18.212 Lecture 2 02/19/2021

last time: The Catalan numbers

$$C_n := \#\text{Dyck paths of length } 2n$$

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

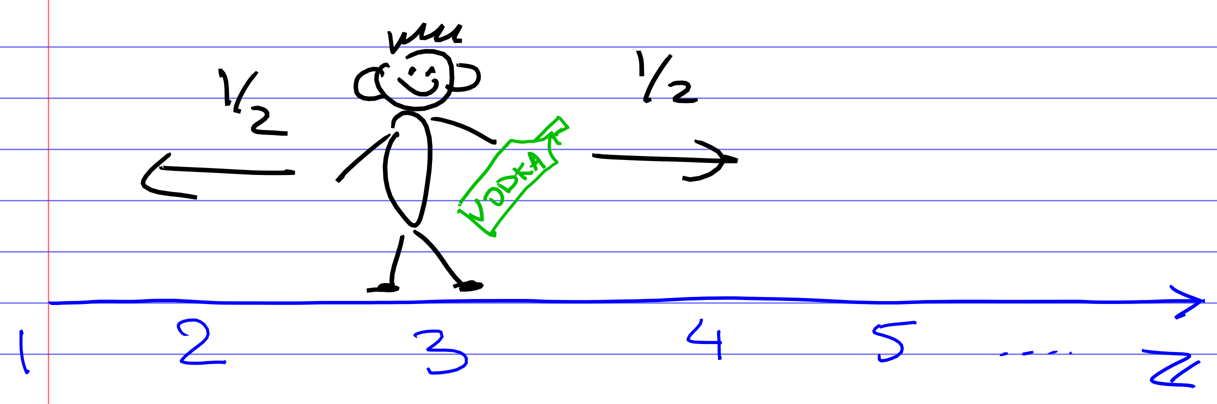
generating function:

$$C(x) := C_0 + C_1 x + C_2 x^2 + \dots$$

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

Application: Drunkard's Walk

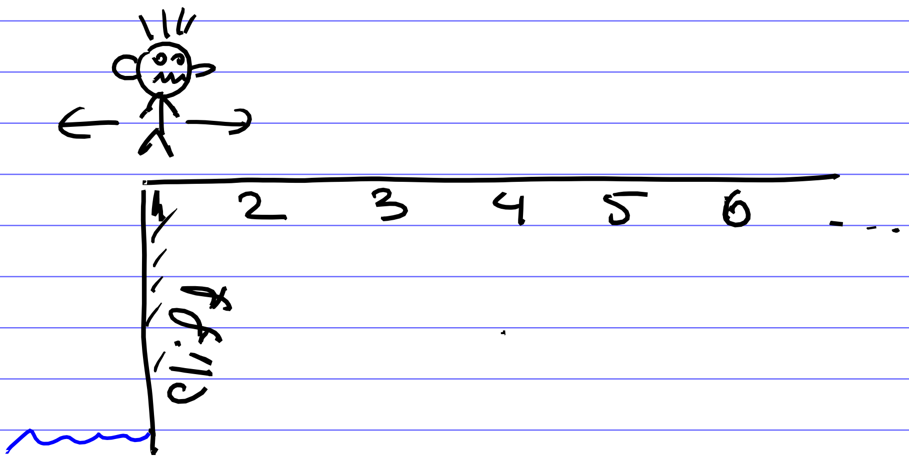
A simple random walk on the integer line $\mathbb{Z}$.



The drunkard makes a step to the right with probability $1/2$ or a step to the left with probability $1/2$, and continues randomly walking like this.

Now assume that there is a cliff at $x = 0$. Initially, the drunkard starts at $x = 1$ (right at the edge of the cliff)

If he steps off the cliff after some number of steps, he falls.



<u>Problem</u>: What is the probability that the drunard does not fall off the cliff?

