



COMBINATORIAL INTERPRETATIONS OF CONVOLUTIONS OF THE CATALAN NUMBERS

Steven J. Tedford

Department of Mathematics, Misericordia University, Dallas, Pennsylvania
stedford@misericordia.edu

Received: 5/23/10, Revised: 8/18/10, Accepted: 9/27/10, Published: 1/13/11

Abstract

We reintroduce an interpretation of the k th-fold self convolution of the Catalan numbers by showing that they count the number of words in symbols X and Y , where the total number of Y 's is k more than the total number of X 's, and at no time are there more Y 's than k plus the number of X 's. Using this, we exhibit some of the wide variety of combinatorial interpretations of the k th-fold self convolution of the Catalan numbers. Finally, we show how these numbers appear as the last column in a truncated Pascal's triangle.

1. Introduction

The Catalan numbers, $\{c_n\}_{n=0}^{\infty}$, are well known for having a large variety of combinatorial interpretations. In fact Stanley has a list of 176 different interpretations [9]. Furthermore it would seem natural that there would be a large variety of combinatorial interpretations for the convolutions of the Catalan numbers as well. However, these interpretations are either not well known, or have not been stated explicitly. In fact, the literature either contains explicit formulas for the values of these convolutions, as in [2], and [8], or contain explicit formulas for a combinatorial interpretation of the Catalan numbers, as in [1], [5], and [8]. However, no reference contains a proof showing the combinatorial interpretation without using the explicit formula.

We will attempt to rectify this situation by determining a large variety of combinatorial interpretations of the higher convolutions of the Catalan numbers. We do this by considering the generating functions of both the combinatorial interpretation and of the convolutions of the Catalan numbers. Furthermore, we show that the convolutions of the Catalan numbers appear as the terminal column in a corresponding truncated Pascal's triangle. We begin with some definitions, then show some of the combinatorial interpretations of higher convolutions of the Catalan numbers, and finally end with the occurrence of these numbers in truncated Pascal's triangles.

2. Definitions

We let $\mathbb{Z}_{\geq 0}$ denote the set of non-negative integers. We begin with the general definition of the sequence of Catalan numbers, $\{c_n\}_{n=0}^{\infty}$.

Definition 1. For $n \in \mathbb{Z}_{\geq 0}$, the n th Catalan number, c_n , is defined by

$$c_n = \frac{1}{n+1} \binom{2n}{n} = \frac{(2n)!}{(n+1)!n!}.$$

Although the above definition gives an explicit statement of c_n , the Catalan numbers are often defined in terms of a combinatorial problem. The particular interpretation we will use is as the number of “proper” words in a two letter alphabet.

Let \mathcal{W} be the set of all finite words on $\{X, Y\}$. For any $w \in \mathcal{W}$, define $|w|$ to be the length of w , i.e. the number of symbols that appear in w . For any $1 \leq i \leq |w|$, let w_i represent the initial segment of w with length i . Finally, $n_X(w)$ and $n_Y(w)$ will be the number of occurrences of X and Y , respectively, in w .

With this, we define

Definition 2. A word $w \in \mathcal{W}$ is *proper* if it satisfies the following two properties:

1. For some $n \in \mathbb{Z}_{\geq 0}$, $n_X(w) = n$, $n_Y(w) = n$,
2. For all $1 \leq i \leq 2n$, $n_Y(w_i) \leq n_X(w_i)$.

If $\mathcal{W}(n)$ is the set of proper words with length $2n$, then it is well known [4] that

Proposition 3. For $n \in \mathbb{Z}_{\geq 0}$, $c_n = |\mathcal{W}(n)|$.

We now generalize the idea of proper words to k -proper words:

Definition 4. Let $k \in \mathbb{Z}_{\geq 0}$. A word $w \in \mathcal{W}$ is *k -proper* if it satisfies the following two properties:

1. For some $n \in \mathbb{Z}_{\geq 0}$, $n_X(w) = n$, $n_Y(w) = n + k$,
2. For all $1 \leq i \leq 2n + k$, $n_Y(w_i) \leq n_X(w_i) + k$.

