RATIONAL CATALAN NUMBERS

JONATHAN YU

Abstract. The Catalan numbers are a fundamental sequence in combinatorics. Some of the things that they count are Dyck paths and binary trees. The Catalan numbers can be extended to the rational Catalan numbers, which count things such as rational Dyck paths and rational binary trees.

1. Introduction

The Catalan numbers are among the most ubiquitous number sequences in combinatorics. Named after Eugène Charles Catalan in 1838, they now appear in over 200 different contexts. The Catalan number C_n counts the number of Dyck paths, full binary trees, balanced parentheses, triangulations of a convex polygon, and so much more.

A common generalization of the Catalan numbers is the rational Catalan numbers. Like the Catalan numbers, these count a number of combinatorial objects, including rational Dyck paths and rational binary trees. In this paper, we will explore the Catalan numbers and derive its formula and recurrence. Then we will generalize to the rational Catalan numbers and prove its formula.

2. Preliminaries

Definition 2.1 (Dyck path)**.** A *Dyck path* is a lattice path between $(0, 0)$ and (n, n) consisting of $(0, 1)$ and $(1, 0)$ steps that does not cross above the line $y = x$ (see figure 1).

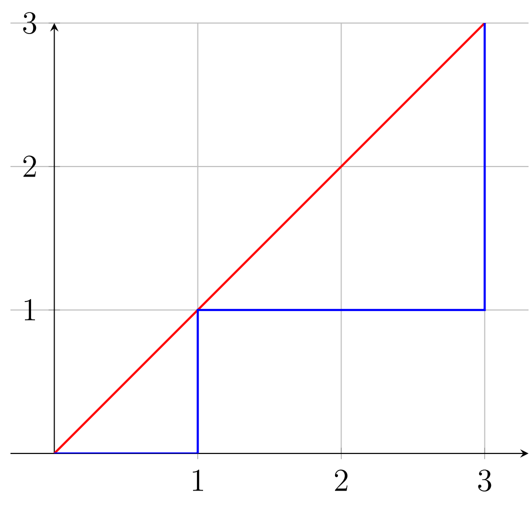

Figure 1. A Dyck path from $(0, 0)$ to $(3, 3)$

Date: December 8, 2025.

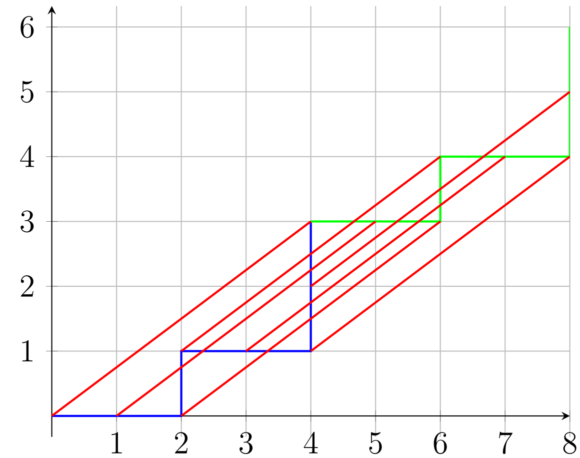

FIGURE 2. Cyclic shift of paths

Definition 2.2 (Semilength)**.** The *semilength* of a Dyck path is the number of $(0,1)$ steps (or equivalently the number of $(1,0)$ steps).

Definition 2.3 (Catalan number)**.** The nth *Catalan number* C_n is

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

Proposition 2.4. *The number of Dyck paths of semilength n is C_n.*

Lemma 2.5 (Cycle lemma)**.** *Let $a_1, a_2, \ldots, a_n$ be a sequence such that $a_i \in \{-1, 1\}$ and $\sum_{i=1}^{n} a_i = k > 0$. Then exactly k cyclic shifts of the sequence always have a positive partial sum.*

Proof. See [1] for a proof of the Cycle Lemma. □

The following proof has been adapted from [2].