Let's find all possibilities for the drunkard to fall off the cliff

$$\text{Prob}\binom{\text{falls off}}{\text{the cliff}} = \overset{\leftarrow}{\frac{1}{2}} +$$

$$+ \overset{\rightarrow}{\frac{1}{2}} \cdot \overset{\leftarrow}{\frac{1}{2}} \cdot \overset{\leftarrow}{\frac{1}{2}} + \overset{\rightarrow}{\frac{1}{2}} \overset{\leftarrow}{\frac{1}{2}} \cdot \overset{\rightarrow}{\frac{1}{2}} \cdot \overset{\leftarrow}{\frac{1}{2}} \overset{\leftarrow}{\frac{1}{2}}$$

$+ \ldots$

all these possibilites correspond to Dyck paths with one extra left step in the end.

$$= \frac{1}{2} C_0 + \left(\frac{1}{2}\right)^3 C_1 + \left(\frac{1}{2}\right)^5 C_2 + \ldots$$

the Catalan numbers

$$= \frac{1}{2} C\left(\frac{1}{4}\right).$$

Let us now plug this into the expression $C(x) = \frac{1-\sqrt{1-4x}}{2x}$

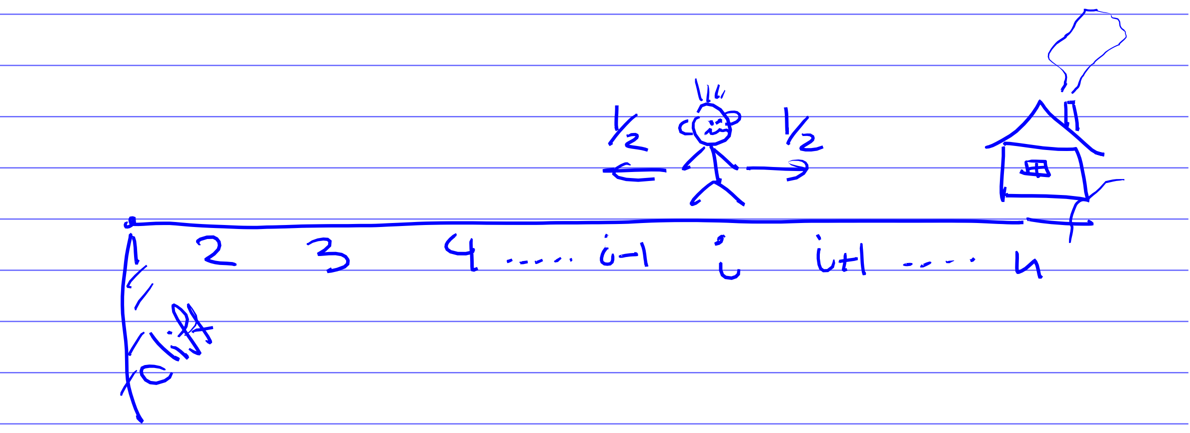
$$\text{Prob}\binom{\text{falls of}}{\text{the cliff}} = \frac{1}{2} \frac{\left(1-\sqrt{1-\frac{4}{4}}\right)}{2 \cdot \frac{1}{4}} = 1$$

Not very good news for the drunkard.

<u>Proposition</u>. $\text{Prob}\binom{\text{drunkard}}{\text{falls}} = 1.$

There is a simpler solution of the drunkard's walk problem that does not use the Catalan numbers.

Let's give the drunkard some hope. Assume that his house is located at $x=n$ (and the cliff is still at $x=0$)



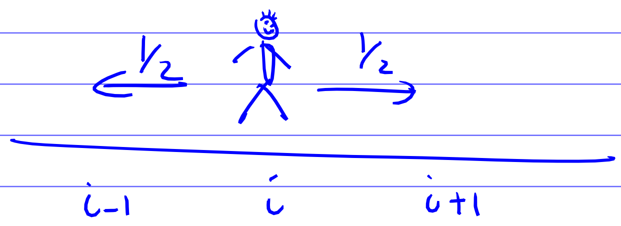
- the drunkard starts at $x=i \in \{0, 1, 2, \dots, n\}$.
  - If he first reaches $x=0$ (cliff) he falls.
  - If he first reaches $x=n$ (house), he goes to sleep.

Let p_i be the probability that the drunkard reaches the house before reaching the cliff.

