ 1 \text{ not singleton}}} w(\pi)=\sum_{i=1}^{k-2}\Big(\sum_{\alpha\in NC(i-1)} w(\alpha)\Big)\Big(\sum_{\beta\in NC(k-i-1)}\overline{w}(\beta)\Big).$$

Ideally, one would like to argue that a partition π counted by the left hand side is constructed from a pair of partitions α and β from the right hand side, where $i+1$ is the smallest element (other than 1) which lies in the same block of π as 1, and that $w(\pi) = w(\alpha)\overline{w}(\beta)$. The first part of this statement is true, but there are partitions π for which the desired relation on weights does not hold. Therefore, we will examine the error $\delta(\pi) := w(\pi) - w(\alpha)\overline{w}(\beta)$ as α and β run over their ranges of summation, and we will complete the proof of the Lemma by showing that the total error is zero.

First we consider the case $i = 1$. In this case, $\left(\sum w(\alpha)\right)\left(\sum \overline{w}(\beta)\right) = P_1^*(t)((1-t)^{k-1} + tP_{k-1}^*(t))$. On the other hand, the total weight of the partitions π in this case is $\sum w(\pi) = [t(P_{k-2}^*(t) - \frac{1-(1-t)^{k-2}}{t}) + 1] + [P_{k-1}^*(t) - (P_{k-2}^*(t) + (1-t)^{k-2})]$. The expression in the first bracket corresponds to those π in which the elements 1 and 2 form a block of cardinality 2, thereby making the element 2 contribute to $m(\pi)$; note also that if the partition induced on $\{3, 4, \ldots, k-1\}$ has only singleton blocks, then the total weight of π is equal to 1. Similarly, the expression in the second bracket corresponds to those π under consideration in which the block containing 1 and 2 has cardinality 3 or more. In this case, $w(\pi) = w(\beta)$, but 2 must not form a singleton block in β. All β's in $NC(k-2)$ in which 2 *is* a singleton block have total w-weight equal to $P_{k-2}^*(t) + (1-t)^{k-2}$, and thus we subtract this quantity from $P_{k-1}^*(t)$.

Therefore, the total error over the partitions π with $i = 1$ is $[t(P_{k-2}^*(t) - \frac{1-(1-t)^{k-2}}{t}) + 1] + [P_{k-1}^*(t) - (P_{k-2}^*(t) + (1-t)^{k-2})] - [P_1^*(t)((1-t)^{k-1} + tP_{k-1}^*)] = (1-t)(P_{k-1}^*(t) - P_{k-2}^*(t) - (1-t)^{k-2})$.

When $i \geq 2$, the error $\delta(\pi) = w(\pi) - w(\alpha)\overline{w}(\beta) = 0$ if 2 is not a singleton in α and β is arbitrary. However, $\delta(\pi) \neq 0$ in three remaining situations. We look at each in turn.

When $\alpha = 2/3/\ldots/i$ and $\beta = i+1/i+2/\ldots/k-1$, then $\delta(\pi) = 1 - [(1-(1-t)^i)/t] \cdot 1$.

A partition π under consideration for which $\alpha = 2/3/\ldots/i$ and $\beta \neq i+1/i+2/\ldots/k-1$, has weight $w(\pi) = tw(\beta) = \overline{w}(\beta)$. Hence, in this case, $\sum \delta(\pi) = [tP_{k-i}^*(t) + (1-t)^{k-i} - 1] - \frac{1-(1-t)^i}{t}[tP_{k-i}^*(t) + (1-t)^{k-i} - 1]$.

When $\alpha \neq 2/3/\ldots/i$ but has 2 as a singleton block, we have $w(\pi) = tw(\alpha)\overline{w}(\beta)$ because 2 contributes to the weight of π though it does not contribute to the weight of α. The total error in this case is $\sum \delta(\pi) = (t-1)[P_{i-1}^*(t) - \frac{1-(1-t)^{i-1}}{t}][tP_{k-i}^*(t) + (1-t)^{k-i}]$.