For $n \geq 0$, let $\mathcal{W}(n, k)$ be the set of k -proper words with length $2n + k$. A result which follows immediately from the definition is:

Proposition 5. $\mathcal{W}(n) = \mathcal{W}(n, 0)$.

This directly implies that

Corollary 6. For $n \in \mathbb{Z}_{\geq 0}$, $c_n = |\mathcal{W}(n, 0)|$.

We will relate the number of k -proper words to self-convolutions of the Catalan numbers. So we now define the convolution of a sequence. We choose to do this in terms of generating functions.

Definition 7. Given any sequence $\{a_n\}_{n=0}^\infty$, the generating function for the sequence, $A(z)$, is the formal power series

$$A(z) = \sum_{n=0}^\infty a_n z^n.$$

With this,

Definition 8. Given a sequence $\{a_n\}_{n=0}^\infty$ with generating function $A(x)$, the k th-fold self convolution of $\{a_n\}_{n=0}^\infty$ is the sequence which has $A^{k+1}(x)$ as a generating function. We will let $a(n, k)$ denote the n -th term of the k th-fold self convolution of $\{a_n\}_{n=0}^\infty$.

We now give a few examples to illustrate the convolutions of some given sequences.

Example 9. Suppose $\{a_n\}_{n=0}^\infty$ is the constant sequence of 1's. Then the terms of the first convolution of $\{a_n\}_{n=0}^\infty$ can be found as follows:

$$\begin{aligned} A^2(z) &= (1 + z + z^2 + z^3 + \dots) \cdot (1 + z + z^2 + z^3 + \dots) \\ &= 1 + (z + z) + (z^2 + z^2 + z^2) + (z^3 + z^3 + z^3 + z^3) + \dots \\ &= 1 + 2z + 3z^2 + 4z^3 + \dots \\ &= \sum_{n=0}^\infty (n + 1)z^n. \end{aligned}$$

Thus $a(n, 1) = n + 1$.

Example 10. Suppose $\{a_n\}_{n=0}^\infty$ is any sequence. Then the second convolution of $\{a_n\}_{n=0}^\infty$ can be found as follows:

$$\begin{aligned} A^3(z) &= (a_0 + a_1z + a_2z^2 + \dots)(a_0 + a_1z + a_2z^2 + \dots)(a_0 + a_1z + a_2z^2 + \dots) \\ &= a_0a_0a_0 + (a_1a_0a_0 + a_0a_1a_0 + a_0a_0a_1)z \\ &\quad + (a_2a_0a_0 + a_1a_1a_0 + a_1a_0a_1 + a_0a_2a_0 + a_0a_1a_1 + a_0a_0a_2)z^2 + \dots \end{aligned}$$

This implies that

$$\begin{aligned} a(0, 2) &= a_0a_0a_0, \\ a(1, 2) &= a_1a_0a_0 + a_0a_1a_0 + a_0a_0a_1, \\ a(2, 2) &= a_2a_0a_0 + a_1a_1a_0 + a_1a_0a_1 + a_0a_2a_0 + a_0a_1a_1 + a_0a_0a_2. \end{aligned}$$

If, in fact, we use the sequence of Catalan numbers, then

$$\begin{aligned} c(0, 2) &= 1 \cdot 1 \cdot 1 = 1, \\ c(1, 2) &= 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1 = 3, \\ c(2, 2) &= 2 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 1 + 1 \cdot 2 \cdot 1 + 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot 2 = 9. \end{aligned}$$

As these examples exhibit, the terms in the k th-fold self convolution of $\{a_n\}_{n=0}^\infty$ can be computed directly by

$$a(n, k) = \sum_{\substack{0 \leq i_1, i_2, \dots, i_k, i_{k+1} \leq n \\ \sum_{j=1}^{k+1} i_j = n}} a_{i_1} a_{i_2} \cdots a_{i_k} a_{i_{k+1}},$$

where the summation is over all combinations of $k + 1$ subscripts which add to n .