The probabilities $p_0, p_1, \dots, p_n$ satisfy the relations:

- For $1 \leq i \leq n-1$,

$$\boxed{p_i = \frac{1}{2} p_{i-1} + \frac{1}{2} p_{i+1}}$$



- $\boxed{p_0 = 0}$ the drunkard starts at $x=0$ & falls right away
- $\boxed{p_n = 1}$ the drunkard starts at the house $(x=n)$ & goes to sleep right away

It is not hard to show that this system of linear equations for $p_0, p_1, \ldots, p_n$ has a unique solution:

$$p_i = \frac{i}{n} \quad \text{for } i = 0, 1, \ldots, n.$$

In particular, $p_1 = \frac{1}{n}$

Taking the limit as $n \to \infty$ we obtain the solution of the original problem without the house.

$$\text{Prob}\begin{pmatrix}\text{does not} \\ \text{fall off} \\ \text{the cliff}\end{pmatrix} = \lim_{n \to \infty} p_1 =$$

$$= \lim_{n \to \infty} \left(\frac{1}{n}\right) = 0.$$

back to the Catalan numbers...

last time we proved

Theorem $C_n := \#\left\{\begin{matrix}\text{Dyck paths}\\ \text{of length } 2n\end{matrix}\right\}$

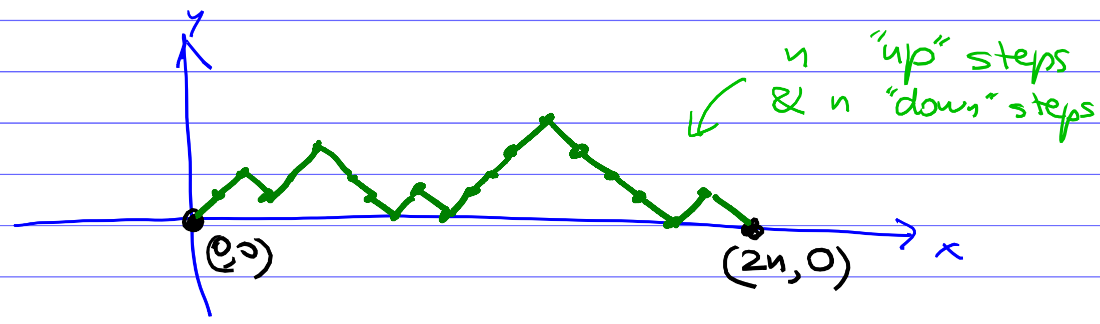
$$= \frac{1}{n+1}\binom{2n}{n}$$

using generating functions.

Let's give another proof

Proof #2: by reflection

C_n = # Dyck paths from $(0,0)$ to $(2n,0)$

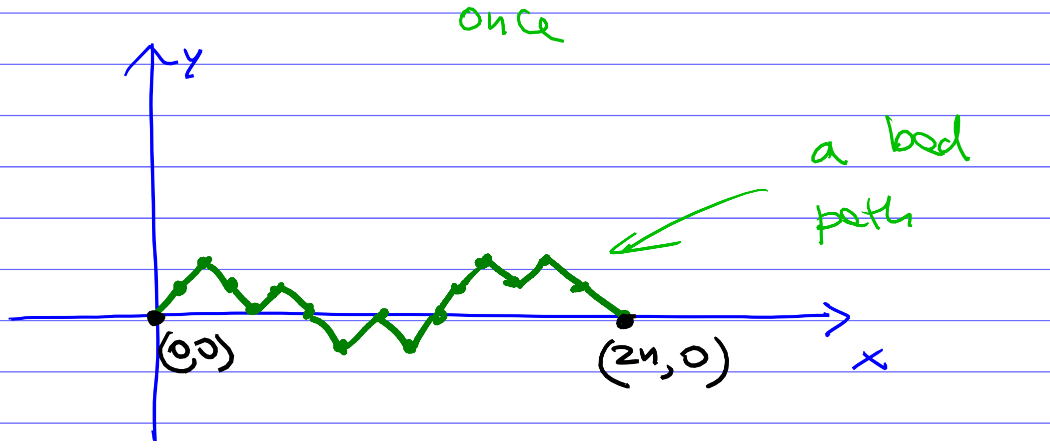


all lattice paths from $(0,0)$ to $(2n,0)$, i.e. any arrangements of n "up" & n "down" steps