Proof of proposition 2.4. Consider the paths from $(0,0)$ to $(n+1,n)$ that do not go above the line $y = \frac{n}{n+1}x$. The first step must be $(1,0)$, since a $(0,1)$ step would put us above the line. Then we have a path from $(1,0)$ to $(n+1,n)$. Observe that any path that crosses the line $y = x - 1$ must also cross $y = \frac{n}{n+1}x$ and vice versa, so we have a Dyck path of semilength n. This tells us that there are the same number of paths from $(0,0)$ to $(n+1,n)$ that do not go above the line passing through the two points as there are Dyck paths.

Let E represent a $(1,0)$ step and N a $(0,1)$ one. Any path from $(0,0)$ to $(n+1,n)$ can then be represented by a permutation of $n+1$ Es and n Ns. Consider a cyclic shift of a given string of Es and Ns. We claim that exactly one of these paths stays below the line $y = \frac{n}{n+1}x$. To see this, we concatenate two copies of the path, one from $(0,0)$ to $(n+1,n)$ and the other from $(n+1,n)$ to $(2n+2,2n)$. We draw lines parallel to $y = \frac{n}{n+1}x$ going through each of the points on the bottom copy (see figure 2).

There is a bijection between the Es and Ns and a sequence of 1s and -1s. In our above construction, there is exactly 1 more E step than N steps. By theorem 2.5, exactly one of these sequences is strictly positive, so its corresponding path is the unique path that forms a Dyck path.

There are $\binom{2n+1}{n}$ ways to permute $n+1$ Es and n Ns. Since only one of the $2n+1$ cyclic shifts yields a Dyck path, there are a total of $\frac{1}{2n+1}\binom{2n+1}{n}$ Dyck paths. Finally,

$$\frac{1}{n+1}\binom{2n}{n} = \frac{(2n)^{\underline{n}}}{(n+1)n!} = \frac{(2n)^{\underline{n-1}}}{n!} = \frac{(2n+1)^{\underline{n}}}{(2n+1)n!} = \frac{1}{2n+1}\binom{2n+1}{n}.$$

This completes the proof. □

Proposition 2.6 (Catalan recurrence)**.** *The Catalan number C_{n+1} satisfies*

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}.$$

The following proof is from [2].

Proof. Consider a Dyck path from $(0,0)$ to $(n+1,n+1)$. Let $(k+1,k+1)$ for $0<k<n$ be the first place it hits the line $y=x$ and consider the path from $(0,0)$ to $(k+1,k+1)$. Because we assumed that $(k+1,k+1)$ was the first place that the path hit $y=x$, the first step must be $(1,0)$ and the last $(0,1)$. But this just gives us a Dyck path of semilength k, so there are C_k such paths. The path from $(k+1,k+1)$ to $(n+1,n+1)$ is a Dyck path of semilength $n-k$, giving us C_{n-k} paths.

If $k=0$ or $k=n$, our Dyck path does not hit the line $y=x$ anywhere besides $(0,0)$ and $(n+1,n+1)$. This means that our first step must be $(1,0)$ and our last $(0,1)$. This gives us a Dyck path of semilength n, so we have C_n such paths in each case. We see that $C_0C_n = C_nC_0 = C_n$, so the recurrence works even when $k=0$ or $k=n$.

Considering each of the C_kC_{n-k} paths for $0 \le k \le n$ gives us all possible Dyck paths of semilength $n+1$, so summing over all k yields the result. □

3. Rational Catalan Numbers

The rational Catalan numbers are a generalization of the Catalan numbers. They count things similar to that of the classical Catalan numbers.

Definition 3.1 (Rational Dyck path)**.** Let $m,n \in \mathbb{Z}^+$ such that $\gcd(m,n)=1$. A *rational Dyck path* is a lattice path between $(0,0)$ and (m,n) consisting of $(0,1)$ and $(1,0)$ steps that does not cross above the line $y = \frac{n}{m}x$ (see figure 3).