A routine calculation simplifies the sum of the errors from all the different cases to

$$\sum_{\pi \in NC(k-1)} \delta(\pi) = (1-t)\left[P_{k-1}^*(t) - (1-t)^{k-2} - \sum_{j=1}^{k-2} (tP_j^*(t) + (1-t)^j)P_{k-1-j}^*(t)\right].$$

Inductively, this completes the proof since the expression in the bracket is zero by the recurrence relation for $P_{k-1}^*(t)$. $\square$

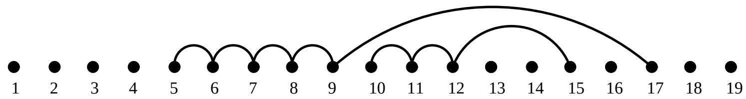

Figure 3.8: Illustration to the proof of Remark 3.4: the non-crossing partition with precisely one "filler"point, corresponding to the subset $\{5, 10, 13, 15, 17\}$.

A different proof of the fact that the polynomials satisfying the recurrence of Lemma 3.2 are given by the expression of Lemma 3.3 was found by F.W. Schmidt [20] by means of generating functions. Schmidt's proof uses the fact stated below as Remark 3.7-b, which we include among other consequences of our combinatorial interpretation of the polynomials $P_k(t)$ as $\sum_{\pi\in NC(k-1)} w(\pi)$. The other remarks, suggested by the table of Figure 3.7 which shows the first few polynomials $P_k(t)$, are made transparent by our combinatorial model.

Remark 3.4 *For each $k > 0$, we have $P_k(0) = 2^{k-1}$.*

Upon setting $t = 0$, the weight w will vanish on all non-crossing partitions except for the trivial partition $1/2/\ldots/k-1$, whose weight becomes k, and for those partitions having $m = 1$, whose weight becomes 1. Thus, $P_k(0) = k + \#\{\pi \in NC(k-1)\colon m(\pi) = 1\}$. The non-crossing partitions of $[k-1]$ which have exactly one "filler" (i.e., $m = 1$) are in bijective correspondence with the subsets of $[k-1]$ whose cardinality is 2 or more.

To describe the correspondence, consider an arbitrary such subset $a_1 < a_2 < \ldots < a_j$. In the corresponding partition, each element $p < a_1$ or $p > a_j$ will form a singleton block. If $j = 2$, then we form one more block, namely $\{a_1, a_1+1, \ldots, a_2\}$. If $j = 3$, then we form the block $\{a_1, a_1+1, \ldots, a_2-1, a_3\}$ and let each of the other elements form a singleton block. If $j > 3$, then a block $\{a_1, a_1+1, \ldots, a_2-1, a_j\}$ is formed as in the case $j = 3$, and the construction is completed recursively by constructing a non-crossing partition of $\{a_2, a_2+1, \ldots, a_j-1\}$ based on the subset $a_2 < a_3 < \ldots < a_{j-1}$ (whose cardinality is still 2 or more).

For example, let $k - 1 = 19$. The partition shown in Figure 3.8,

$$1 \;/\; 2 \;/\; 3 \;/\; 4 \;/\; 5\;6\;7\;8\;9\;17 \;/\; 10\;11\;12\;15 \;/\; 13 \;/\; 14 \;/\; 16 \;/\; 18 \;/\; 19$$

corresponds to the subset $\{5, 10, 13, 15, 17\}$, and the partition

$$1 \;/\; 2 \;/\; 3 \;/\; 4 \;/\; 5\;6\;7\;8\;9\;17 \;/\; 10\;11\;12\;13\;14\;15 \;/\; 16 \;/\; 18 \;/\; 19$$

corresponds to the subset $\{5, 10, 15, 17\}$.

It is easy to see that the partitions arising from this construction are all the partitions having $m = 1$ and that the correspondence is a bijection. Therefore, $\#\{\pi \in NC(k-1)\colon\ m(\pi) = 1\} = \sum_{j=2}^{k-1} \binom{k-1}{j} = 2^{k-1} - k$.

Remark 3.5 *For each $k > 0$, we have $P_k(1) = C_{k-1}$.*

This is obvious, since setting $t = 1$ makes the weight $w(\pi)$ equal to 1 for all $\pi \in NC(k-1)$.

