Proceedings of the 3rd International Conference on Applied Physics and Mathematical Modeling
DOI: 10.54254/2753-8818/2025.DL29880

Catalan Numbers from the Perspective of Convolution Inversion and Generalized Dyck Paths

Zihan Zhao[1*], Yuanzhi Zhu[2]

[1]*College of Science, Mathematics and Technology, Wenzhou-Kean University, Wenzhou, China*
[2]*Faculty of Science and Technology, Beijing Normal University-Hong Kong Baptist University United International College, Zhuhai, China*
**Corresponding Author. Email: zzihan@kean.edu*

Abstract. Catalan numbers are a rich and extensively studied class of combinatorial numbers in combinatorics. They have a wide range of applications, possess various combinatorial interpretations, and are often used to solve various combinatorial counting problems. They are usually denoted by C_n . Among their many combinatorics implications, this article will focus on their utility in Dyck path counting and, using this counting method, especially the idea of convolution inversion, present several other combinatorial counting examples and generalizations.

Keywords: Catalan number, Combinatorics, Dyck Path, Convolution

1. Introduction

Let's review one of the simplest definitions of a Catalan number: Consider a walk on a two-dimensional coordinate grid, starting from the origin and moving to (n,n) , where n is a given positive integer. Each step is allowed to be only one unit length upward or rightward, and the walk path must always lie below the line segment $(0,0)$ to (n,n) , that is, below the function $y = x$. The total number of such paths is the n th Catalan number [1]. Below is a brief proof of this result, followed by a result and proof for a variant problem, which introduces the innovative aspects of this paper.

When counting lattice paths, note that if we ignore the restriction conditions, the total number is $\binom{2n}{n}$. Now, considering the restriction conditions, we require the number of "legal" paths that do not cross the diagonal.

A natural way is to find the number of all paths that cross the diagonal. In figure 1, draw the line $y = x + 1$ in the graph, and [2] reflect the portion of all "illegal" paths from $(0,0)$ to $(k, k+1)$ across this line. This establishes a one-to-one correspondence with all paths from $(-1,1)$ to (n, n) , whose number is $\binom{2n}{n-1}$, Subtracting this part gives

$$\binom{2n}{n} - \binom{2n}{n-1} = \frac{1}{n+1}\binom{2n}{n} = C_n \tag{1}$$

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2025.DL29880

Which is the classical standard Catalan number, whose convolutional recursive relation and generating function are:

$$C_{n+1} = \sum_{i=0}^{n} C_i C_{n-i} \tag{2}$$

$$C\left(x\right) = \frac{1-\sqrt{1-4x}}{2x} \tag{3}$$

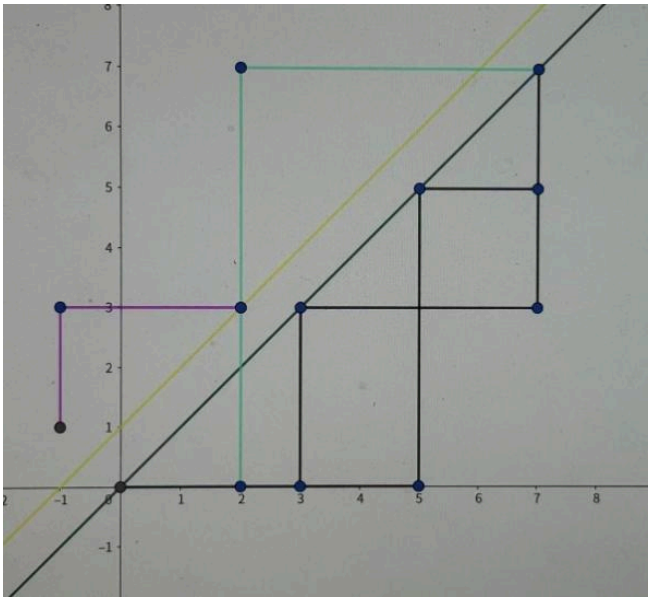


Figure 1. A mirror method

After completing the counting of classical Dyck paths, we consider changing the constraint conditions, such as counting the total number of paths from $(0,0)$ to (n, n) in a grid that do not touch the diagonal strictly.

