Catalan numbers

Gábor Hetyei

There are many combinatorial problems whose solution is a Catalan number. Here we discuss only some of the most famous ones.

1 Lattice paths

Imagine that there are n persons holding a 5 dollar bill and n persons holding s 10 dollar bill in front of a box office. The ticket costs 5 dollars and at the beginning the box office has no cash. How many ways are there to line up the $2n$ people if we want to make sure the box office never runs out of change? *We define the Catalan number C_n as the answer to this question.*

The problem is equivalent to the following *lattice path enumeration* problem: how many lattice paths are there, starting from $(0,0)$ consisting of n northeast steps (from (x,y) to $(x+1,y+1)$) and of n southeast steps (from (x,y) to $(x+1,y-1)$) that never go below the x-axis. Indeed, we can associate a northeast step to each person holding a 5 dollar bill, and southeast step to each person holding a 10 bill. As we parse the people standing in line, we obtain a lattice path. The box office never runs out of change exactly when the lattice path remains above the x-axis.

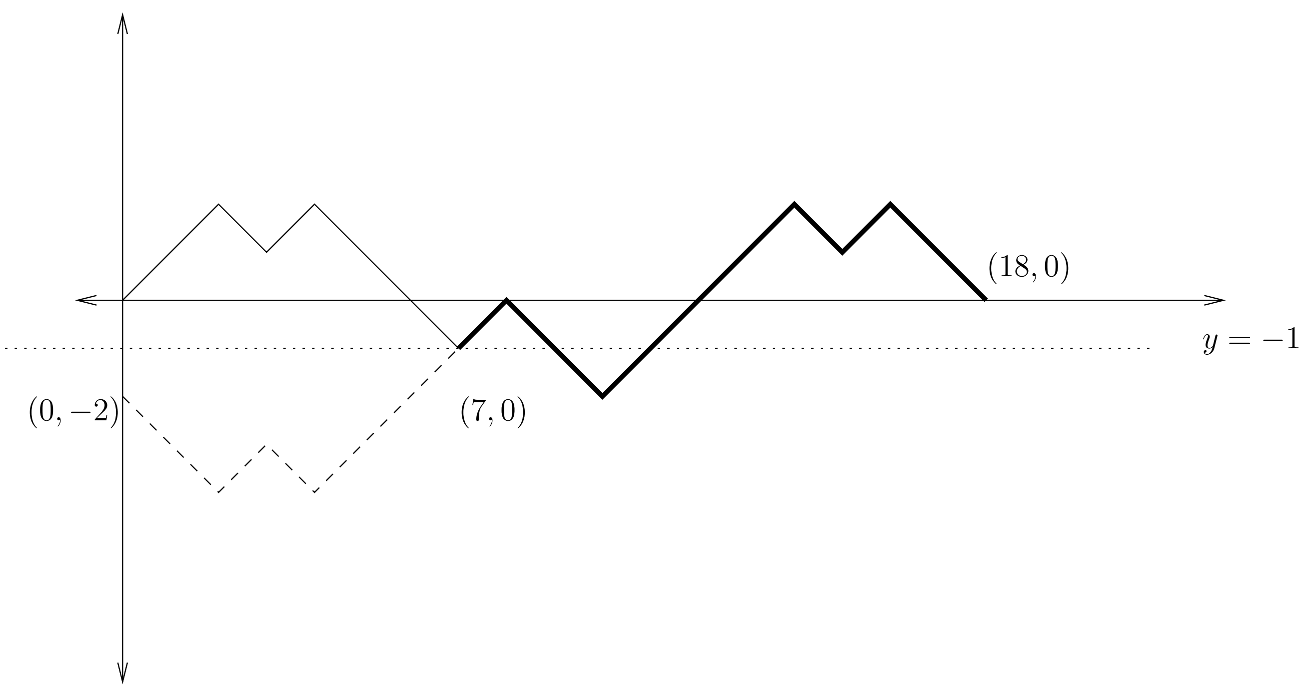

Figure 1: An example of the use of the reflection principle

The lattice path visualization allows us to solve the problem using the *reflection principle*. Let us note first that the number of all lattice paths that have exactly n northeast steps and n southeast steps is $\binom{2n}{n}$. All such lattice paths begin at $(0,0)$ and end at $(2n,0)$. From the number of all lattice paths we want to subtract the number of all *"bad"* lattice paths, i.e., the number of lattice path from $(0,0)$ to $(2n,0)$ going below the x axis. Since each step ends at a lattice point (that is, a point with integer coordinates), a "bad" lattice path necessarily has a point on the line $y=-1$. Let us reflect the part of a "bad" lattice path before its first point on the line $y=-1$ to the line $y=-1$. We obtain a lattice path from $(0,-2)$ to $(2n,0)$. Conversely, any lattice path from $(0,-2)$ to $(2n,0)$ has at least one point on the line $y=-1$. Reflecting the part before the first point on the line $y=-1$ to the line $y=-1$ yields a "bad" lattice path from $(0,0)$ to $(2n,0)$. The reflection described above provide a bijection. Hence the number of "bad" lattice paths from $(0,0)$ to $(2n,0)$ is the same as the number of all lattice paths from $(0,-2)$ to $(2n,0)$. A lattice path beginning at $(0,-2)$ and consisting of northeast and southeast steps only, ends at $(2n,0)$ exactly when it has $n+1$ northeast steps and $n-1$ southeast steps. The number of all such lattice paths is $\binom{2n}{n+1}$. Thus the Catalan number C_n is given by

$$C_n = \binom{2n}{n} - \binom{2n}{n+1} = \binom{2n}{n}\left(1-\frac{n}{n+1}\right) = \frac{1}{n+1}\binom{2n}{n}.$$