3. Equality of the Sequences

We intend to show that the sequence, $\{|\mathcal{W}(n, k)|\}_{n=0}^\infty$, of k -proper words and the sequence, $\{c(n, k)\}_{n=0}^\infty$, of k th-fold self convolution of the Catalan numbers are equal. We do this by showing that these two sequences have the same generating function. We consider each sequence separately.

3.1. Catalan Numbers

Let $C(z) = \sum_{n=0}^\infty c_n z^n$ be the generating function of $\{c_n\}_{n=0}^\infty$, and thus $C^{k+1}(z)$ is the generating function of $\{c(n, k)\}_{n=0}^\infty$.

It follows immediately from *Segner's recurrence relation* for the Catalan numbers [7] that the following equation holds:

$$C(z) = zC^2(z) + 1. \tag{1}$$

From this, we obtain the following equation involving higher powers of the Catalan generating function:

$$zC^{k+2}(z) = C^{k+1}(z) - C^k(z). \tag{2}$$

3.2. k -Proper Words

We now consider the generating function for the number of k -proper words. Let $W_k(z)$ denote the generating function of $\{|\mathcal{W}(n, k)|\}_{n=0}^\infty$:

$$W_k(z) = \sum_{n=k}^\infty |\mathcal{W}(n, k)| z^n.$$

In order to determine $W_k(z)$ for $k \geq 0$, we begin with the following lemma.

Lemma 11. For $n, k \in \mathbb{Z}_{\geq 0}$, $|\mathcal{W}(n - 1, k + 2)| = |\mathcal{W}(n, k + 1)| - |\mathcal{W}(n, k)|$

Proof. Let $w \in \mathcal{W}$ with $|w| \geq 1$. One of two possibilities can occur, we consider each separately.

Suppose $w = Xw'$ with $w' \in \mathcal{W}$. In this situation, $|w'| = |w| - 1$ and for $1 \leq i \leq |w'|$, $n_X(w_{i+1}) = n_X(w'_i) + 1$ and $n_Y(w_{i+1}) = n_Y(w'_i)$. Thus, if $w \in \mathcal{W}(n, k + 1)$, then $n_X(w) = n$, $n_Y(w) = n + k + 1$, and $n_Y(w_i) \leq n_X(w_i) + k + 1$. This implies that $n_X(w') = n - 1$, $n_Y(w') = (n - 1) + k + 2$, and $n_Y(w'_i) = n_Y(w_{i+1}) \leq n_X(w_{i+1}) + k + 1 = n_X(w'_i) + k + 2$. Thus $w' \in \mathcal{W}(n - 1, k + 2)$.

Now suppose $w = Yw'$ with $w' \in \mathcal{W}$. Now, $|w'| = |w| - 1$ and for $1 \leq i \leq |w'|$, $n_X(w_{i+1}) = n_X(w'_i)$ and $n_Y(w_{i+1}) = n_Y(w'_i) + 1$. Thus, if $w \in \mathcal{W}(n, k + 1)$, then $n_X(w) = n$, $n_Y(w) = n + k + 1$, and $n_Y(w_i) \leq n_X(w_i) + k + 1$. This implies that $n_X(w') = n$, $n_Y(w') = n + k$, and $n_Y(w'_i) = n_Y(w_{i+1}) - 1 \leq n_X(w_{i+1}) + k = n_X(w'_i) + k$. Thus $w' \in \mathcal{W}(n, k)$.

These two cases together imply that

$$|\mathcal{W}(n, k + 1)| = |\mathcal{W}(n - 1, k + 2)| + |\mathcal{W}(n, k)|,$$

thus proving the lemma. □

Theorem 12. For $k \in \mathbb{Z}_{\geq 0}$, $zW_{k+2}(z) = W_{k+1}(z) - W_k(z)$.