In figure 2, notice that the answer we want for this problem can be obtained by simply subtracting the number of paths that just touch the diagonal from all the “legal” paths, therefore, if we denote by S_n the count of paths that do not touch the diagonal strictly, we can obtain the following recurrence relation [3]:

$$C_n - S_n = \sum_{\mathrm{i=1}}^{\mathrm{n\text{-}1}} \mathrm{S_i C_{n\text{-}i}} \tag{4}$$

$$C_n = \sum_{\mathrm{i=0}}^{\mathrm{n}} \mathrm{S_i C_{n\text{-}i}} \tag{5}$$

Construct the generating function [4]

DOI: 10.54254/2753-8818/2025.DL29880

$$\begin{cases} C(x) = \sum_{i=0}^{\infty} C_n x^n \\ S(x) = \sum_{i=0}^{\infty} S_n x^n \end{cases} \tag{6}$$

$$C(x) * S(x) = \sum_{n=0}^{\infty} \sum_{i=0}^{n} S_i C_{n-i} x^n = S_0C_0 + (S_1C_0 + S_0C_1)x + (S_2C_0 + S_1C_1 + S_0C_2)x^2 + ...$$

$$= C_1x + C_2x^2 + C_3x^3 + ... = C(x) - 1$$

$$S\left(x\right) = 1 - \frac{1}{C(x)} = \frac{1 - \sqrt{1 - 4x}}{2} = xC\left(x\right) \tag{7}$$

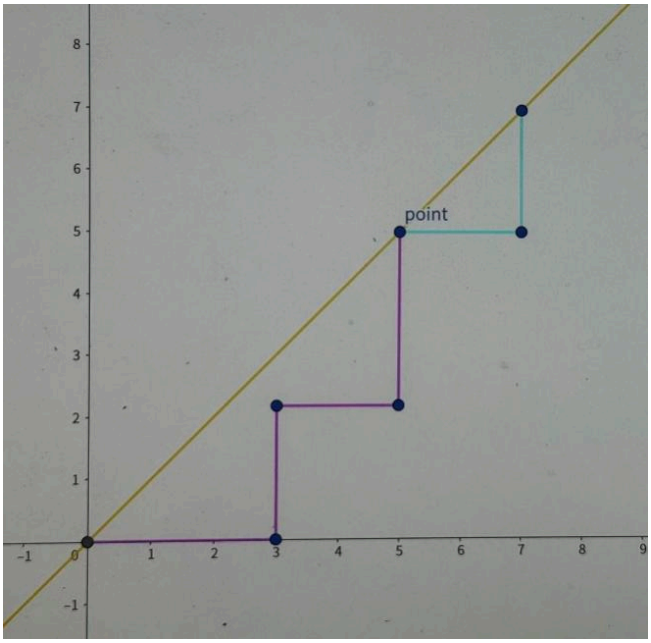


Figure 2. A classification method

This tells that for every term of the $S(x)$, the coefficient walks up one step compared to $C(x)$. Thus, we have:

$$S_n = C_{n-1} \tag{8}$$

In solving the original Dyck Path problem, two key deductions and operations are valuable. First, we note that for any "illegal path" that crosses the diagonal, there exists a coordinate (a moment) at which it first crosses the constraining segment. We classify all illegal paths based on this coordinate. Second, based on the classification results in step 1, we construct a bijection by taking symmetries, mapping each illegal path to a unique element in another set (the general path from $(-1,1)$ to (n, n)), thus achieving a complete counting without duplication or omission [5].