$$= \binom{2n}{n}.$$

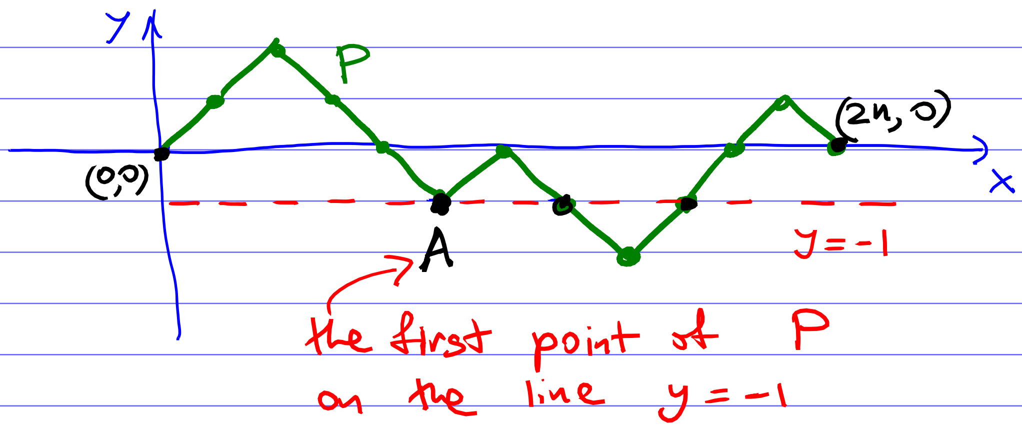
So $C_n = \binom{2n}{n} - \#\left\{\begin{matrix}\text{"bad" paths}\\ \text{from } (0,0) \text{ to } (2n,0)\end{matrix}\right\}$

bad paths are the paths that go below the x-axis at least once



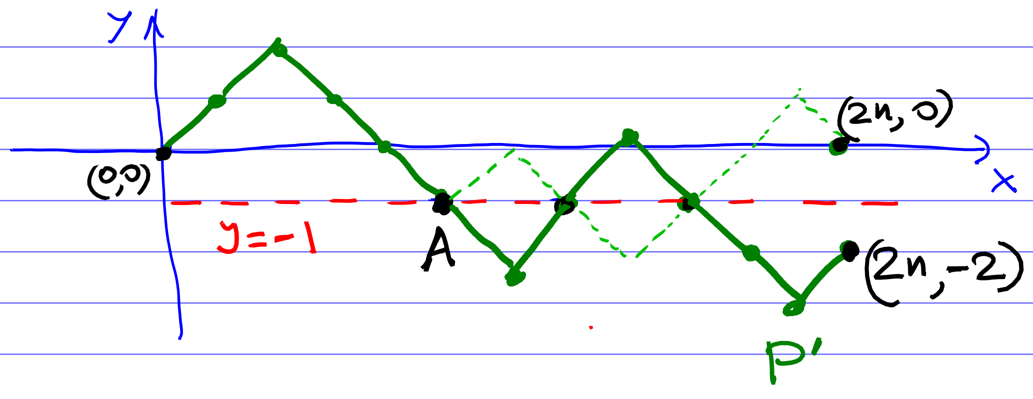
So it is enough to figure out how to count # of bad paths.

Observation. The bad paths are exactly the lattice paths from $(0,0)$ to $(2n,0)$ that contain at least one point on the line $y=-1$.



Let A be the <u>first</u> point of a bad path on the line $y=-1$.

Let P' be the path obtained by reflecting the segment of P between the points A and $(2n,0)$ around the line $y=-1$.



The map $P \mapsto P'$ is a bijection between the "bad" lattice paths from $(0,0)$ to $(2n,0)$ and <u>all</u> lattice paths from $(0,0)$ to $(2n,-2)$.

To see that this is a bijection we need to construct the inverse map $P' \mapsto P$.

The inverse map: Find the first point A of P' on the line $y=-1$ & reflect the segment of P' between A & the end-point $(2n,-2)$ around the line $y=-1$.

The lattice paths from $(0,0)$ to $(2n,-2)$ contain $(n-1)$ "up" steps & $(n+1)$ "down" steps.

So the number of such lattice paths is $\binom{2n}{n-1}$.

We deduce

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

all lattice paths from $(0,0)$ to $(2n,0)$

bad lattice paths from $(0,0)$ to $(2n,0)$

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

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

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

as needed. $\square$

This is a nice proof...

However, ideally, in combinatorics we like to find proofs that don't involve subtraction.

Is there a subtraction-free combinatorial proof of the formula $C_n = \frac{1}{n+1}\binom{2n}{n}$?