Proof. Recall that $|\mathcal{W}(0, k)| = 1$ for all $k \geq 0$. Thus:

$$\begin{aligned} zW_{k+2}(z) &= \sum_{n=0}^{\infty} |\mathcal{W}(n, k + 2)| z^{n+1} = \sum_{n=1}^{\infty} |\mathcal{W}(n - 1, k + 2)| z^n \\ &= \sum_{n=1}^{\infty} (|\mathcal{W}(n, k + 1)| - |\mathcal{W}(n, k)|) z^n \\ &= \sum_{n=1}^{\infty} |\mathcal{W}(n, k + 1)| z^n - \sum_{n=1}^{\infty} |\mathcal{W}(n, k)| z^n \\ &= \sum_{n=0}^{\infty} |\mathcal{W}(n, k + 1)| z^n - \sum_{n=0}^{\infty} |\mathcal{W}(n, k)| z^n \\ &= W_{k+1}(z) - W_k(z). \end{aligned}$$

□

With this, we immediately obtain the following corollary.

Corollary 13. For $n, k \in \mathbb{Z}_{\geq 0}$, $c(n, k) = |\mathcal{W}(n, k)|$.

4. Combinatorial Interpretations of $c(n, k)$.

With Corollary 13, we can now obtain a large number of combinatorial interpretations of the k th-fold self-convolution of the Catalan numbers. We begin by interpreting k -proper words as a particular subset of proper words.

Theorem 14. For $n \geq 0$ and $k \geq 0$, $c(n, k)$ counts the number of proper words $w \in \mathcal{W}(n + k)$ with $w_k = XX \cdots X$.

Proof. Consider the mapping from $\mathcal{W}(n, k)$ to $\mathcal{W}(n + k)$ which takes the word $w \in \mathcal{W}(n, k)$ to the word $X^k w$, where X^k denotes k copies of X . This mapping is clearly a bijection between $\mathcal{W}(n, k)$ and $\{w \in \mathcal{W}(n + k) : w_k = XX \cdots X\}$. Thus, using Corollary 13, this set has size $c(n, k)$. \square

Example 15. We now exhibit this particular interpretation of $c(n, k)$ by considering a few particular values of both n and k . Consider, in particular, $\mathcal{W}(4)$:

$$\mathcal{W}(4) = \{XXXXYYYYY, XXXYXYYYY, XXXYYXYXY, XXXYYYXY, \\ XYXXYYY, XYXYXYXY, XYXXXYXY, XYXXXYXY, \\ XYXXXYXY, XYXXXYXY, XYXXXYXY, XYXXXYXY, \\ XYXXXYXY, XYXXXYXY\}.$$

Notice, of the 14 elements of $\mathcal{W}(4)$, exactly 9 of them begin with at least two X 's.

$$\mathcal{W}(2, 2) = \{XXYYY, XYXY, XYXY, XYXY, XYXY, \\ YXYXY, YXYXY, YXYXY, YXYXY\}.$$

This agrees with the fact that $c(2, 2) = 9$.

With Theorem 14 and the standard bijections between the different combinatorial interpretations (many of which are exhibited in [3]), we obtain the following combinatorial interpretations of $c(n, k)$. (there are many more, one for each interpretation of the Catalan numbers.)

Corollary 16. For $n \geq 0$ and $k \geq 0$, $c(n, k)$ counts the following:

1. The number of balanced strings of $n + k$ brackets with an initial string of at least k left brackets.
2. The number of lattice paths consisting of either North or East steps, which do not pass above the line $y = x$, which start at $(k, 0)$ and end at $(n + k, n + k)$.
3. The number of rooted plane trees with $n + k$ edges whose leftmost branch has length at least k .
4. The number of increasing functions f on $[n + k] = \{1, 2, \dots, n + k\}$ where $f(i) = 1$ if $1 \leq i \leq k$.
5. The number of full binary trees with $2(n + k)$ edges which have only single vertices on the left side of the first k branches.
6. The number of standard Young tableaux of size $(n + k, n + k)$ where the first term in the second row is greater than k .

Another immediate corollary of Theorem 14 is

Corollary 17. For $n, k \in \mathbb{Z}_{\geq 0}$, $c(n, k) \leq c(n + k)$.