DOI: 10.54254/2753-8818/2025.DL29880

2. Generalized catalan numbers and Dyck paths under other restrictions

Now let's change the constraints on the walk path and see the counting results under other constraints, which is a generalization [6]. One of the most common generalizations is to stretch the walk area, that is, the end point of the path becomes (kn, n), and the path must always be above the line $y = \frac{1}{k}x$. The path total result in this case is the generalized Catalan number, that is [7]

$$C_n^{(k)} = \frac{1}{kn+1}\binom{(k+1)n}{n} \tag{9}$$

Another very interesting example is to go from the origin to (m, n), where m and n are given non-negative integers. Again, each step can only go one unit to the right or up, and the path must not be higher than the line segment from the origin to (m, n). Here is the number of paths [8]:

$$\# = \begin{cases} \frac{1}{m+n}\binom{m+n}{n} & if\ gcd(m,n) = 1 \\ \frac{1}{m+n}\sum_{d|gcd(m,n)} \varphi(d)\binom{\frac{m+n}{d}}{\frac{n}{d}} & if\ gcd(m,n) \neq 1 \end{cases} \tag{10}$$

The proof of these two results is omitted here.

3. Counting a more general Dyck path

Let us now turn to several more generalized problems concerning the enumeration of Dyck paths. Their solutions rely on combinatorial reasoning similar to that employed in deriving Catalan numbers, as well as on certain applications of the principle of inclusion–exclusion, namely, classifying all the excluded paths according to the “illegal” positions that render them inadmissible.

Consider this walking problem: On a finite line segment with a beginning at $(-n,\ 0)$ and an end at $(n,\ 0)$, an object starts from the origin and moves left or right one unit per second, always preventing it from exceeding the boundaries on both sides [9]. How many paths are there in total within time $t(t \geq m)$?

The answer to this question is:

$$a_n\left(t\right) = \sum_{m=-n}^{n}\left(\sum_{i=\left\lceil -\frac{t+m}{4(n+1)}\right\rceil}^{\left\lfloor \frac{t-m}{4(n+1)}\right\rfloor}\binom{t}{\frac{t+m}{2}+2i\left(n+1\right)} - \sum_{j=\left\lceil -\frac{t+m}{4(n+1)}-\frac{1}{2}\right\rceil}^{\left\lfloor \frac{t-m}{4(n+1)}-\frac{1}{2}\right\rfloor}\binom{t}{\frac{t+m}{2}+\left(2j+1\right)\left(n+1\right)}\right) \tag{11}$$

It's easy to see that the expression within the summation is formally very similar to Catalan numbers. This is because its solution is based on the same principles we emphasized earlier: classification and constructive symmetry.

A complete proof is given below.

First, we may assume that the endpoint is fixed at m. In the unconstrained case, we then have:

$$\begin{cases} N(+1) + N(-1) = t \\ N(+1) - N(-1) = m \end{cases} \tag{12}$$

DOI: 10.54254/2753-8818/2025.DL29880

Therefore, under unconstrained conditions, the total number of paths from the origin to m in t steps are:

$$\binom{t}{\frac{t+m}{2}}$$

Here we adopt the convention that any binomial coefficient with non-integer or negative entries is taken to be zero; in particular, if t and m have different parity then the total number of paths is zero.

Next, we exclude all paths that violate the boundary. For any path S that attains an out-of-bounds position, there exists a first such time τ . Preserve the entire trajectory after time τ (excluding the step at τ), and reflect the initial segment of the path up to and including τ about the vertical line $x = n$ (or, symmetrically, about $x = -n$). This reflection yields two image endpoints:

$$x_1 = 2n + 2\,;\; x_2 = -2n - 2$$

Consequently, every path to be excluded can be associated with a path from the origin terminating at one of these image endpoints. However, this correspondence is not one-to-one: paths that violate the boundary on both sides are counted twice by the single reflections. Therefore, the reflection–exclusion procedure must be iterated, alternately adding and subtracting the counts at each step, until the domain's boundaries are exhausted. In this way one obtains a recursive, stepwise elimination process — an application of the inclusion–exclusion principle:

$$x_3 = 4n + 4\,;\; x_4 = -4n - 4$$

$$\cdots$$

$$x_{2k-1} = 2^k(n+1)\,;\; x_{2k} = -2^k(n+1)$$

$$= \binom{t}{\frac{t+m}{2}} - \binom{t}{\frac{t+m}{2}+2(n+1)} - \binom{t}{\frac{t+m}{2}-2(n+1)} + \binom{t}{\frac{t+m}{2}+4(n+1)} + \binom{t}{\frac{t+m}{2}-4(n+1)} - \ldots$$

$$= \sum\binom{t}{\frac{t+m}{2}+2k(n+1)} - \sum\binom{t}{\frac{t+m}{2}+(2k+1)(n+1)} \tag{13}$$