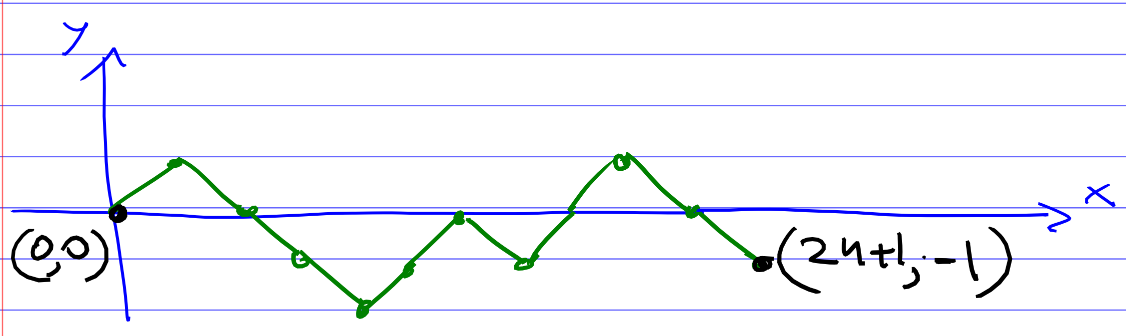
Proof #3 : cyclic shifts

Let's slightly rewrite the formula:

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

$\binom{2n+1}{n}$ = # all sequences $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{2n+1}$ with n entries $\varepsilon_i = 1$ and $n+1$ entries $\varepsilon_j = -1$.

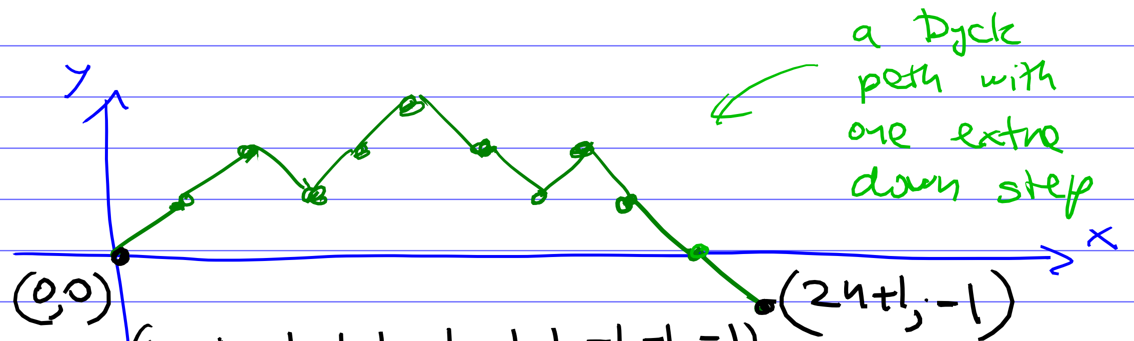
= # all lattice paths from $(0,0)$ to $(2n+1,-1)$.



$(1,-1,-1,-1,1,1,-1,1,1,-1,-1)$

C_n = # paths as above (with n up steps & $n+1$ down steps) that go below the x-axis only at the last step

these are exactly Dyck path with one extra "down" step attached in the end



$(1,1,-1,1,1,-1,-1,1,-1,-1,-1)$

Equiv. C_n = # $(\varepsilon_1, \dots, \varepsilon_{2n+1})$ with n 1's & $n+1$ (-1)'s s.t.

$$(*) \quad \left[\begin{array}{l} \varepsilon_1+\varepsilon_2+\dots+\varepsilon_i \geq 0 \\ \forall\ i=1,2,\dots,2n \end{array}\right.$$

(but $\varepsilon_1+\varepsilon_2+\dots+\varepsilon_{2n+1} = -1$)

How to see that among all sequences with n 1's & $(n+1)$ (-1)'s exactly $\frac{1}{2n+1}$ satisfy this condition?

For a sequence $(\varepsilon_1, \dots, \varepsilon_{2n+1})$ with n 1's & $(n+1)$ (-1)'s consider its $2n+1$ cyclic shifts:

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

$(\varepsilon_2, \varepsilon_3, \varepsilon_4, \dots, \varepsilon_{2n+1}, \varepsilon_1)$

$(\varepsilon_3, \varepsilon_4, \varepsilon_5, \dots, \varepsilon_{2n+1}, \varepsilon_1, \varepsilon_2)$

$\vdots$

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

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

$2n+1$ cyclic shifts

Lemma. All these $2n+1$ sequences are different from each other. Exactly one of them satisfy the condition $(*)$.

