

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.

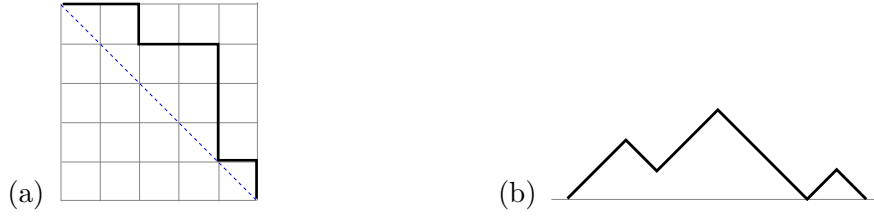


Figure 1.1: Dyck paths with steps (a) $(1,0)$ and $(0,-1)$ and (b) $(1,1)$ and $(1,-1)$.

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 $(2n,-2)$.

Indeed, this map is injective, and there is an inverse: take any lattice path from $(0,0)$ to $(2n,-2)$, take the first point where it touches the line $y = -1$, and reflect the right part – we get a bad path.

Now, the paths from $(0,0)$ to $(2n,-2)$ contain $n-1$ steps up and $n+1$ steps down, so the number of all paths from $(0,0)$ to $(2n,-2)$ is $\binom{2n}{n+1}$. Thus,

$$C_n = \binom{2n}{n} - \binom{2n}{n+1} = \frac{(2n)!}{n!n!} - \frac{(2n)!}{(n-1)!(n+1)!} = \frac{(2n)!}{n!n!} - \frac{n}{n+1} \frac{(2n)!}{n!n!} = \frac{1}{n+1} \binom{2n}{n}$$

□

Proof 2: by cyclic shifts. First, observe that

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

Interpret $\binom{2n+1}{n}$ as all lattice paths from $(0,0)$ to $(2n+1,-1)$, i.e. paths with n steps up and $n+1$ steps down. Then C_n is the number of those paths (call them *good*) that go below the x -axis at the last step only. We want to show that good paths constitute precisely $1/(2n+1)$ part of all paths.

Let us switch to sequences of ± 1 . Take any sequence $(\varepsilon_1, \dots, \varepsilon_{2n+1})$ of length $2n+1$ adding up to -1 , and consider all its $2n+1$ *cyclic shifts*:

$$(\varepsilon_1, \dots, \varepsilon_{2n+1}), (\varepsilon_2, \varepsilon_3, \dots, \varepsilon_{2n+1}, \varepsilon_1), (\varepsilon_3, \dots, \varepsilon_{2n+1}, \varepsilon_1, \varepsilon_2), \dots, (\varepsilon_{2n+1}, \varepsilon_1, \dots, \varepsilon_{2n}).$$

Now the proof follows from the following statement which we leave as an exercise

Exercise. (1) Show that all cyclic shifts are distinct.

(2) Show that out of $2n + 1$ shifts of one sequence precisely one is good. (Hint: consider the leftmost lowest point of the path).

□

Definition 1.8. Let $(a_n)_{n \in \mathbb{Z}_{\geq 0}}$ be a sequence of non-negative integers. A formal power series $\sum_{k=0}^{\infty} a_k x^k$ is called a *generating function* of (a_n) .

Example 1.9. Let $a_n = 1$ for every n . The the generating function is $A(x) = 1 + x + \cdots = 1/(1 - x)$.

In general, how to find a closed formula for the generating function $A(x)$ of a sequence (a_n) ? One can follow the following plan:

- write a recurrence relation on a_n ;
- interpret the recurrence relation as an equation on $A(x)$;
- solve the equation and get an explicit expression for a_n .

We will proceed along this plan to get a *third proof* of the theorem.

Lemma 1.10 (Recurrence on Catalan numbers).

$$C_n = C_0 C_{n-1} + C_1 C_{n-2} + \cdots + C_{n-1} C_0 = \sum_{k=1}^n C_{k-1} C_{n-k}$$

Proof. Take any Dyck path of length $2n$, and consider the first point $2k > 0$ where the path touches the x -axis. Then on the right there is a Dyck path of length $2n - 2k = 2(n - k)$. On the left there is a Dyck path of length $2k$ which stays strictly above x -axis, which means that we can think of it as a Dyck path of length $2k - 2$ between points $(1, 1)$ and $(2k - 1, 1)$. Both Dyck paths on the left and on the right are arbitrary, so there are precisely $C_{k-1} \cdot C_{n-k}$ Dyck paths which touch x -axis for the first time at $2k$, and the result follows.

□

Defining $C_0 = 1$, we can now recursively compute any Catalan number.