Remark 3.6 *The degree of $P_k(t)$ is equal to $k-1$ and the leading coefficient is $(-1)^{k-1}$.*

The degree and leading coefficient are determined by $w(1/2/\ldots/k-1) = \frac{1-(1-t)^k}{t}$ since, no two "fillers" being consecutive points, all other partitions in $NC(k-1)$ have weight t^{m-1}, for some $m \leq \frac{k-1}{2}$.

Remark 3.7 *The Catalan numbers satisfy the relations*

a) $\sum_{j\geq 1} (-1)^{j-1} j \binom{n-j}{j} C_{n-j} = 2^n - n - 1$

and

b) $\sum_{j\geq 0} (-1)^{j} \binom{n-j}{j} C_{n-j} = 1.$

These relations follow immediately from setting, in turn, $t = 0$ and $t = 1$ in Lemma 3.3 and using Remarks 3.4 and 3.5 above, with $n = k-1$.

Non-crossing partitions can be coded using colored Motzkin paths (as in [26]) and the weight $w(\pi)$ on $NC(k-1)$ can be described in lattice path terms:

Corollary 3.8 *Let $M(n)$ be the set of colored Motzkin paths which start at the origin and end at the point $(n,0)$, that is, lattice paths whose steps are $(+1,+1)$ (North-East), $(+1,-1)$ (South-East), or $(+1,0)$ (horizontal), and in which a horizontal step is colored red if it is at zero abscissa and is colored either red or blue if it is at a positive abscissa. To each such path p we associate its number, $s(p)$, of occurrences of two consecutive steps of the form (NE,red) or (blue,red) or (blue, SE) or (NE,SE). Then*

$$\sum_{p\in M(n)} t^{s(p)} = \sum_{j\geq 0} (-1)^j (1-t)^j \binom{n-j}{j} C_{n-j}.$$

Proof. The correspondence between colored Motzkin paths and non-crossing partitions is as follows. The i^{th} step of the path, starting from the origin, determines the block in which the point i is placed in the partition: horizontal red step: i forms a singleton block; NE step: i is the initial element of a non-singleton block; SE step: i is adjoined as maximum element to the block with largest minimum from among the blocks available that do not yet have a maximum assigned; horizontal blue step: i is adjoined (as an "intermediate" element) to the block with largest minimum from among the blocks available that do not yet have a maximum assigned. It is easy to check that the resulting partition is indeed non-crossing and that this correspondence is invertible.

Under this correspondence, "filler" points are precisely those produced by the second of two consecutive steps of the form (NE,red) or (blue,red) or (blue, SE) or (NE,SE), the first two cases giving singleton "fillers" and the last two cases giving "fillers" which are maxima in their blocks. Thus, the statistic $s(p)$ on a colored Motzkin path is equal to the number of "fillers" of the corresponding non-crossing partition, and $\sum_{p\in M(n)} t^{s(p)} = \sum_{\pi\in NC(n)} t^{m(\pi)}$. This last sum, in view of our proof of Lemma 3.3, is equal to $(1-t)^{n+1} + tP^*_{n+1}(t) = (1-t)^{n+1} + tP_{n+1}(t)$. After substituting the expression of Lemma 3.3 for P_{n+1}, a routine calculation completes the proof. $\square$

4 Dyck words counted according to the number of exterior pairs

We now turn to the second statistic, the number of exterior pairs. The main result of this section is the following theorem.

Theorem 4.1 *Let $g_k(t)$ be the generating function for Dyck paths whose number of exterior pairs is k, counted according to their length. Then $g_0(t) = \frac{1-t}{1-2t}$ and*

$$g_k = \frac{t^{k+2}(1-t)}{(1-2t)^{2k+1}} \sum_{i=0}^{k-1} (-1)^i \binom{k-1}{i} C_{i+1} t^i (1-t)^i,$$

where $C_m = \frac{1}{m+1}\binom{2m}{m}$ is the m^{th} Catalan number.

The starting point of our proof will be to show, using Theorem 2.3, that $g_k(t)$ is expressible in terms of a polynomial $R_k(t)$ given by a certain recurrence relation (Lemma 4.2). Then, in Lemma 4.3, we prove an explicit formula for $R_k(t)$ by means of a combinatorial interpretation in terms of weighted colored non-crossing partitions.