Example. $n = 3$

$(1, -1, -1, 1, 1, -1, -1)$

$(-1, -1, 1, 1, -1, -1, 1)$

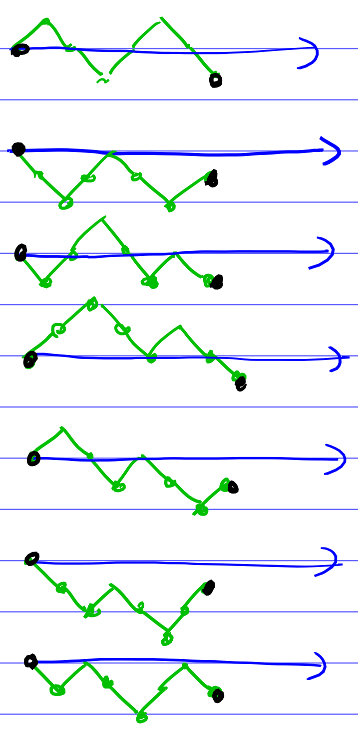
$(-1, 1, 1, -1, -1, 1, -1)$

the good one → $(1, 1, -1, -1, 1, -1, -1)$

$(1, -1, -1, 1, -1, -1, 1)$

$(-1, -1, 1, -1, -1, 1, 1)$

$(-1, 1, -1, -1, 1, 1, -1)$



Exercise. Prove this lemma.

Now we deduce that

C_n = # "good" lattice paths from $(0,0)$ to $(2n+1, -1)$

$= \frac{1}{2n+1}$ # all lattice paths from $(0,0)$ to $(2n+1, -1)$

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

as needed. $\square$

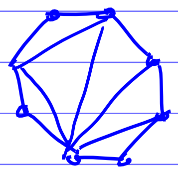
Other combinatorial interpretations of the Catalan numbers C_n:

- # Dyck paths of length $2n$

- # triangulations of an $(n+2)$-gon

Example

$n=6$



- # valid parenthesizations of $n+1$ letters

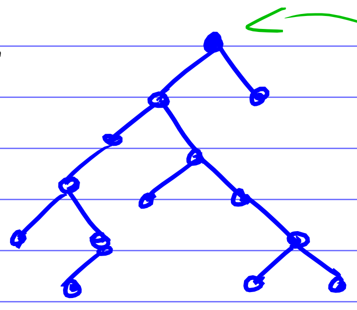
Example. $n=5$

(a ((b c)((d e) f)))

- # plane binary trees with n vertices.

Example.

$n=14$

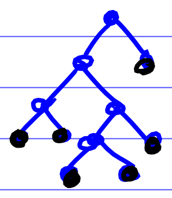


a plane binary tree has a root, & each vertex has at most 1 left child & at most 1 right child.

- # complete plane binary tree with $n+1$ leaves

Example

$n=5$



a plane binary tree is complete if each vertex v has exactly children or v is a leaf (no vertices with 1 child are allowed)

Exercise.

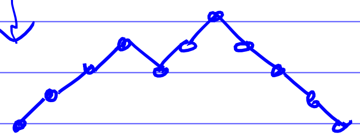
Find bijections between all these different combinatorial objects.

$$\{\text{parenthesizations}\} \xrightarrow{?} \{\text{Dyck paths}\}$$

the left parenthesis "(" $\rightsquigarrow$ "up" step

the right parenthesis ")" $\rightsquigarrow$ "down" step

$$(a((b\ c)((d\ e)\ f)))$$

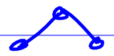


We get a Dyck path with the correct numbers of up & down steps, but ....

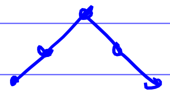
<u>Warning</u>. This is not a bijection!

<u>Example</u>.

$n = 1$ $\quad (a\ b) \rightsquigarrow$



$n = 2$ $\quad ((a\ b)\ c)$, $(a\ (b\ c))$ $\rightsquigarrow$



A bijection between parenthesizations & Dyck paths is a little bit less obvious...