MATH 320: COMBINATORICS.

COUNTING PATHS
VICTOR H. MOLL

The goal of these notes is to describe the counting of paths explained in class. In these problems one usually has a starting point and an ending point and some description of what the path is suppose to do.

`Example 1`. Suppose we want to count all the paths going from $(0,0$ to (a,b), with a, $b \in \mathbb{N}$. The possible steps are $(i,j) \mapsto (i+1,j)$ (moving one unit to the right) and $(i,j) \mapsto (i,j+1)$ (moving one unit up). Then, in order to get to (a,b), you must take $a+b$ steps. Of these a must be horizontal and b vertical. Now mark the steps as s_1, $s_2, \cdots, s_{a+b}$. You must choose a of them, which must be horizontal. Once you do that, the rest is completely decided. Therefore, the number of such paths is

$$\binom{a+b}{a}.$$

`Example 2`. This is more complicated. In this case we take paths with steps $(i,j) \mapsto (i+1,j+1)$ (this is diagonally up, it will be called NE). The other steps are $(i,j) \mapsto (i+1,j-1)$, called SE. The goal is to count all the paths from $(0,0)$ to $(n,0)$; that is, the path will start and end on the x-axis. In order to make the problem more interesting, we will also require that the path must stay above the x-axis.

The first observation is that the numbers of NE steps must be the same as the number of SE steps, since the ups and down must be same. Therefore the total number of steps `must be even`. At this point we need to introduce a variable: let

$$C_n = \text{ the number of } \texttt{legal} \text{ paths starting at } (0,0) \text{ and ending at } (n,0).$$

A legal path is one composed by the two type of steps above and staying above the x-axis, with the appropriate initial and ending points. Therefore $C_n =$ if n is odd.

In order to get an expression for C_n we will divide them into *disjoint classes*. `This is not easy`. It turns out that it is convenient to define the set

$$\mathcal{O}_k = \text{legal paths with first hit of the } x \text{ axis at position } k. \tag{1}$$

The possible choices of k are 2, 4, $\cdots, n-2$, n. It should be clear that each legal path is precisely in one of the classes $\mathcal{O}_k$. The addition principle states

Date: September 18, 2019.

that

$$C_n = \sum_{k=2}^{n} |\mathcal{O}_k| , \tag{2}$$

where $|X|$ denotes the number of elements in the set X.

The question is now to determine how many elements are in the set $\mathcal{O}_k$. Any such path starts at $(0,0)$, stays above the x-axis until it reaches $(k,0)$ (this is the first time it hits this axis) and then is followed by an arbitrary legal path starting at $(k,0)$ and ending at $(2n,0)$. The total number of these type of paths is now obtained by multiplying the number of choices for the first part with the number of choices of the second type. The count of the second type is easy: move any such legal path k units to the left. Then you get a legal path starting at $(0,0)$ and ending at $(n-k,0)$. (Recall that both n and k are even. Therefore the total number of the second type is simply C_{n-k}. The first type of paths are not all the legal ones. Here we have restricted the positions, except the first and last, to be *above* the axis. `Now comes a good idea`: ignore the first and last step and treat the level one above the x-axis as the new horizontal axis. Then you get all possible `legal paths` of length $k-2$. Therefore the number of paths of the first kind is C_{k-2}. In order to convince yourself of this, do the case $n=10$. This gives the recurrence

$$C_n = \sum_{k=2}^{n} C_{k-2} C_{n-k}. \tag{3}$$

For example, when $n=2$, we must get $C_2 = 1$. Replacing in (3) gives

$$C_2 = C_0^2. \tag{4}$$

Therefore $C_0 = 1$ is imposed by the recurrence. Think that is also consistent with the definition. *It is a little strange.* There is one such path: start at $(0,0)$ and do not move.

Since n and k are even, write $n=2m$ and $k=2j$. Then (3) becomes

$$C_{2m} = \sum_{j=1}^{m} C_{2j-2} C_{2m-2j} \tag{5}$$

and now only `even indices` are involved. Actually (3), with $C_1 = 0$, will show you that $C_{odd} = 0$, as you expect.