Incorporating the constraints of the domain, we obtain:

$$a_{0\to m,\,n}\left(t\right) = \sum_{i=\left\lceil -\frac{t+m}{4(n+1)} \right\rceil}^{\left\lfloor \frac{t-m}{4(n+1)} \right\rfloor} \binom{t}{\frac{t+m}{2} + 2i(n+1)} - \sum_{j=\left\lceil -\frac{t+m}{4(n+1)} - \frac{1}{2} \right\rceil}^{\left\lfloor \frac{t-m}{4(n+1)} - \frac{1}{2} \right\rfloor} \binom{t}{\frac{t+m}{2} + (2j+1)(n+1)} \tag{14}$$

It suffices to take the finite sum of the above expression over $m = -n, \ldots, +n$; this yields the result stated earlier.

In particular, if we set the endpoint $\mathrm{m} = 0$, the boundary $\mathrm{n} = 1$, and take t to be the even integer $2r$, the model produces the following elegant combinatorial identity:

DOI: 10.54254/2753-8818/2025.DL29880

$$2^r = \sum_{i=\lceil -\frac{r}{4} \rceil}^{\lfloor \frac{r}{4} \rfloor} \left(\frac{2r}{r+4i} \right) - \sum_{j=\lceil -\frac{r}{4} - \frac{1}{2} \rceil}^{\lfloor \frac{r}{4} - \frac{1}{2} \rfloor} \left(\frac{2r}{r+2(2j+1)} \right) \tag{15}$$

This result can be extended to the discrete two-dimensional setting: on the lattice formed by the intersections of the families of lines $x = -n, \ldots, n$ and $y = -n, \ldots, n$, a particle starts at $(0,0)$ and at each time moves along an edge between lattice points. The total number of its trajectories up to time t is:

$$b_n\left(k\right) = \sum_{k=0}^{t} \binom{t}{k} a_n\left(k\right) a_n\left(t-k\right) \tag{16}$$

which is also a form of convolution, since the movements take place on x-axis and the movements take place on y-axis are independent.

4. Another combinatorics problems with a similar approach

4.1. Problem

(1) On an infinite square chessboard, there is a king. Each move allows it to move to one of the eight surrounding cells.

(2) On an infinite regular triangular chessboard, there is a king. Each move allows it to move to one of the six surrounding cells.

(3) On an infinite k-dimensional lattice chessboard, there is a king. Each move allows it to move to one of all the surrounding cells.

Problem: Let the square where the king currently resides be the initial position. Under the different constraints above, give a theoretical analytical solution for the probability that the king will never return to its initial position within n steps.

4.2. Solution

(1) Understanding the conjugation relationship, we can set the total number of string groups satisfying "length n; all strings in the group have the same number of $+1$ s and -1 s at the nth position" to u_n. In other words, u_n encompasses all paths that are at the initial position at the end of step n.

The following general terms can be indirectly noted from the fact that there can be no empty steps:

$$u_n = \begin{cases} \sum_{i=0}^{k} \binom{2k}{2i} \cdot \binom{2i}{i} \cdot \left(\sum_{j=0}^{i} \binom{2i}{2j} \cdot \binom{2k-2i+2j}{k-i+j} \right), & if\ n = 2k, k \in Z+ \\ \sum_{i=1}^{k-1} \binom{2k-1}{2i} \cdot \binom{2i}{i} \cdot \left(\sum_{j=1}^{i} \binom{2i}{2j-1} \cdot \binom{2k-2-2i+2j}{k-1-i+j} \right), & if\ n = 2k-1,\ k \in Z+ \end{cases} \tag{17}$$

$$u_1 = 0,\ u_2 = 8$$

For the case where n is even, the reason u_n admits the following representation is as follows. First, among the $2k$ horizontal steps choose $2i$ of them to be nonzero; of those $2i$ chosen steps take exactly i to be $+1$ and i to be -1, while the remaining horizontal steps are 0. Second, from