Lemma 4.2

$$\begin{cases} g_0(t) = \dfrac{1-t}{1-2t} \\ g_k(t) = \dfrac{t^{k+2}(1-t)}{(1-2t)^{2k+1}} R_k \quad \text{if} \quad k > 1, \end{cases}$$

where R_k is a polynomial satisfying the recurrence

$$\begin{cases} R_1(t) = 1 \\ R_k(t) = t^2(1-t)^2 \displaystyle\sum_{i=1}^{k-2} R_i R_{k-1-i} + (1-2t+2t^2) R_{k-1} \quad \text{if} \quad k > 1. \end{cases}$$

Proof. Recall from Theorem 2.3 that the generating function $G(q,t)$ for Dyck paths counted according to their number of exterior pairs and length is related to the generating function $F(q,t)$ of Dyck paths counted according to their pyramid

$R_1(t) =$	1								
$R_2(t) =$	1	$-2t$	$+2t^2$						
$R_3(t) =$	1	$-4t$	$+9t^2$	$-10t^3$	$+5t^4$				
$R_4(t) =$	1	$-6t$	$+21t^2$	$-44t^3$	$+57t^4$	$-42t^5$	$+14t^6$		
$R_5(t) =$	1	$-8t$	$+38t^2$	$-116t^3$	$+240t^4$	$-336t^5$	$+308t^6$	$-168t^7$	$+42t^8$

Figure 4.9: Table of the polynomials $R_k(t)$ for $1 \leq k \leq 5$

weight and length, by a simple change of variables: $G = G(q,t) = F(q^{-1}, qt)$. Therefore, from the equation of Theorem 2.3 satisfied by $F(q,t)$ we obtain

$$qt(1-t)G^2 - (1+qt-2t)G + (1-t) = 0.$$

since $G(q,t) = \sum_{k\geq 0} g_k(t)q^k$, this yields

$$t(1-t)\sum_{k\geq 0}\sum_{i=0}^{k} g_i g_{k-i} q^{k+1} - (1-2t)\sum_{k\geq 0} g_k q^k - t\sum_{k\geq 0} g_k q^{k+1} + 1 - t = 0,$$

whence we obtain

$$\begin{cases} (1-2t)g_0 - 1 + t = 0 \\ t(1-t)\displaystyle\sum_{i=0}^{k-1} g_i g_{k-1-i} - (1-2t)g_k - tg_{k-1} = 0. \end{cases}$$

This can be restated as the recurrence

$$\begin{cases} g_0 = \dfrac{1-t}{1-2t} \\ g_k = \dfrac{t(1-t)}{1-2t}\displaystyle\sum_{i=0}^{k-1} g_i g_{k-1-i} - \dfrac{t}{1-2t} g_{k-1} \quad (k \geq 1), \end{cases}$$

and the proof is easily completed by induction on k. □

Lemma 4.3 *For every $k > 0$, the polynomial $R_k(t)$ is given by the explicit formula*

$$R_k(t) = \sum_{i=0}^{k-1} (-1)^i \binom{k-1}{i} C_{i+1} t^i (1-t)^i,$$

where $C_m = \frac{1}{m+1}\binom{2m}{m}$ is the m^{th} Catalan number.

Proof. As in the case of Lemma 3.3 where we treated the polynomials $P_k(t)$, the proof will consist of first giving a combinatorial interpretation of the above expression for $R_k(t)$ and then, using this interpretation, establishing that the expression satisfies the recurrence relation of Lemma 4.2.