One can use the recurrence to produce the values

$$C_2 = 1,\ C_4 = 2,\ C_6 = 5,\ C_8 = 14,\ C_{10} = 42, \tag{6}$$

and the OEIS site gives the `Catalan numbers` as the first hit. There are many other sites, but if you use the recurrence to generate enough values, you will see that the only possible choice is

$$C_{2n} = \frac{1}{n+1}\binom{2n}{n}. \tag{7}$$

The rest of the notes proves this. One needs to have an idea: a common one is to introduce the generating function

$$f(x) = \sum_{n=0}^{\infty} C_n x^n. \tag{8}$$

Note that we have included also the values C_n with n odd. In other problems one does not have extra information on the coefficients.

The form of (3) suggests to compute the square of $f(x)$. This might not be clear at first, but experience helps.

Suppose you have two series

$$A(x) = \sum_{k=0}^{\infty} a_k x^k \quad \text{and} \quad B(x) = \sum_{j=0}^{\infty} b_k x^j \tag{9}$$

and then you multiply them

$$\begin{aligned} A(x)B(x) &= \left(\sum_{k=0}^{\infty} a_k x^k\right) \times \left(\sum_{j=0}^{\infty} b_j x^j\right) \\ &= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} a_k b_j x^{j+k}. \end{aligned}$$

Think about all the points (k, j) that you need to sum over. They form a `lattice` of points on the unit quadrant with integer coefficients. The point is that, since we see the expression $j + k$, it might be convenient to use this as a new variable. Let

$$r = j + k \tag{10}$$

and then observe that r runs from 0 to ∞. When r has a fixed value, say $r = 10$, then you are covering a line with slope -1. The possible choices of k are $0, 1, \cdots, r$ and then j is determined as $j = r - k$. Therefore

$$\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} a_k b_j x^{j+k} = \sum_{r=0}^{\infty}\left(\sum_{k=0}^{r} a_k b_{r-k}\right) x^r. \tag{11}$$

This shows that the coefficient of $A(x)B(x)$ of x^r is

$$\sum_{k=0}^{r} a_k b_{r-k}. \tag{12}$$

This is called the `convolution` of the sequences $\{a_k\}$ and $\{b_j\}$. Observe the similarity with the convolution for continuous functions

$$(u * v)(x) = \int_0^x u(y)v(x-y)\,dy. \tag{13}$$

The calculation above shows that the square of the generating function $f(x)$ is given by

$$f^2(x) = \sum_{r=0}^{\infty} \left(\sum_{k=0}^{r} C_k C_{r-k} \right) x^r. \tag{14}$$

This looks very much like (3). To put in the right form, we need to start the sum at $k = 2$. This is easy to fix:

$$\sum_{k=0}^{r} C_k C_{r-k} = \sum_{k=2}^{r+2} C_{k-2} C_{r+2-k} \tag{15}$$

And now you can use (3) to write the right-hand side of (15) as C_{r+2}. `Think about this`. Then

$$f^2(x) = \sum_{r=0}^{\infty} C_{r+2} x^2. \tag{16}$$

Now we need to write the right-hand side of (16) in terms of f. Observe that

$$\begin{aligned} f^2(x) &= \sum_{r=0}^{\infty} C_{r+2} x^r \\ &= x^{-2} \sum_{r=0}^{\infty} C_{r+2} x^{r+2} \\ &= x^{-2} \sum_{m=2}^{\infty} C_m x^m. \end{aligned}$$

In this last sum is f except that the first two terms are missing. Using $C_0 = 1$ and $C_1 = 1$ it gives

$$f^2(x) = x^{-2} \left(f(x) - 1 \right). \tag{17}$$

This simplifies to

$$x^2 f^2(x) - f(x) + 1 = 0. \tag{18}$$

This is a quadratic equation that gives two choices (from the $\pm$):

$$f(x) = \frac{1 \pm \sqrt{1 - 4x^2}}{2x^2}. \tag{19}$$

Since $f(0) = C_0 = 1$, the correct sign must be the negative one

$$f(x) = \frac{1 - \sqrt{1 - 4x^2}}{2x^2}. \tag{20}$$

Now remember that

$$f(x) = \sum_{n=0}^{\infty} C_n x^n, \tag{21}$$

and since (20) has only even powers of x, you see again that $C_{odd} = 0$. It is nive when things work out to something you know.