The first 11 Catalan numbers C_n are shown in Table 1,

n	0	1	2	3	4	5	6	7	8	9	10
C_n	1	1	2	5	14	42	132	429	1430	4862	16796

Table 1: The Catalan numbers C_n for $n \leq 10$.

2 A recurrence for the Catalan numbers

Another way to solve the lattice path counting problem presented in Section 1 is to find a recurrence as follows. Consider any lattice path from $(0,0)$ to $(2n,0)$ that never goes below the x axis. Assume k is the least positive integer such that $(2k,0)$ is on the lattice path. The part of the lattice path between $(0,0)$ and $(2k,0)$ necessarily begins with a northeast step from $(0,0)$ to $(1,1)$, ends with a southeast step from $(2k-1,1)$ to $(2k,0)$ and never goes below the line $y=1$ between $(1,1)$ and $(2k-1,1)$. The number of lattice paths from $(1,1)$ to $(2k-1,1)$ that never go below the line $y=1$ is C_{k-1} and there are C_{n-k} ways to continue the lattice path from $(2k,0)$ to $(2n,0)$. We obtain the following recurrence:

$$C_n = \sum_{k=1}^{n} C_{k-1}C_{n-k}.$$

This recurrence allows to write a quadratic equation for the generating function $C(x) := \sum_{n=0}^{\infty} C_n x^n$. Solving this quadratic equation allows us to find the Catalan numbers. (To complete this proof is a bonus problem.)

3 Parenthesizations

Consider a product $x_1 \cdots x_n$ of n variables. How many ways are there to put parentheses that unambiguously indicate the order to perform the multiplications? For example, for $n = 3$ the answer is 2: we may write $x_1(x_2x_3)$ or $(x_1x_2)x_3$. For general n the answer is the Catalan number C_{n-1}. This is easily seen by converting the problem to counting the number of ways to enter the product into a *reverse polish notation* (RPN) calculator. An RPN calculator has no equal sign, nor parentheses, only an ENTER (E) button and operation buttons. Hitting E after typing a number stores the on the top of the stack, hitting the multiplication symbol $\times$ replaces the top two numbers in the stack with their product. Thus to perform $(x_1(x_2x_3))x_4$, we would enter $x_1Ex_2Ex_3E \times \times x_4E\times$ in an RPN calculator. This string is uniquely described by listing only the symbols E and $\times$: the above parenthesization is associated to the string $EEE \times \times E\times$. A string of symbols E and $\times$ describes a valid parenthesization of $x_1 \cdots x_n$ if it contains n copies of E, $n - 1$ copies of $\times$, and in every initial segment of the string the number of symbols E exceeds the number of symbols $\times$ by at least one. Indeed, the difference between the number of symbols E and the number of symbols $\times$ in an initial segment is the current number of numbers stored in the stack, which can never be less than 1 after entering the first number.

We have shown that the number of valid parenthesizations of $x_1 \cdots x_n$ is the same as the number of lattice paths from $(0, 0)$ to $(2n - 1, 1)$ containing n northeast steps and $n - 1$ southeast steps such that the lattice path never goes below the line $y = 1$ after the first step (which is a northeast step). Clearly the number of such lattice paths is the same as the number of lattice paths from $(0, 0)$ to $(2n - 2, 0)$ that never goes below the x axis. As seen in Section 1, the answer is C_{n-1}.

4 Counting circular arrangements

Consider a circular arrangement of $n+1$ copies of 1 and n copies of -1. Associate to each 1 a northeast step and to each -1 a southeast step. It is not difficult to show that there is exactly one 1 such that starting to read the circular arrangement from there, and reading all numbers exactly once we obtain a lattice path from $(0, 0)$ to $(2n + 1, 1)$ that never goes below the line $y = 1$ after the first step. (To prove this is a bonus question). There are $\binom{2n}{n}$ ways to write a linear arrangement of $n+1$ copies of 1s and n copies of -1 that begins with a 1. Consider two such linear arrangements equivalent if they can be obtained by breaking up the same circular arrangement. Each equivalence class has $n + 1$ linear arrangements in it, and there is exactly one linear arrangement in each class that encodes a lattice path from $(0, 0)$ to $(2n + 1, 1)$ that never goes below the line $y = 1$ after the first step. This argument provides yet another proof of $C_n = \binom{2n}{n}/(n + 1)$.