DOI: 10.54254/2753-8818/2025.DL29880

the vertical steps select at least $2(k-i)$ of them to be nonzero, and among those selected take half to be $+1$ and half to be -1, with the remainder set to 0. This construction guarantees that at every time step the king moves by one unit in at least one coordinate. The case of odd nn follows by the same reasoning.

Here we invoke the foregoing method of "classification and exclusion." Let v_n denote the number of admissible paths — that is, those paths for which, for every positive integer $N < n$, the king is not at the origin at time N. Observe that any path not counted in v_n must have some $s < n$ such that, after step s, it never returns to the origin. In other words, we classify and exclude illegal paths according to the time of their last "illegal", this gives a convolution equality in the same form as the Catalan numbers:

$$v_n = 8^n - \sum_{i=0}^{n-2} u_{n-i} v_i \tag{18}$$

$$v_0 = 1,\ v_1 = 8$$

The probability can be expressed as:

$$p_n = \frac{v_n}{8^n} \tag{19}$$

So far, the first question has been solved. The graph of $p_n (n \leq 300)$ is in figure 3 below.

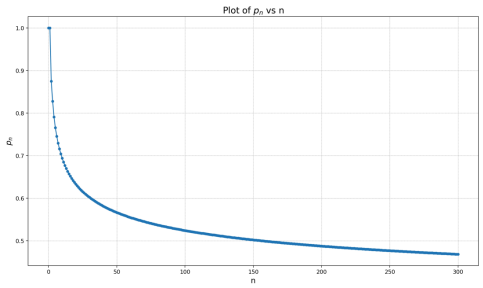


Figure 3. The probability pn on the classical checkerboard

(2) The second question, the six-direction walk problem, differs only in the derived expression of u_n, while the rest of the ideas are the same. The solution is as follows:

$$u_n = \begin{cases} \sum_{i=0}^{k} \left(\frac{2k}{2i}\right) \cdot \left(\frac{2i}{i}\right) \cdot \left(\sum_{j=0}^{i} \left(\sum_{x+y=2j} \left(\frac{i}{x}\right) \cdot \left(\frac{i}{y}\right) \cdot \left(\frac{2k-2i}{k-i+j-x}\right)\right)\right), & if\ n = 2k, k \in Z+ \\ \sum_{i=0}^{k} \left(\frac{2k-1}{2i}\right) \cdot \left(\frac{2i}{i}\right) \cdot \left(\sum_{j=1}^{i} \left(\sum_{x+y=2j-1} \left(\frac{i}{x}\right) \cdot \left(\frac{i}{y}\right) \cdot \left(\frac{2k-1-2i}{k-i+j-1-x}\right)\right)\right), & if\ n = 2k-1, k \in Z+ \end{cases} \tag{20}$$

$$u_1 = 0,\ u_2 = 6$$

$$\begin{gathered} v_n = 6^n - \sum_{i=0}^{n-2} u_{n-i} v_i \\ v_0 = 1, v_1 = 6 \end{gathered} \tag{21}$$

$$p_n = \frac{v_n}{6^n} \tag{22}$$

DOI: 10.54254/2753-8818/2025.DL29880

The graph of $p_n(n \leq 300)$ is in figure 4 below.