We consider the set $NC^c(k)$ of non-crossing partitions of $[k]$ whose points $p \in [k]$ are colored purple or green as follows. Starting with $\pi \in NC(k)$, we construct a *colored non-crossing partition*, π^c, by coloring independently each singleton block $\{s\}$, $s > 1$, either purple or green, and coloring every other point purple. Thus, if π has S singleton blocks whose elements are larger than 1, then π will give rise to 2^S colored partitions. The weight of a point $p \in [k]$ is defined to be

$$w(p) := \begin{cases} 1 & \text{if p=1} \\ & \text{or p forms a singleton block colored green;} \\ -t(1-t) & \text{otherwise.} \end{cases}$$

Finally, if π^c is a colored non-crossing partition, then its weight is defined to be $w(\pi^c) := \prod_{p\in[k]} w(p)$, and we claim that

$$R_k^*(t) := \sum_{\pi^c \in NC^c(k)} w(\pi^c) = \sum_{i=0}^{k-1} (-1)^i \binom{k-1}{i} C_{i+1} t^i (1-t)^i.$$

Indeed, every $\pi^c \in NC^c(k)$ which has precisely $i+1$ purple points arises in exactly one fashion from an arbitrary non-crossing partition on $i+1$ elements which will be colored purple (counted by the Catalan number), to which we adjoin $k-i-2$ singleton blocks colored green. As elements of $[k]$, the $i+1$ purple points must include 1, hence their placement in $[k]$ is done in one of $\binom{k-1}{i}$ possible ways. The fact that 1 is always purple gives the correct range for i.

We now show that $R_k^*(t)$ satisfies the recurrence of Lemma 4.3, hence $R_k^*(t) = R_k(t)$, which will complete the proof.

Let $\pi^c \in NC^c(k)$. If 1 is not a singleton block in π^c, let j be the smallest element other than 1 which is in the same block as 1. Suppose $j > 2$ and consider the two partitions π_1^c and π_2^c induced by π^c on the sets $\{2, 3, \ldots, j-1\}$ and $\{j, j+1, \ldots, k\}$, respectively. Note that if the element 2 is colored purple in π^c, then each of π_1^c and π_2^c is a non-crossing partition colored according to the coloring scheme defined above, and that $w(\pi^c) = w(\pi_1^c)w(\pi_2^c)[(-1)t(1-t)]^2$. The last factor follows from the fact that 2 and j being purple points larger than 1 have nontrivial weight in π^c though they have trivial weight in π_1^c and π_2^c, respectively. The sum of the weights of all the colored non-crossing partitions of this type is equal to

$$t^2(1-t)^2 \sum_{j=3}^{k} R_{j-2}^* R_{k-j+1}^*.$$

The remaining partitions π^c, that is, those where 1 is a singleton block or where 1 and 2 are in the same block, can be obtained from the colored non-crossing partitions ρ^c on the set $\{2, 3, \ldots, k\}$ which we modify in either of three ways. The first possibility is to adjoin 1 to ρ^c as a singleton block. The second is to adjoin 1 to the block of ρ^c which contains 2. The third is to adjoin 1 to the block of ρ^c containing 2, and then separate the element 2 into a singleton block which we color green. Since 1 must always be colored purple, the third modification preserves the number of purple points, while the first two give a π^c with one more purple point than ρ^c has. Hence

$$w(\pi^c) = \begin{cases} w(\rho^c)(-t(1-t)) & \text{in the first two cases;} \\ w(\rho^c) & \text{in the third case.} \end{cases}$$

Summing over all partitions ρ^c contributes to $R_k^*(t)$ a total of $(-2t(1-t)+1)R_{k-1}^*$.

Therefore, $R_k^*(t)$ satisfies the same recurrence as $R_k(t)$ for $k \geq 2$, and $R_1^*(t) = 1 = R_1(t)$, hence $R_k^*(t) = R_k(t)$ for all k. □

Note that in the table of Figure 2.4 one identifies – with the aid of [24] – sequences of coefficients of Chebyshev polynomials, or linear combinations of such coefficients. The generating function of the inverses of Chebyshev polynomials of the first kind is $T^{-1}(x,y) = \frac{1}{2}\left(\frac{1-x^2y^2}{1-2x+x^2y^2}+1\right)$, and it can be shown that $T^{-1}(x,y) = \sum_{n\geq 0} \tau_n(x) y^{2n}$, where $\tau_n(x) = \frac{(1-x)x^{2n}}{(1-2x)^{n+1}}$. Thus, by Theorem 4.1, we have

$$g_k(t) = \frac{1}{t^{3k-2}} \tau_{2k}(t) R_k(t).$$

The relation between the Catalan numbers and orthogonal polynomials is well-known. Indeed, a similar relation appears in [25] where the enumeration of Dyck paths according to their order (as defined in section 2.1) gives rise to the quotient of two inverse Chebyshev polynomials of the second kind. More generally, it is in fact known (see [26]) that Dyck paths provide a combinatorial interpretation of the moments of the Chebyshev polynomials of the second kind.