5. Truncating Pascal’s Triangle

In [10], Young, Taylor, and Zwicker showed that the Catalan numbers arise in the last column when the rows of Pascal’s Triangle are truncated immediately after the central column. We show that by changing where the rows are truncated, different convolutions of the Catalan numbers appear in the last column.

As with many combinatorial number theory objects, Pascal’s Triangle can be defined in many different ways. We will use the recursive definition of Pascal’s Triangle.

Definition 18. Pascal’s Triangle, \mathcal{P} , is defined by

$$\mathcal{P}(n, r) = \begin{cases} 0 & \text{if } n < 0 \text{ or } r > n, \\ 1 & \text{if } n = r = 0, \\ \mathcal{P}(n - 1, r) + \mathcal{P}(n - 1, r - 1) & \text{otherwise.} \end{cases}$$

We show the terms of the first five rows of \mathcal{P} in Figure 1. In Figure 2, we show the values of the first six rows. We show the terms of the first six rows of \mathcal{P} and their values in Figure 1. We also label the columns of Pascal’s triangle starting with labeling the central column 0. These column numbers are indicated in the figure.

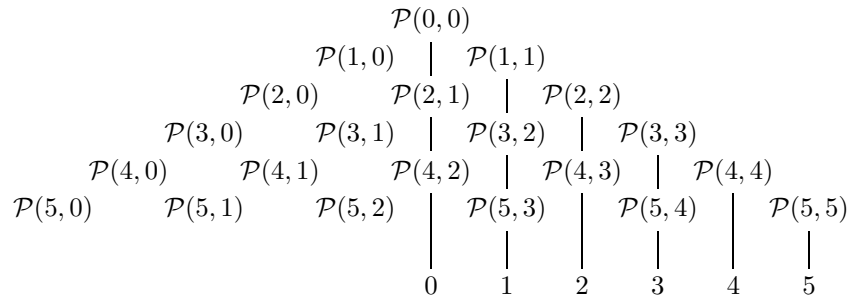


Figure 1: Pascal’s Triangle with columns labeled

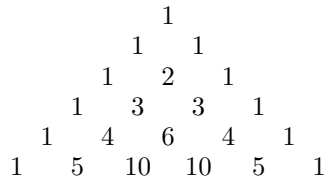


Figure 2: \mathcal{P}

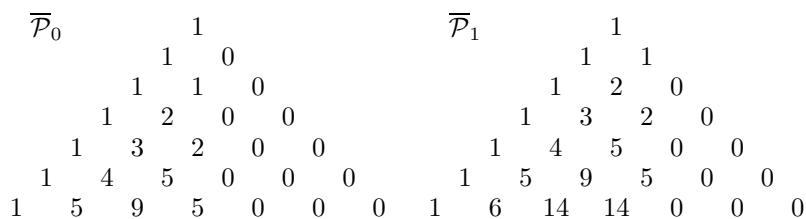


Figure 3: $\overline{\mathcal{P}}_0$ and $\overline{\mathcal{P}}_1$

We use this recursive definition to build a collection of new triangles, each similar to Pascal’s Triangle.

Definition 19. For any $k \in \mathbb{Z}_{\geq 0}$, the k th truncated Pascal’s Triangle, denoted $\overline{\mathcal{P}}_k$, is defined by the following:

$$\overline{\mathcal{P}}_k(n, r) = \begin{cases} 0 & \text{if } n < 0 \text{ or } r > \lfloor \frac{n}{2} \rfloor + k, \\ 1 & \text{if } n = r = 0, \\ \overline{\mathcal{P}}_k(n - 1, r) + \overline{\mathcal{P}}_k(n - 1, r - 1) & \text{otherwise.} \end{cases}$$

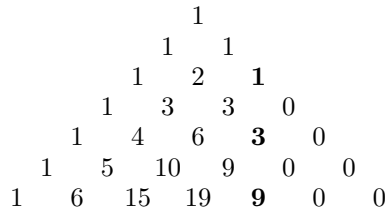
From a procedural point of view, we have replaced elements of Pascal’s Triangle with zeros after column k and have recalculated the remaining entries of the triangle using Pascal’s Identity. Figure 3 shows $\overline{\mathcal{P}}_0$ and $\overline{\mathcal{P}}_1$.

