April 3rd, 2018

Lecture 23

The first hour in the lecture was spent solving the Fibonacci recurrence relation

$$a_n = a_{n-1} + a_{n-2} \quad , \; n \geq 2, \quad a_0 = 1, \; a_1 = 1$$

(Equivalently, $a_n = a_{n-1} + a_{n-2}$, $n \geq 3$, $a_1 = 1$, $a_2 = 2$)

in two different ways:

(a) Via characteristic equations

(b) Via generating functions and their functional equations

We had done these earlier in the class lectures without giving details. The details are available in the notes for Lecture 21 on March 25. On pages 8 and 9 the problem is done using characteristic equations and on pages 14 and 15 it is done using generating functions and functional equations. Please look at these four pages in the notes for the March 25 class (i.e., the notes for Lecture 21).

The rest of the last lecture was spent understanding Catalan numbers in a different way. We begin the discussion on that in the next page.

Catalan numbers:

Recall that for $n \geq 0$, the n^{th} Catalan number is

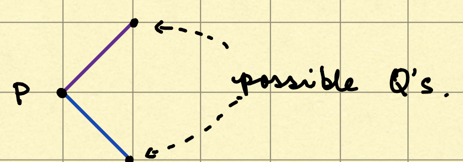
$$\frac{1}{n+1}\binom{2n}{n}$$

and that this is the solution of the recurrence relation

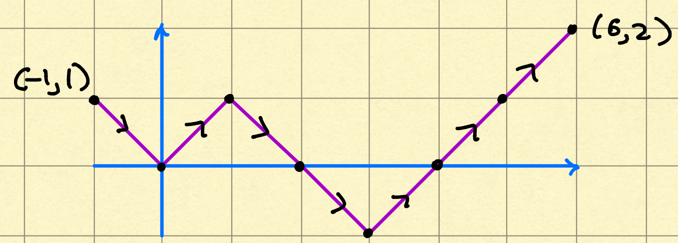
$$a_n = \sum_{i=0}^{n-1} a_i a_{n-1-i} \quad (n \geq 1), \quad a_0 = 1.$$

The techniques used were functional equations and Taylor's series formula for $(1+y)^{1/2}$. Let us give a combinatorial interpretation for Catalan numbers as well as for the above recurrence relation.

Walks: We will be considering discrete walks on the x-y plane each step of which is a move from an integer point (m,n) ((m,n) is an integer point if both m and n are integers) to either $(m+1, n+1)$ or to $(m+1, n-1)$. This means each step is a move from an integer point P to an integer point Q in either the north east direction or in the south east direction by a distance of $\sqrt{2}$.



One step or move in the walk.



A walk from $(-1,1)$ to $(6,2)$.

Observations

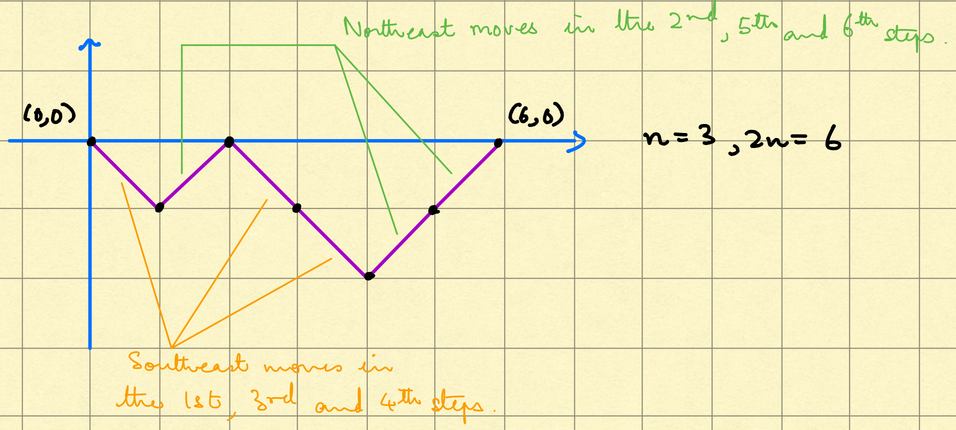
- Suppose a walk begins at $(0,0)$. The number of steps taken (m,n) is m, regardless of what the second coordinate n is. More generally, the number of steps taken to go from (p,q) to (m,n) is $m-p$ regardless of what q or n are. This is because with every move the x-coordinate definitely increases by 1, whereas the y-coordinate could increase or decrease by 1.
- Suppose a walk begins at $(0,0)$ and ends at $(0,r)$, then r must be even. This is because the only way one can reach $(0,r)$ from $(0,0)$ is if the number of northeast steps equals the number of southeast steps. In particular $r=2k$. More generally if a walk begins at (m,n) and ends at (m,q) (x-coordinates of both points the same) then the walk has an even number of steps.
- If a walk begins at $(0,0)$ and ends at (n,k) then $n+k$ must be even, since at each stage we are increasing the x-coordinate by one and either increasing or decreasing the x-coordinate by one: if $x+y$ is even then $(x+1)+(y+1)$ as well as $(x+1)+(y-1)$ is even. Since we are starting at $(0,0)$, the assertion is clear.