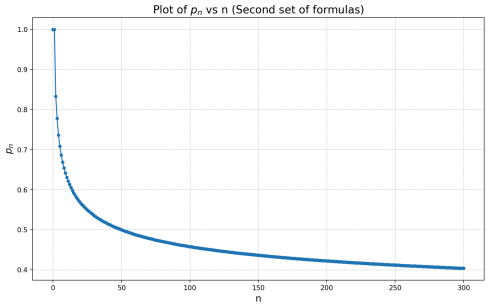


Figure 4. The probability pn on the triangular checkerboard

(3) For the third question, an elegant way is discovered to combine the exclusion of vacancies and the inclusion-exclusion principle [10]. Under the premise that the structure of remains unchanged, the general formula of u_n can be written as follows:

$$u_n = \sum_{j=0}^{n} (-1)^{n-j} \binom{n}{j} \left(\sum_{i=0}^{\left[\frac{j}{2}\right]} \binom{j}{2i} \binom{2i}{i} \right)^k \tag{23}$$

This formulation may be understood as follows: start with the set of all unconstrained n-step trajectories, subtract those with at least one idle step, then add back those with at least two idle steps, then subtract those with at least three idle steps and so on in succession.

Considering that each position in the k-dimensional lattice has $(3^k - 1)$ neighboring positions, we have:

$$v_n = (3^k - 1)^n - \sum_{i=0}^{n-2} u_{n-i} v_i \tag{24}$$

$$v_0 = 1,\ v_1 = 3^k - 1$$

$$p_n = \frac{v_n}{(3^k - 1)^n} \tag{25}$$

When k is 2, the result is equivalent to the verbose expression of the first question.

This solves all three problems.

In fact, for u_n and v_n, as with the Catalan numbers, the analogous objects of interest are their generating functions. Noting that $v_1 = 0$, we therefore have:

$$\begin{cases} u(x) = \sum_{i=0}^{\infty} u_n x^n \\ v(x) = \sum_{i=0}^{\infty} v_n x^n \end{cases} \tag{26}$$

$$u\left(x\right) * v\left(x\right) = \sum_{n=0}^{\infty} \sum_{i=0}^{n} u_i v_{n-i} x^n = \sum_{n=0}^{\infty} \left(3^k - 1\right)^n \tag{27}$$

DOI: 10.54254/2753-8818/2025.DL29880

However, since this series diverges, it cannot provide any further analytic contribution to the result.

5. Conclusion

This paper provides a detailed summary and extension of the Catalan number methodology for enumerating Dyck paths. Starting from a convolution perspective, it employs the anchoring of "boundary-violation features," in combination with reflections and the principle of inclusion–exclusion, to address a related combinatorial problem. The first part introduces two fundamental counting problems directly associated with the Catalan numbers; the second part briefly reviews the definition of generalized Catalan numbers; and the third and fourth parts delve into the methodological framework of Dyck path enumeration, presenting analytic solutions to three distinct versions of the "king's walk problem."

References

[1] Lattice path enumeration. (2015). Handbook of Enumerative Combinatorics, 613–702. https: //doi.org/10.1201/b18255-16
[2] Bertrand, J. (1887). Solution d'un problème. Comptes Rendus, 105, 369.
[3] Roman, S. (2019). An introduction to Catalan numbers.
[4] Flajolet, P., & Sedgewick, R. (2009). Analytic Combinatorics. Cambridge University Press.
[5] Stanley, R. P. (2015). Catalan numbers. Cambridge University Press.
[6] Fuss, N. I. (1791). Solutio quaestionis, quot modis polygonum in partes triangulares sectum in plures triangulos dividere possit. Nova Acta Academiae Scientiarum Imperialis Petropolitanae, 9, 243–251.
[7] Zhou, R. D. P., & Yan, S. H. F. (2017, January). The Raney numbers and (s, s+1)-core partitions - Sciencedirect. https: //www.sciencedirect.com/science/article/pii/S0195669816300622
[8] Bizley, M. T. (1954). Journal of the Institute of Actuaries, 80(1), 55–62. https: //doi.org/10.1017/s002026810005424x
[9] Feller, W. (1968). An Introduction to Probability Theory and Its Applications, Vol. 1. Wiley.
[10] Bóna, M. (2007). Introduction to Enumerative Combinatorics. McGraw Hill.