In [10], Young, Taylor, and Zwicker showed that the n th term of column 1 (starting with $n = 0$), in $\overline{\mathcal{P}}_1$ is $c_{n+1} = c(n, 1)$. In fact

Theorem 20. Let $k \in \mathbb{Z}_{\geq 0}$, the n th term (starting with $n = 0$) of column k in $\overline{\mathcal{P}}_k$ is $c(n, k)$.

Proof. In Pascal’s Triangle, $\mathcal{P}(n, r)$ counts the number of paths from the point $(0, 0)$ to the point (n, r) consisting only of moves going either southeast or southwest. Similarly, $\overline{\mathcal{P}}_k(n, r)$ counts the number of paths starting at $(0, 0)$ to (n, r) consisting of moves going either southeast or southwest which do not go past column k . Since the n th term of column k in $\overline{\mathcal{P}}_k$ occurs at position $(m, \lfloor \frac{m}{2} \rfloor + k)$, where $m = k + 2n$, such a path will consist of n moves southwest and $n + k$ moves to the southeast. Furthermore, since this path does not pass column k , the number of southeast moves never exceeds the number of southwest moves plus k . Thus, the n th term of column k (n starting at 0) in $\overline{\mathcal{P}}_k$ is $|\mathcal{W}(n, k)|$. By Corollary 13, this is $c(n, k)$. \square

Figure 4 shows the first rows of triangle $\overline{\mathcal{P}}_2$, with the last column highlighted, to further show this result.

Figure 4: $\overline{\mathcal{P}}_2$

6. Summary

A generating function proof has been presented of a known result relating convolutions of the Catalan numbers to a combinatorial interpretation regarding words in two letters. With this proof, additional interpretations of these convolutions have become accessible. Finally, we have shown a new way to find the Catalan numbers and their convolutions in Pascal-like triangles.

Although explicit formulas for the Catalan numbers and their convolutions are known [8], an explicit formula for the other terms of $\overline{\mathcal{P}}_k$ have not been determined for general k .

Acknowledgements We would like to thank the anonymous referee, whose comments greatly improved this article.

References

- [1] André, D., *Solution directe du problème résolu par M. Bertrand*, Comptes Rendus Acad. Sci. Paris **105** (1887), 436–437.
- [2] Catalan, E., *Sur les nombres de Segner*, Rend. Circ. Mat. Pal. **1** (1887), 190–201.
- [3] Choo, Koo-Guan, *Catalan numbers*, <http://www.maths.usyd.edu.au/u/kooc/catalan.html>
- [4] Hilton, P. and Pedersen, J., *The ballot problem and Catalan numbers* Nieuw Archeif voor Wiskunde **8** (1990), 209–216.
- [5] Hilton, P. and Pedersen, J., *Catalan numbers, their generalizations, and their uses*, The Math. Intell. **13** (1991), 64–75.
- [6] Hoggart Jr., V. E. and Bicknell, M., *Pascal, Catalan, and general sequence convolution arrays in a matrix*, Fibonacci Quarterly **14** (1976), 135–143.
- [7] Koshy, T., Catalan Numbers with Applications, Oxford University Press, New York, 2009.

- [8] Larcombe, P. J. and French, D. R., *The Catalan number k -fold self-convolution identity: The original formulation*, Journal of Combinatorial Mathematics and Combinatorial Computing **46** (2003), 191–204.
- [9] Stanley, R. P., Enumerative Combinatorics, Volume 2, Catalan Addendum, Cambridge University Press 1999, <http://www-math.mit.edu/~rstan/ec/catadd.pdf>.
- [10] Young, S., Taylor, A., and Zwicker, W., *Counting quota systems: A combinatorial question from social choice theory*, Mathematics Magazine **68** (1995), 331–343.