Example 1.11.

$$C_1 = C_0 \cdot C_0 = 1; \quad C_2 = C_0 \cdot C_1 + C_1 \cdot C_0 = 1 + 1 = 2; \quad C_3 = C_0 \cdot C_2 + C_1 \cdot C_1 + C_2 \cdot C_0 = 2 + 1 + 2 = 5.$$

Lemma 1.12. The generating function $C(x)$ satisfies the equation $xC(x)^2 - C(x) + 1 = 0$.

Proof. We know that $C_n = \sum_{k=1}^n C_{k-1} C_{n-k}$. Multiplying by x^n , we get

$$C_n x^n = \sum_{k=1}^n C_{k-1} C_{n-k} x^n = x \sum_{k=1}^n (C_{k-1} x^{k-1}) (C_{n-k} x^{n-k})$$

Now, summing on $n > 0$, we get on the left $\sum_{n=1}^{\infty} C_n x^n = C(x) - C_0 = C(x) - 1$. On the right, we get

$$\begin{aligned} \sum_{n=1}^{\infty} \sum_{k=1}^n x(C_{k-1}x^{k-1})(C_{n-k}x^{n-k}) &= x \sum_{n=1}^{\infty} \sum_{i=0}^{n-1} (C_i x^i)(C_{n-i-1}x^{n-i-1}) = \\ &= x \sum_{m=0}^{\infty} \sum_{i=0}^m (C_i x^i)(C_{m-i}x^{m-i}) = x \cdot C(x) \cdot C(x), \end{aligned}$$

so we obtain the equation $C(x) - 1 = xC(x)^2$, which is precisely what we wanted. $\square$

Lemma 1.13.

$$C(x) = \frac{1 - \sqrt{1 - 4x}}{2x}$$

Proof. Solving quadratic equation with respect to $C(x)$ (while considering x as a parameter), we see that $C(x) = \frac{1 \pm \sqrt{1 - 4x}}{2x}$. The reason to choose the sign is the following: we want $\lim_{x \rightarrow 0} C(x) = C(0)$, where we already know $C(0) = C_0 = 1$. Now, if we consider the positive sign, then $\lim_{x \rightarrow 0} C(x) = \infty$, while the negative sign gives the required limit. $\square$

We are left to extract the explicit expression for C_n from the generating function. For this, we recall the definition of a generalization of binomial coefficients.

Definition 1.14. Given $\alpha \in \mathbb{R}$ and $k \in \mathbb{N}$, the *binomial coefficient* $\binom{\alpha}{k}$ is given by

$$\binom{\alpha}{k} = \frac{\alpha(\alpha - 1)(\alpha - 2) \dots (\alpha - k + 1)}{k!}$$

Exercise. Show that

$$(1 + y)^\alpha = \sum_{k \geq 0} \binom{\alpha}{k} y^k$$

Proof 3: via generating function. We now apply the exercise above to $\alpha = 1/2$, $y = -4x$. We get

$$\sqrt{1 - 4x} = \sum_{k \geq 0} \binom{\frac{1}{2}}{k} (-4x)^k = 1 + \binom{\frac{1}{2}}{1} (-4x) + \binom{\frac{1}{2}}{2} (-4x)^2 + \dots,$$

so

$$C(x) = \frac{1 - \sqrt{1 - 4x}}{2x} = -\frac{1}{2} \left(\binom{\frac{1}{2}}{1} (-4) + \binom{\frac{1}{2}}{2} (-4)^2 x + \dots \right) = -\frac{1}{2} \left(\sum_{n \geq 0} \binom{\frac{1}{2}}{n+1} (-4)^{n+1} x^n \right).$$

Therefore,

$$\begin{aligned} C_n &= -\frac{1}{2} \binom{\frac{1}{2}}{n+1} (-4)^{n+1} = -\frac{1}{2} \frac{\frac{1}{2}(\frac{1}{2} - 1)(\frac{1}{2} - 2) \dots (\frac{1}{2} - n)}{(n+1)!} (-4)^{n+1} = \\ &= -\frac{1}{2} \frac{\frac{1}{2}(-\frac{1}{2})(-\frac{3}{2}) \dots (-\frac{2n-1}{2})}{(n+1)!} (-4)^{n+1} = \frac{1 \cdot 3 \cdot \dots \cdot (2n-1)}{(n+1)!} \frac{2^{2n+2}}{2^{n+2}} = \frac{(2n)!}{(2^n n!)(n+1)!} \cdot 2^n = \frac{1}{n+1} \binom{2n}{n} \end{aligned}$$

as required. $\square$

1.3 Further examples and Applications

Example 1.15 (Drunkard's walk). A drunkard walks randomly along a line in steps of length 1 to the left or right with probability $1/2$. The starting point is $x = 1$. The walk is terminated when the drunkard reaches the point $x = 0$. Question: what is the probability the walk terminates?