An examination of Figure 4.9 suggests several properties of the polynomials $R_k(t)$, stated in the next remarks. The combinatorial interpretation of $R_k(t)$ given in the proof of Lemma 4.3 provides easy explanations for these properties.

Remark 4.4 *For each $k > 0$, $R_k(0) = 1$.*

Indeed, $R_k(0)$ counts the colored non-crossing partitions whose weight is 1, but the weight is equal to 1 if and only if all points except 1 are colored green. This occurs for a unique coloring of the unique partition which has only singleton blocks.

Remark 4.5 *For all $k > 0$, $R_k(1) = 1$.*

Setting $t = 1$ annihilates the weight of all purple points, except for the point 1. Hence, the colored partition of the previous remark will also be the only one contributing to $R_k(1)$.

Remark 4.6 *The degree of the polynomial $R_k(t)$ is equal to $2(k-1)$, and the leading coefficient of $R_k(t)$ is equal to the Catalan number C_k.*

The maximum power of t occurs from the colored partitions all of whose points are purple. There are C_k such colored partitions, one from each non-crossing partition on $[k]$, and each has weight $w(\pi^c) = [-t(1-t)]^{k-1} = t^{k-1}(t-1)^{k-1}$.

We close with a reformulation of Lemma 4.3 in terms of colored Motzkin paths.

Corollary 4.7 *Let $M(n)$ be the set of colored Motzkin paths, as defined in Corollary 3.8. To each step of such a path p we assign a weight: the first step always has weight 1, every horizontal red step (unless it is the first step of p) has weight $1+z$, and every other step has weight z. If the weight of the path is $r(p)$, equal to the product of the weights of all the steps, then*

$$\sum_{p \in M(n)} r(p) = \sum_{i=0}^{n-1} \binom{n-1}{i} C_{i+1} z^i.$$

Proof. Notice, as in the proof of Corollary 3.8, that the horizontal red steps in colored Motzkin paths are those which correspond to singleton blocks in non-crossing partitions. Thus, the argument is similar to that used in the proof of Lemma 4.3 to establish the combinatorial interpretation of the polynomials $R_k^*(t)$. In that case, we had $z = -t(1-t)$. □

By taking $z = -1$ in Corollary 4.7, we obtain an additional consequence of Lemma 4.3.

Corollary 4.8 *If $NC^{s \subseteq \{1\}}(n)$ denotes the collection of non-crossing partitions of $[n]$ in which at most the point 1 forms a singleton, and if $NC^{s=\phi}(n)$ denotes the collection of non-crossing partitions of $[n]$ with no singleton blocks, then*

$$R_n(-\omega) = (-1)^{n-1} NC^{s \subseteq \{1\}}(n) = (-1)^{n-1}(NC^{s=\phi}(n) + NC^{s=\phi}(n-1)),$$

where ω is a primitive cube root of unity.

Proof. Letting $z = -1$ in Corollary 4.7, gives the inclusion-exclusion expression for $NC^{s \subseteq \{1\}}(n)$, up to a factor of $(-1)^{n-1}$. On the other hand, $z = -1$ is equivalent to $t = -\omega$ in Lemma 4.3, while the second equality follows by a simple counting argument. □

Acknowledgments

The first author wishes to thank Maylis Delest and Jean-Marc Fédou under whose guidance this work originated. The second author expresses her appreciation for the hospitality extended to her by the research groups at LaBRI. We also thank Jean-Guy Penaud for his interest in this work.

References

[1] J. Bonin, L. Shapiro, R. Simion, *Some q–analogues of the Schröder numbers arising from combinatorial statistics on lattice paths*, J. Stat. Planning & Inference, to appear.

[2] M. Bousquet-Mélou, "q-énumération de polyominos convexes", Publications du LaCIM, Université du Québec à Montréal 9, 1991.

