The Catalan Numbers

As we have seen, there are several counting sequences which occur over and over; such as the Fibonacci + Stirling numbers. Perhaps the classiest such sequence is the Catalan numbers.

N	0	1	2	3	4	5	6	7
C_N	1	1	2	5	14	42	132	429

$\underbrace{2 \quad 5 \quad 14 \quad 42}_{\text{the mark of Catalan}}$

How many ways to multiply $N+1$ numbers with N multiplications?

$N=2$ $\quad (x_0\, x_1)\, x_2 \qquad x_0\, (x_1\, x_2) \qquad C_2 = 2$

$N=3$ $\quad x_0(x_1(x_2\, x_3)) \qquad ((x_0\, x_1)\, x_2)\, x_3 \qquad C_3 = 5$

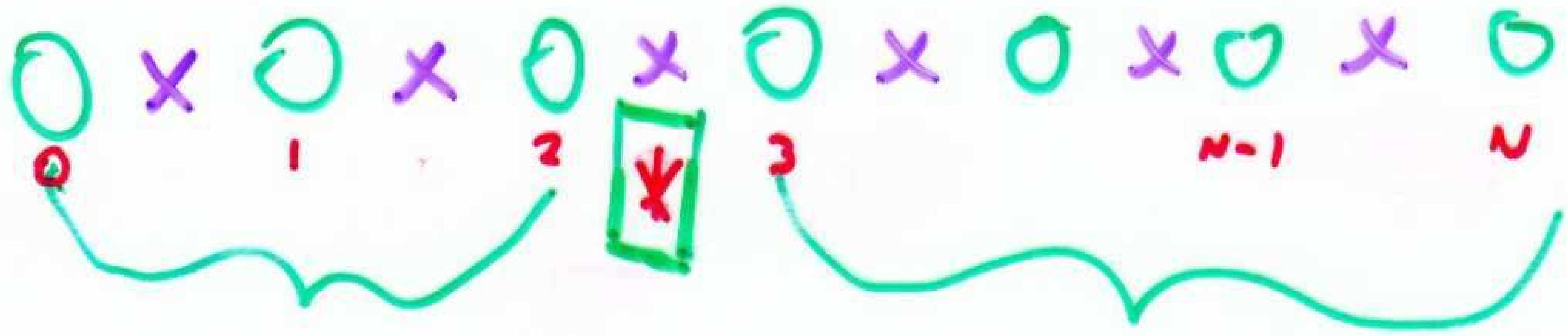
$x_0((x_1\, x_2)\, x_3) \qquad (x_0(x_1\, x_2))\, x_3$

$(x_0\, x_1)(x_2\, x_3)$

These are associated with the number of well formed parenthesations:

( )( )( ) | (( ))( ) | ( )(( )) | (( )( )) | ((( )))

Multiplying $N+1$ numbers implies N places to put a $*$. Any of them can be the outermost in the parenthesisation, which leaves two smaller strings which must be parenthesised.



$$C_N = C_0 \cdot C_{N-1} + C_1 C_{N-2} \cdots + C_{N-1} C_0$$

$$= \sum_k C_k \cdot C_{N-1-k} + (N=0)$$

This is the recurrence which describes the Catalan numbers and helps explain why they are ubiquitous - it is the simplest possible convolution.

$$C(z) = \sum_{k,N} C_k C_{N-1-k} z^N + \sum_{N=0} z^N$$

$$= \sum C_k z^k \sum C_{N-1-k} z^{N-k} + 1$$

$$= C(z) \cdot z \cdot C(z) + 1$$

Expanding this to a power series is interesting

$$z\,C(z)^2 - C(z) + 1 = 0$$

By the quadratic formula

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

It happens that $-$ is right, so $C(z) = \dfrac{1-\sqrt{1-4z}}{2z}$

To expand, use the binomial theorem:

$$\sqrt{1-4z} = \sum_{k\ge 0} \binom{1/2}{k}(-4z)^k = \sum_{n\ge 0} \frac{(1/2)^{\underline{k}}(-4)^k z^k}{k!}$$

$$= 1 + \sum_{k\ge 1} \frac{1}{2k}\binom{-1/2}{k-1}(-4z)^k$$

So, $$\frac{1-\sqrt{1-4z}}{2z} = \sum_{k\ge 1} \frac{1}{k}\binom{-1/2}{k-1}(-4z)^{k-1}$$