- If we walk from $(0,0)$ to (p,q) then $-p \leq q \leq p$. This is because in the extreme case of always moving northeast, in p steps we arrive at (p,p), and at the other extreme (always southeast) we end up at $(p,-p)$ in p steps, and any other walk of p-steps will be between these extremes.
- The number of walks from $(0,0)$ to $(2n,0)$ is
$$\binom{2n}{n}.$$

This is seen as follows: any walk from $(0,0)$ to $(2n,0)$ must have an equal number of northeast and southeast moves, and therefore each is n in number. If the southeast moves are made in the $i_1^{th}, i_2^{th}, \ldots, i_n^{th}$ steps then $\{i_1, i_2, \ldots, i_n\}$ is a choice of n numbers from $\{0,1,2,\ldots,2n\}$. Conversely, suppose we pick n numbers from $\{0,1,2,\ldots,2n\}$ say $i_1, i_2, \ldots, i_n$, then one can construct a walk with $2n$ steps in which the southeast moves are at the $i_1^{th}, i_2^{th}, \ldots, i_n^{th}$ step and in the remaining n steps are northeast moves, and this walk will have to terminate at $(2n, 0)$.

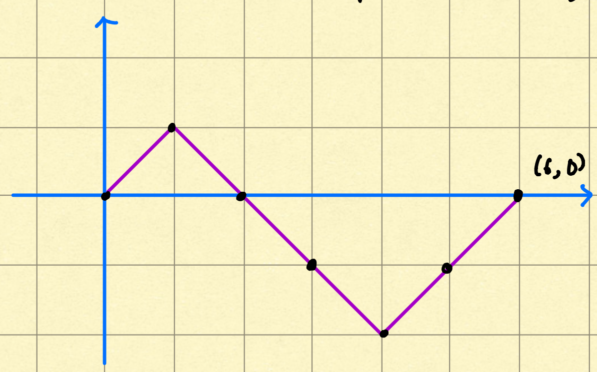
For example suppose $n=3$. Let us pick three numbers from the set $\{1,2,3,4,5,6\}$

Say we pick 1, 3, 4. Then the walk from $(0,0)$ to $(6,0)$ is :



$n=3$, $2n=6$

If instead we had picked 2, 3, 4, the walk would be:

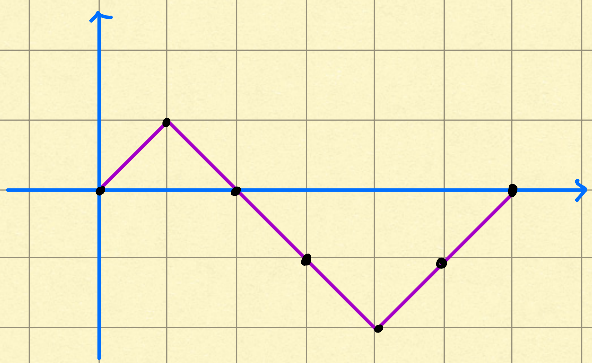


$n=3$, $2n=6$

Fix n. Call a walk a good walk if it is from $(0,0)$ to $(2n, 0)$ and never dips below the x-axis (it is allowed to touch the x-axis). A walk from $(0,0)$ to $(2n, 0)$ is a bad walk if it does dip below the x-axis.

Picture examples are given below.

A good walk (never dips below x-axis)



A bad walk (dips below the x-axis)

Let $\quad C_n = \#$ of good walks from $(0,0)$ to $(2n, 0)$.