Definition 3.2. The *rational Catalan number $C_{m,n}$* is

$$C_{m,n} = \frac{1}{m+n}\binom{m+n}{m}.$$

Proposition 3.3. *The number of Dyck paths from $(0,0)$ to (m,n) not going above the line $y=\frac{n}{m}x$ is given by $C_{m,n}$.*

The following proof has been adapted from [2].

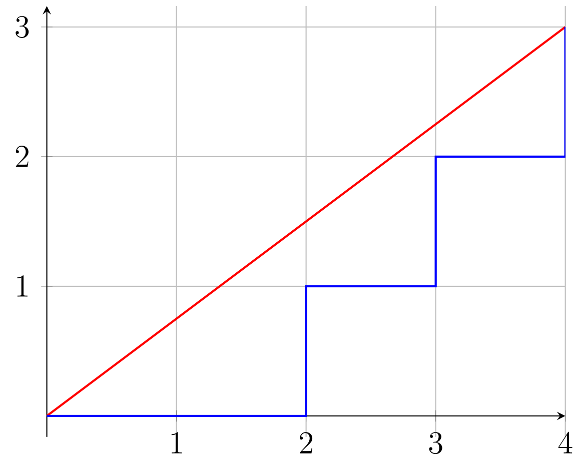

FIGURE 3. A rational Dyck path from $(0,0)$ to $(4,3)$

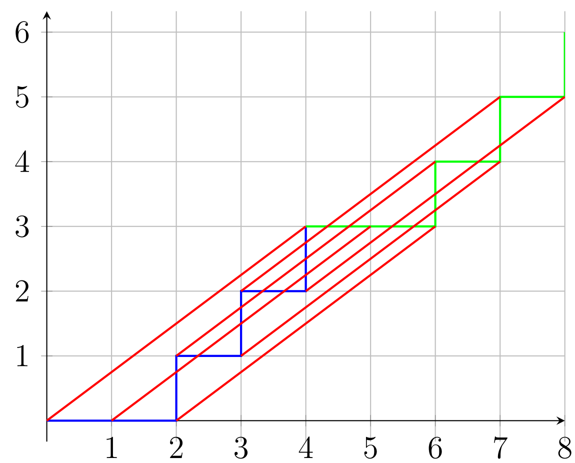

FIGURE 4. Cyclic shift of paths

Proof. Consider a sequence of m $E = (1,0)$ and n $N = (0,1)$ steps. This corresponds to a lattice path from $(0,0)$ to (m,n). Since each path corresponds to a sequence of m Es and n Ns, we have $\binom{m+n}{m}$ such paths.

We claim that exactly 1 of the $m+n$ cyclic shifts of a sequence of Es and Ns lies entirely below the line $y = \frac{n}{m}x$. To see this, we create a copy of the path to get a path from $(0,0)$ to $(2m,2n)$. As before, we draw lines connecting each point on the original path with its corresponding point on the copy (see figure 4).

Since m and n are relatively prime, each of these lines is distinct, meaning that there is exactly 1 line that lies above the rest. This line is the unique line corresponding to a rational Dyck path. Suppose for the sake of contradiction that there was another line that would yield a rational Dyck path. Then it must also be the highest line, or else it would intersect the path at some point other than the endpoints. But there cannot be two highest lines, so we conclude that only 1 of the $m+n$ cyclic shifts gives a rational Dyck path.

Since there are $\binom{m+n}{m}$ ways to order m Es and n Ns and only one of the $m+n$ cyclic shifts gives a rational Dyck path, we get that the number of rational Dyck paths from $(0,0)$ to (m,n) is $\frac{1}{m+n}\binom{m+n}{m}$, or $C_{m,n}$. □

References

[1] Nachum Dershowitz and Shmuel Zaks, *The Cycle Lemma and Some Applications*, Tel Aviv University, School of Computer Science, [Online]. Available: https://www.cs.tau.ac.il/~nachumd/papers/CL.pdf

[2] Simon Rubinstein-Salzedo, *Combinatorics*, Euler Circle, Mountain View, CA.