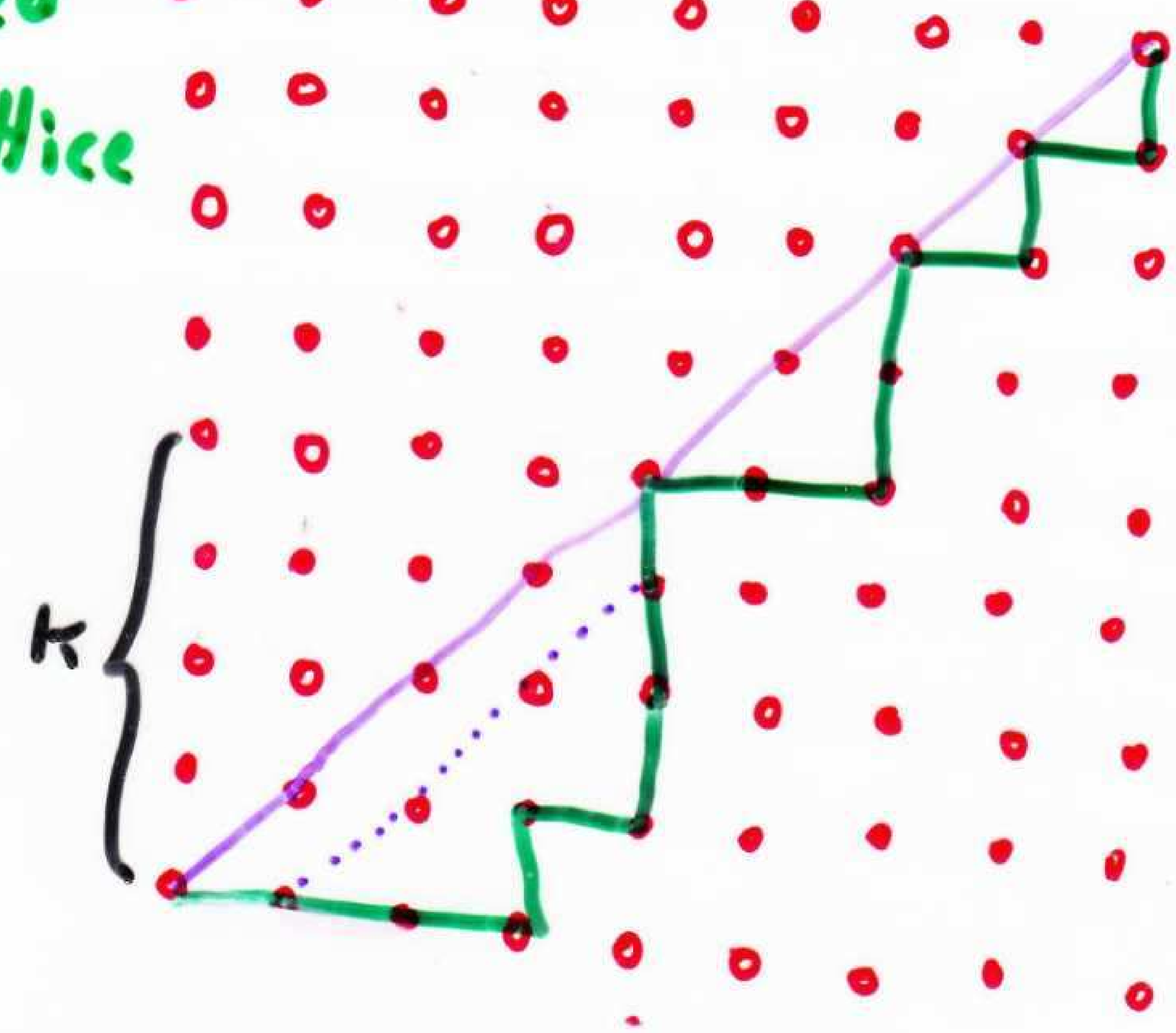
$$= \sum_{N\ge 0} \binom{-1/2}{N}\frac{(-4z)^N}{N+1} = \sum_{N\ge 0}\binom{2N}{N}\frac{z^N}{N+1}$$

Eqn 5.34 in GKP

So $C_N = \binom{2N}{N}\cdot\frac{1}{N+1}$!