Reaching $x = 0$ after $2k + 1$ steps is equivalent to returning to $x = 1$ after $2k$ steps and then going left, so this can be understood as a Dyck path. The probability of any individual path is $(1/2)^{2k+1}$, the number of such paths of length $2k + 1$ is clearly C_k . Thus,

$$P = \sum_{k=0}^{\infty} C_k \frac{1}{2^{2k+1}} = \frac{1}{2} \sum_{k=0}^{\infty} C_k \frac{1}{2^{2k}} = \frac{1}{2} \sum_{k=0}^{\infty} C_k \left(\frac{1}{4}\right)^k = \frac{1}{2} C\left(\frac{1}{4}\right) = \frac{1}{2} \frac{1 - \sqrt{1 - 4 \cdot \frac{1}{4}}}{2 \cdot \frac{1}{4}} = 1$$

We list below some further interpretations of Catalan numbers C_n .

- The number of *plane binary trees* with n vertices.

A plane binary tree can be defined recursively: if not empty, it has a root vertex, a left subtree, and a right subtree, both of which are binary trees (either of them can be empty). When drawing such trees, the root is drawn on the top, with an edge drawn from it to the root of each of its subtrees. All binary trees with 3 vertices are shown in Fig. 1.2.



Figure 1.2: Plane binary trees with 3 vertices

- The number of *complete plane binary trees* with $n + 1$ leaves.

A plane binary tree is complete if it has no vertices of valence 2 (except for the root which is of valence 2), i.e., for every vertex both left and right subtrees are either simultaneously nonempty or empty (in the latter case the vertices are called *leaves*). All complete binary trees with 4 leaves are shown in Fig. 1.3.

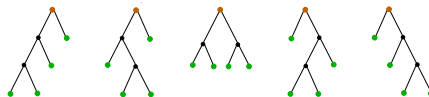


Figure 1.3: Complete plane binary trees with 4 leaves

- The number of *plane trees* with $n + 1$ vertices.

A plane tree can also be defined recursively: it has a root vertex, and, in the case the whole tree is not a single vertex, it has a sequence $(T_1, \dots, T_k)$ of subtrees T_i , $1 \leq i \leq k$, each of which is also a plane tree. In particular, the subtrees of each vertex are ordered; when drawing such trees, the subtrees are drawn from left to right. The root is on the top, with an edge drawn from it to the root of each of its subtrees. All plane trees with 4 vertices are shown in Fig. 1.4.

- The number of *non-crossing matchings* on $2n$ vertices.

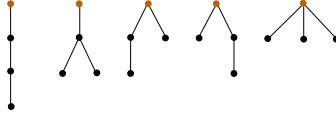


Figure 1.4: Plane trees with 4 vertices

A matching on $2n$ vertices can be drawn as a way to connect $2n$ nodes on the x -axis by arcs, with every node connected to precisely one other node by an arc drawn in the upper halfplane. A matching is *non-crossing* if all arcs can be drawn in a way such that no pair of arcs intersects.

All non-crossing matchings on 6 vertices are shown in Fig. 1.5.

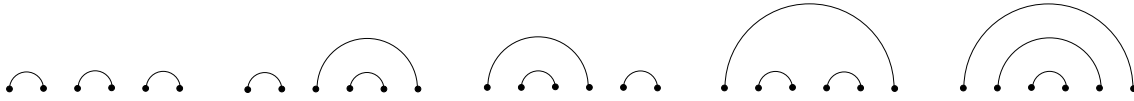


Figure 1.5: Non-crossing matchings on 6 vertices

• The number of *non-nesting matchings* on $2n$ vertices.

A matching is *non-nesting* if all the arcs can be drawn in a way such that no arc is above the other, or, equivalently, there is no pair of arcs with ends at points (a_1, b_1) and (a_2, b_2) such that $a_1 < a_2 < b_2 < b_1$.

All non-nesting matchings on 6 vertices are shown in Fig. 1.6.

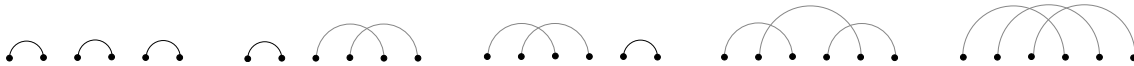


Figure 1.6: Non-nesting matchings on 6 vertices