It is obvious that

$$C_n = \text{Total \# of walks from } (0,0) \text{ to } (2n,0) - \text{\# of bad walks from } (0,0) \text{ to } (2n,0)$$

$$= \binom{2n}{n} - \text{\# of bad walks from } (0,0) \text{ to } (2n,0).$$

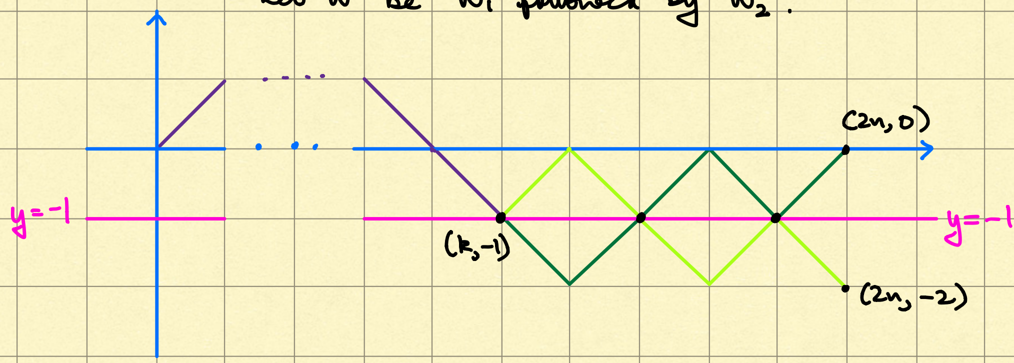
Let us count the number of bad walks.

Every bad walk from $(0,0)$ to $(2n,0)$ has to touch the line $y=-1$ at least once. Suppose the first time a bad walk W does so is at $(k,-1)$.

Then the bad walk W can be split into two walks:

A walk W_1 from $(0,0)$ to $(k,-1)$ and a walk W_2 from $(k,-1)$ to $(2n,0)$. One can reflect the walk W_2 around the line $y=-1$ to get a new walk W_2'.

Let W' be W_1 followed by W_2'.



W_1 = Purple walk

W_2 = Dark green walk

W_2' = Light green walk

$W = W_1 + W_2$

$W' = W_1 + W_2'$

The walk W_2' ends at $(2n,-2)$ and hence the walk W' ends at $(2n,-2)$.

Conversely, suppose one has a walk W' from $(0,0)$ to $(2n,-2)$. It has to touch the line $y=-1$, and suppose the first time it does is at $(k,-1)$. Then $W' = W_1 + W_2'$ where W_1 is a walk from $(0,0)$ to $(k,-1)$ and the W_2' is a walk $(k,-1)$ to $(2n,-2)$. Reflect W_2' about the line $y=-1$ and obtain a walk W_2 from $(k,-1)$ to $(2n,0)$. Call this walk W_2. Set $W = W_1 + W_2$.

Since W passes through $(k,-1)$ it must be a bad walk from $(0,0)$ to $(2n,0)$.

The above considerations show that

# of bad walks from $(0,0)$ to $(2n,0)$ = # of walks from $(0,0)$ to $(2n,-2)$.

Now a walk from $(0,0)$ to $(2n,-2)$ has $2n$ steps and the number of southeast moves must be two more than the number of northeast moves (so that the ending y-coordinate is -2). It follows that such a walk has $n+1$ southeast moves and $n-1$ northeast moves. This means the the number of walks from $(0,0)$ to $(2n,-2)$ is equal to $\binom{2n}{n+1}$ (which equals $\binom{2n}{n-1}$).

Thus

# of bad paths from $(0,0)$ to $(2n,0)$ = $\binom{2n}{n-1}$

This means

$$C_n = \binom{2n}{n} - \binom{2n}{n-1}.$$

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

Proof of claim:

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

$$= \frac{(2n)!}{n!\,(n-1)!}\left\{\frac{1}{n} - \frac{1}{n+1}\right\}$$

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

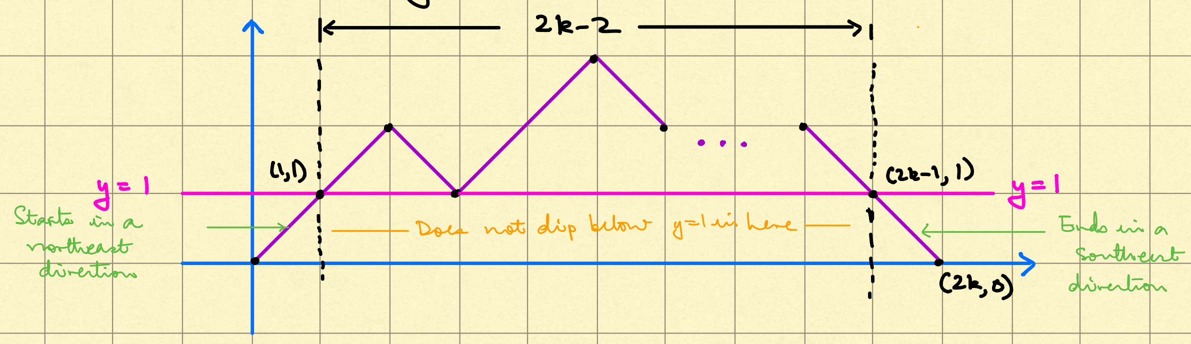
q.e.d.