How many ballot sequences leave an election tied, with one guy never behind?

This can be interpreted as the number of lattice paths that do not go above the main diagonal.



Let k be the first point of contact with the diagonal. The number of ways to get there is C_{k-1}, since no intermediate point hits the diagonal. There are C_{n-k} ways to finish, so

$$C_n = \sum_k C_{k-1} C_{n-k} + (n=0) = \sum_k C_k C_{n-k-1} + (n=0)$$

and the ballot sequences are counted by the Catalan numbers!

How many ways are there to fill in two rows of N boxes so each row & column is sorted?

1	3	4	7	9	10	11	13	17
2	5	6	8	12	14	15	16	18

Think of the numbers as being the number of a voter – the top row is the guy who never trails. This is a special case of the Young tableaux

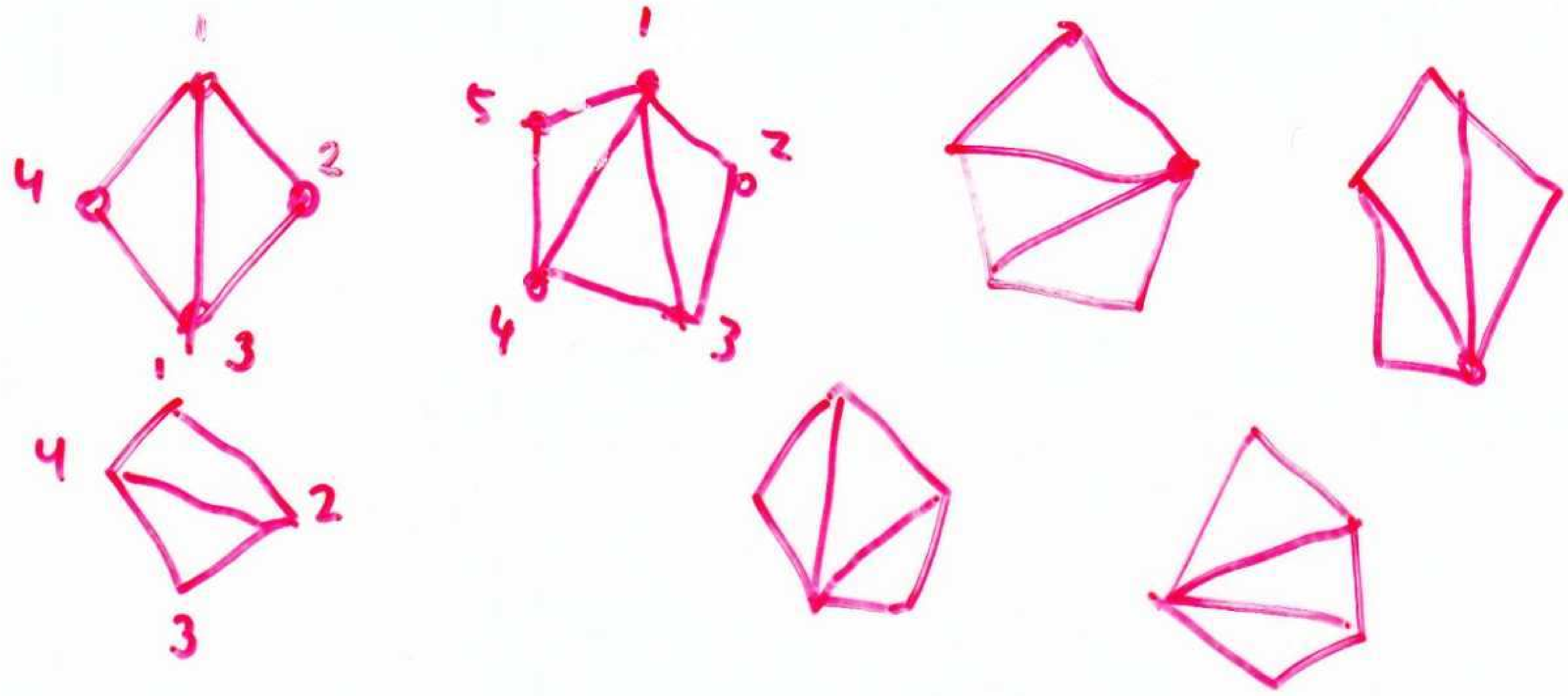
These are counted by the Catalan numbers!