To obtain a formula for C_n, we need a formula for $\sqrt{1-4x^2}$. We start with $\sqrt{1+t}$, expand and then replace t by $-4x^2$. `Let's see what comes out of this`.

Start with the binomial theorem

$$(1+t)^a = \sum_{j=0}^{a} \binom{a}{j} t^j \tag{22}$$

and then put $a = 1/2$. The upper limit of the sum you can replace by ∞, since in the case when a is a positive integer, the added terms (with binomials) are zero. Then, first when $a \in \mathbb{N}$,

$$\binom{a}{j} = \frac{a!}{j!\,(a-j)!} = \frac{a(a-1)(a-2)\cdots(a-j+1)}{j!} \tag{23}$$

and note that there are j factors on top and bottom of the last fraction. Then, when $a = \frac{1}{2}$,

$$\tfrac{1}{2}\left(\tfrac{1}{2}-1\right)\left(\tfrac{1}{2}-2\right)\cdots\left(\tfrac{1}{2}-j+1\right) = (-1)^{j-1}\tfrac{1}{2}\cdot\tfrac{1}{2}\cdot\tfrac{3}{2}\cdots\tfrac{2j-3}{2} \tag{24}$$

The denominator is 2^j, since there are j factors. The numerator is the product of the odd numbers from 1 to $2j-3$. In order to write it as factorials, `here is a nice trick`:

$$\begin{aligned} 1\cdot 3\cdots(2j-3) &= \frac{1\cdot 2\cdot 3\cdot 4\cdots(2j-4)\cdot(2j-3)}{2\cdot 4\cdot 6\cdots(2j-4)} \\ &= \frac{1\cdot 2\cdot 3\cdot 4\cdots(2j-4)\cdot(2j-3)}{(2\cdot 1)\cdot(2\cdot 2)\cdot(2\cdot 3)\cdots(2\cdot(j-2))} \\ &= \frac{(2j-3)!}{2^{j-2}(j-2)!} \end{aligned} \tag{25}$$

This gives, after a small adjustment,

$$\binom{\frac{1}{2}}{j} = \frac{(-1)^{j-1}(2j-2)!}{2^{2j-1}j!\,(j-1)!} \tag{26}$$

for $j \geq 1$. Then

$$(1+t)^{1/2} = 1 + \sum_{j=1}^{\infty} \frac{(-1)^{j-1}(2j-2)!}{2^{2j-1}j!\,(j-1)!} t^j. \tag{27}$$

Now replace t by $-4x^2$ to get

$$(1-4x^2)^{1/2} = 1 - 2\sum_{j=1}^{\infty} \frac{(2j-2)!}{j!\,(j-1)!} x^{2j}. \tag{28}$$

Now use (20) to get

$$
\begin{aligned}
f(x) &= \frac{1-\sqrt{1-4x^2}}{2x^2} \\
&= \sum_{j=1}^{\infty} \frac{(2j-2)!}{j!(j-1)!} x^{2j-2} \\
&= \sum_{j=0}^{\infty} \frac{(2j)!}{(j+1)!\,j!} x^{2j} \\
&= \sum_{j=0}^{\infty} \frac{1}{j+1}\binom{2j}{j} x^{2j}.
\end{aligned} \tag{29}
$$

This proves $C_{odd} = 0$ and

$$
C_{2j} = \frac{1}{j+1}\binom{2j}{j}. \tag{30}
$$

FINALLY, this proves (7).

The numbers C_{2n} are called `Catalan numbers`. They will come back.