

Topics in Combinatorics IV, Term 1¹

1 Catalan numbers

1.1 Definitions

Definition 1.1 (one of many). The n -th Catalan number C_n is the number of sequences $(\varepsilon_1, \dots, \varepsilon_{2n})$ with $\varepsilon_i = \pm 1$ such that

$$\begin{aligned} & \cdot \sum_{i=1}^{2n} \varepsilon_i = 0; \\ & \cdot \sum_{i=1}^k \varepsilon_i \geq 0 \text{ for every } k \leq 2n. \end{aligned}$$

Remark. Sequences in the definition above are called *ballot sequences*.

Example 1.2. $n = 2$: the only sequences of length $2n = 4$ are $(1, 1, -1, -1)$ and $(1, -1, 1, -1)$, so $C_2 = 2$.
 $n = 3$: there are five sequences (list them!), so $C_3 = 5$.

Two equivalent definitions of C_n :

- the number of “bracketings” of a non-associative product of $n + 1$ variables;
- the number of triangulations of a convex $(n + 2)$ -gon on a plane (here by a triangulation we mean a maximal collection of non-crossing diagonals, it automatically subdivides the polygon into triangles).

Example 1.3. For $n = 3$, there are precisely five bracketings:

$$((a_1 a_2) a_3) a_4 \quad (a_1 (a_2 a_3)) a_4 \quad (a_1 a_2) (a_3 a_4) \quad a_1 ((a_2 a_3) a_4) \quad a_1 (a_2 (a_3 a_4)).$$

There are also precisely five triangulations of a pentagon: two non-crossing diagonals share a vertex, and there are five vertices to choose from.

Exercise 1.4. Show that the two definitions above are equivalent to Definition 1.1.

Ballot sequences can be represented by *Dyck paths*: these are paths in a $n \times n$ square going along the grid from one corner (say, $(0, n)$) to the opposite (i.e., $(n, 0)$) and staying above the main diagonal. The bijection with ballot sequences is obvious: $\varepsilon_i = 1$ becomes a move to the right, and $\varepsilon_i = -1$ becomes a move down.

One can draw the top right half of the square only, in this case a Dyck path can be understood as a path from $(0, 0)$ to $(2n, 0)$ with steps $(1, 1)$ and $(1, -1)$ never going below x -axis. This is the model we will usually use in the sequel.

Remark. Paths whose all steps are vectors of $\mathbb{Z}^d$ are called *lattice paths*. The two models above for Dyck paths sit in $\mathbb{Z}^2$ generated, respectively, by $(1, 0)$, $(0, -1)$, and $(1, 1)$, $(1, -1)$.

¹Based on Alex Postnikov's notes

Example 1.5.

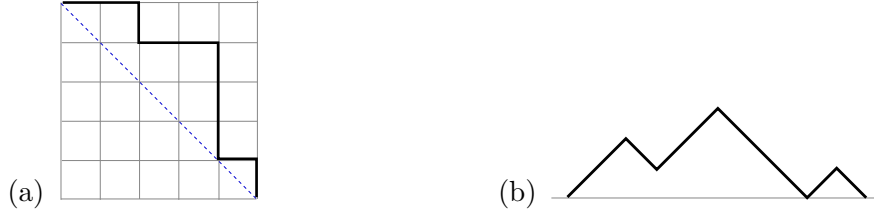


Figure 1.1: Dyck paths with steps (a) $(1,0)$ and $(0,-1)$ and (b) $(1,1)$ and $(1,-1)$.

1.2 Explicit formula for C_n

Theorem 1.6. $C_n = \frac{1}{n+1} \binom{2n}{n}$.

Example 1.7. $C_3 = \frac{1}{4} \binom{6}{3} = \frac{1}{4} \frac{6!}{3!3!} = \frac{6 \cdot 5 \cdot 4}{4 \cdot 6} = 5$; $C_4 = \frac{1}{5} \binom{8}{4} = \frac{1}{5} \frac{8!}{4!4!} = \frac{8 \cdot 7 \cdot 6 \cdot 5}{5 \cdot 8 \cdot 3} = 14$; $C_5 = \frac{1}{6} \binom{10}{5} = 42$.

We will look at three different proofs of the theorem.

Proof 1: by reflection. The number of all lattice paths from $(0,0)$ to $(2n,0)$ with steps $(1,1)$ and $(1,-1)$ is equal to $\binom{2n}{n}$ – choose n steps “up” out of $2n$ steps. We will compute the number of ones which go below the x -axis (call them *bad*) and subtract.

Observe that a path is bad if and only if it intersects the line $y = -1$. Find the first point where it touches the line $y = -1$, and reflect the part of the path to the right of this point with respect to the line $y = -1$. We get a new lattice path, it goes from $(0,0)$ to $(2n,-2)$. We now claim that this map establishes a bijection between *bad* paths from $(0,0)$ to $(2n,0)$ and *all* lattice paths from $(0,0)$ to $(2n,-2)$.

(Proof to be continued) □