How many N+M distance lattice paths between opposite corners in an N×M lattice?

These are counted by binomial coefficients!

Please keep alert! $\binom{M+N}{N}$

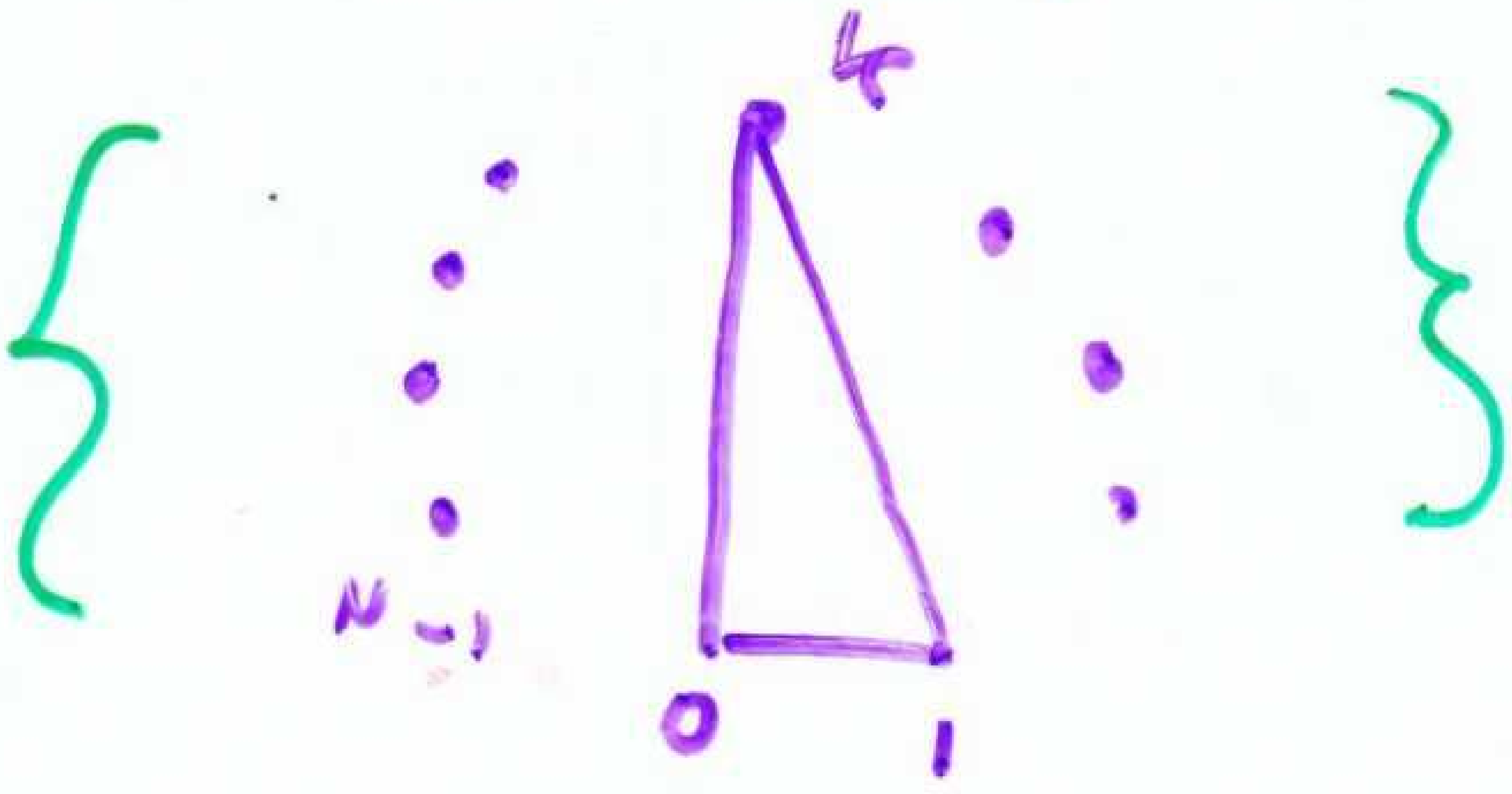
How many ways are there to triangulate a convex polygon?