Thus C_n is the $n^{\underline{th}}$ Catalan number.

__Warning__: In the book (see problem 31, Section 5-5) the notation for the n^{th} Catalan number is C_{2n}. But the usual convention in other books is C_n.

__Recursion relation for Catalan numbers:__

Suppose we have a good walk from $(0,0)$ to $(2k,0)$ such that the path never touches the x-axis except at $(0,0)$ and $(2k,0)$, i.e., the walk touches the x-axis only at the extremities.



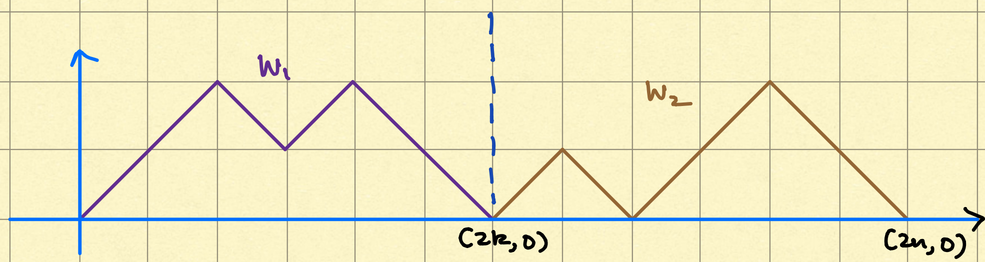
Such a good walk must start in the northeast direction end in the southeast direction and from $(1,1)$ to $(2k-1,1)$

it must stay above the line $y=1$ (it is allowed to touch it). It follows that

$$\#\left[\begin{array}{l}\text{Good walks from } (0,0) \text{ to } (2k,0) \text{ which} \\ \text{don't touch the } x\text{-axis except at end} \\ \text{points}\end{array}\right] = C_{k-1}$$

Now suppose we have a good walk W from $(0,0)$ to $(2n,0)$. The first place it touches the x-axis after $(0,0)$ has to have an even number of steps (since number of northeast moves must equal the number of southeast moves) involved, and so this first place has coordinates $(0,2k)$ where k is some number from 1 to n. This breaks up W into two paths: $W = W_1 + W_2$ where W_1 is a good walk from $(0,0)$ to $(2k,0)$ which never touches the x-axis except at $(0,0)$ and $(2k,0)$, and W_2 is a walk from $(2k,0)$ to $(2n,0)$ which never dips below the x-axis. W_2 can be thought of as a good walk from $(0,0)$ to $(2n-2k,0)$ by shifting the origin from $(0,0)$ to $(2k,0)$.

The number of possibilities for W_1 from our arguments earlier is C_{k-1}, and the number of possibilities for W_2 is C_{n-k}. (See picture on the next page.)



Since k can be any number from 1 to n, we get

Recurrence relation for Catalan numbers $\longrightarrow$ $\boxed{C_n = \sum_{k=1}^{n} C_{k-1} C_{n-k}}$ $\quad n \geq 1$

Note that $C_0 = 1$. So how does the above recurrence relation compare with the one we had earlier for Catalan numbers, namely

$$(*) \longrightarrow \quad a_n = \sum_{i=0}^{n-1} a_i a_{n-1-i} \quad (n \geq 1), \quad a_0 = 1 \quad ?$$

Set $i = k-1$ in the recurrence relation $C_n = \sum_{k=1}^{n} C_{k-1} C_{n-k}$.
The change of variables yields

$$C_n = \sum_{i=0}^{n-1} C_i C_{n-i-1}; \quad n \geq 1, \quad C_0 = 1.$$

This means $\{C_n\}$ is a solution of $(*)$.

Note that $(*)$ is not a linear recurrence relation. We therefore cannot use characteristic equations. In fact the notion of a characteristic equation make no sense for non-linear recurrence relation. The best way of solving this is via generating functions and functional equations as we did in earlier lectures (Lectures 21 and 22).