[3] M. Bousquet-Mélou, *q-énumeration de polyominos convexes*, J. Combin. Th. A, to appear.

[4] L. Carlitz, *Sequences, paths, ballot numbers*, Fibonacci Quart. 10 (1972) 531-549.

[5] L. Carlitz, R. Scoville, *A note on weighted sequences*, Fibonacci Quart. 13 (1975) 303-306.

[6] L. Comtet, "Advanced Combinatorics," D. Reidel, Dordrecht, 1974.

[7] R. Cori, "Un code pour les graphes planaires et ses applications", Astérisque 27, Société Mathématique de France, 1975.

[8] R. Cori, B. Vauquelin, *Planar maps are well labeled trees*, Canad. J. Math. 33 (1981) 1023-1042.

[9] M.P. Delest, J.M. Fédou, *Counting polyominoes using attribute grammars*, Proceedings Workshop on Attribute Grammars, P. Deransart and M. Jourdan, eds., Lecture Notes in Computer Science 461 (1990) 46-60.

[10] M.P. Delest, J.M. Fédou, *Attribute grammars are useful for combinatorics*, Theoretical Computer Science 98 (1992) 65-76.

[11] M. Delest, X.G. Viennot, *Algebraic languages and polyominoes enumeration*, Theoretical Computer Science 34 (1984) 169-206.

[12] A. Denise, *Etude des q-grammaires*, Mémoire de DEA Informatique, Université Bordeaux I, 1991.

[13] P. Edelman, *Multichains, non-crossing partitions and trees*, Discrete Math. 40 (1982) 171-179.

[14] J.M. Fédou, "Grammaires et q-énumération de polyominos", Ph.D. Thesis, Université de Bordeaux I, 1989.

[15] J.M. Fédou, *Enumeration of skew Ferrers diagrams and basic Bessel functions*, J. Stat. Planning & Inference, to appear.

[16] J. Fürlinger, J. Hofbauer, *q-Catalan numbers*, J. Combin. Theory Ser. A 40 (1985) 248-264.

[17] D.E. Knuth, *Semantics of context-free languages*, Math. Sys. Th. 2 (1968) 127-145.

[18] G. Kreweras, *Sur les partitions non croisées d'un cycle*, Discrete Math. 1 (1972) 333-350.

[19] G. Pólya, *On the number of certain lattice polygons*, J. Combin. Theory 6 (1969) 102-105.

[20] F.W. Schmidt, personal communication.

[21] M.P. Schützenberger, *Context-free languages and pushdown automata*, Information and Control 6 (1963) 246-264.

[22] R. Simion, *Combinatorial statistics on non-crossing partitions*, J. Combin. Theory Ser. A, to appear.

[23] R. Simion and D. Ullman, *On the structure of the lattice of non-crossing partitions*, Discrete Math. 98 (1991) 193-206.

[24] N.J.Sloane, "A handbook of integer sequences", Academic Press, New York, 1979.

[25] M. Vauchaussade de Chaumont, X.G. Viennot, *Enumeration of RNA's secondary structures by complexity*, in "Mathematics in Medecine and Biology", V. Capasso, E. Grosso and S.L. Paven-Fontana eds., Lecture Notes in Biomath. 57 (1985) 360-365.

[26] X.G. Viennot, "Une Théorie Combinatoire des Polynômes Orthogonaux Généraux," UQAM Conference Notes, Montréal, 1983.

[27] X.G. Viennot, *Enumerative combinatorics and algebraic languages*, Proceedings FCT'85, L. Budach, ed., Lecture Notes in Computer Science 199 (1985) 450-464.

[28] X.G. Viennot, *A survey of polyominoes enumeration*, in Proc. $4^{ème}$ Colloque Séries Formelles et Combinatoire Algébrique, Montréal 1992, P. Leroux and C. Reutenauer, eds., Publications du LaCIM, Université du Québec à Montr'eal 11 (1992) 399-420.

[29] D. Zeilberger, *A bijection from ordered trees to binary trees that sends the pruning order to the Strahler number*, Discrete Math. 82 (1990) 89-92.