$C_4 = 2$ $\qquad$ $C_5 = 5$

Consider the edge $(0, 1)$. Select the vertex which defines its triangle:

triangulate the $N-k+1$ gon

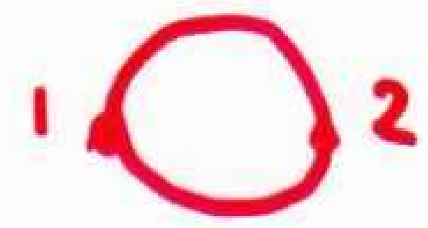


triangulate the k-gon

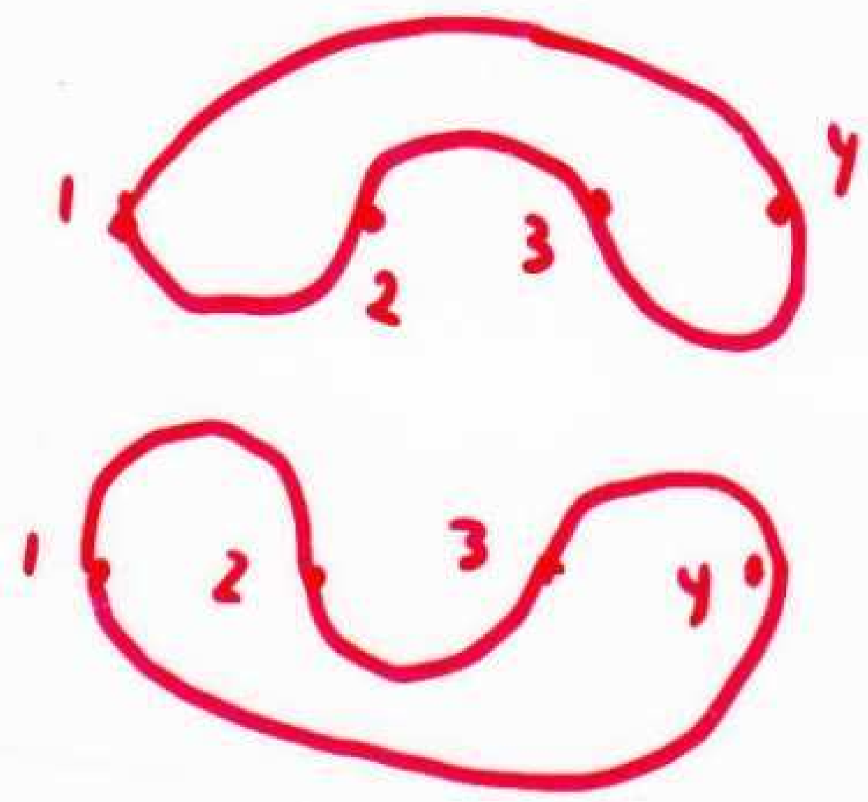
The only tricky thing is that $C_3 = 1$, so the Catalan numbers are shifted two positions:

$$C_N = \sum_k C_k \cdot C_{N-k+1} \qquad C_3 = 1$$

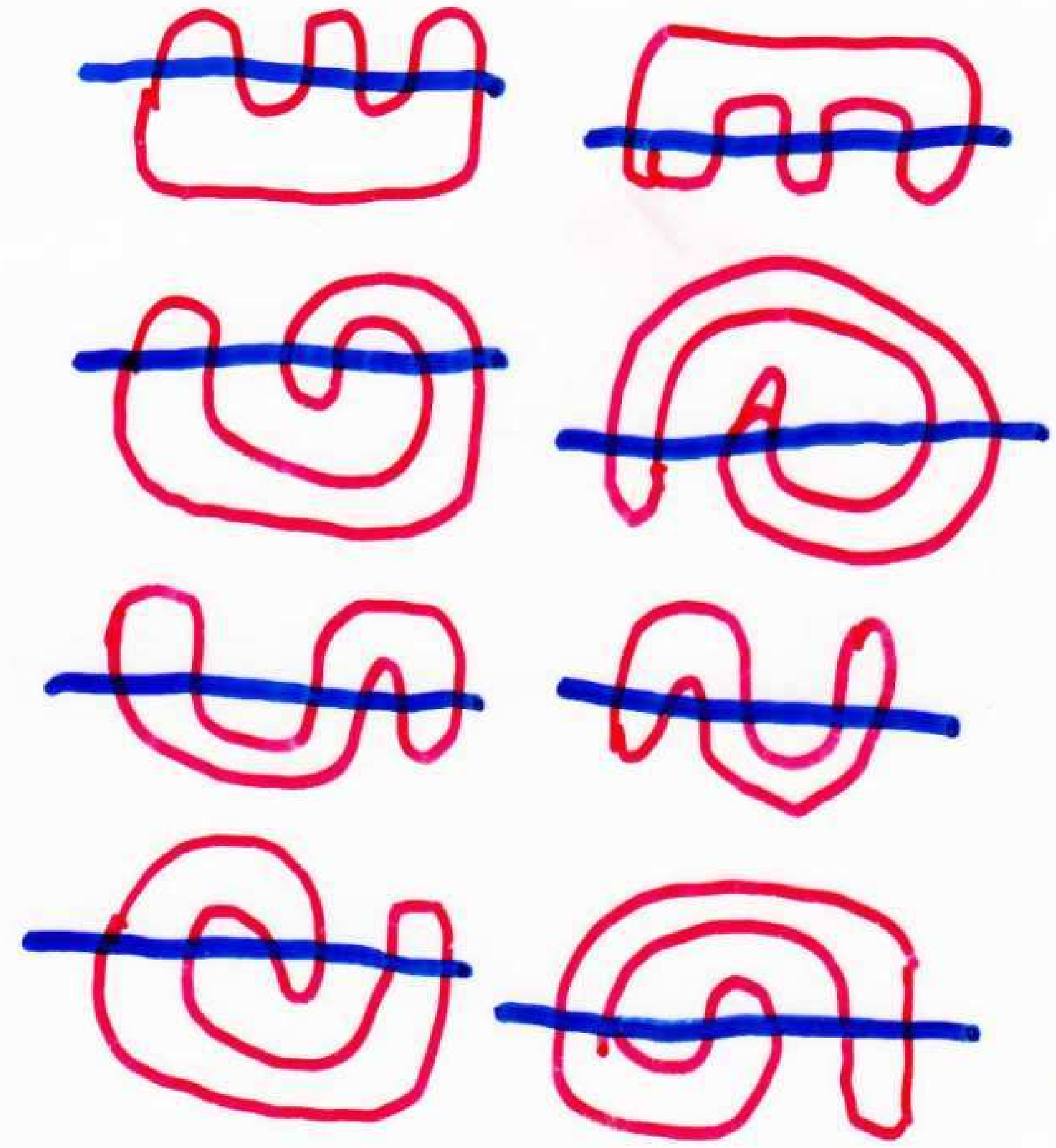
How many Jordan Curves Pass through 2n Points?



$J(1) = 1$



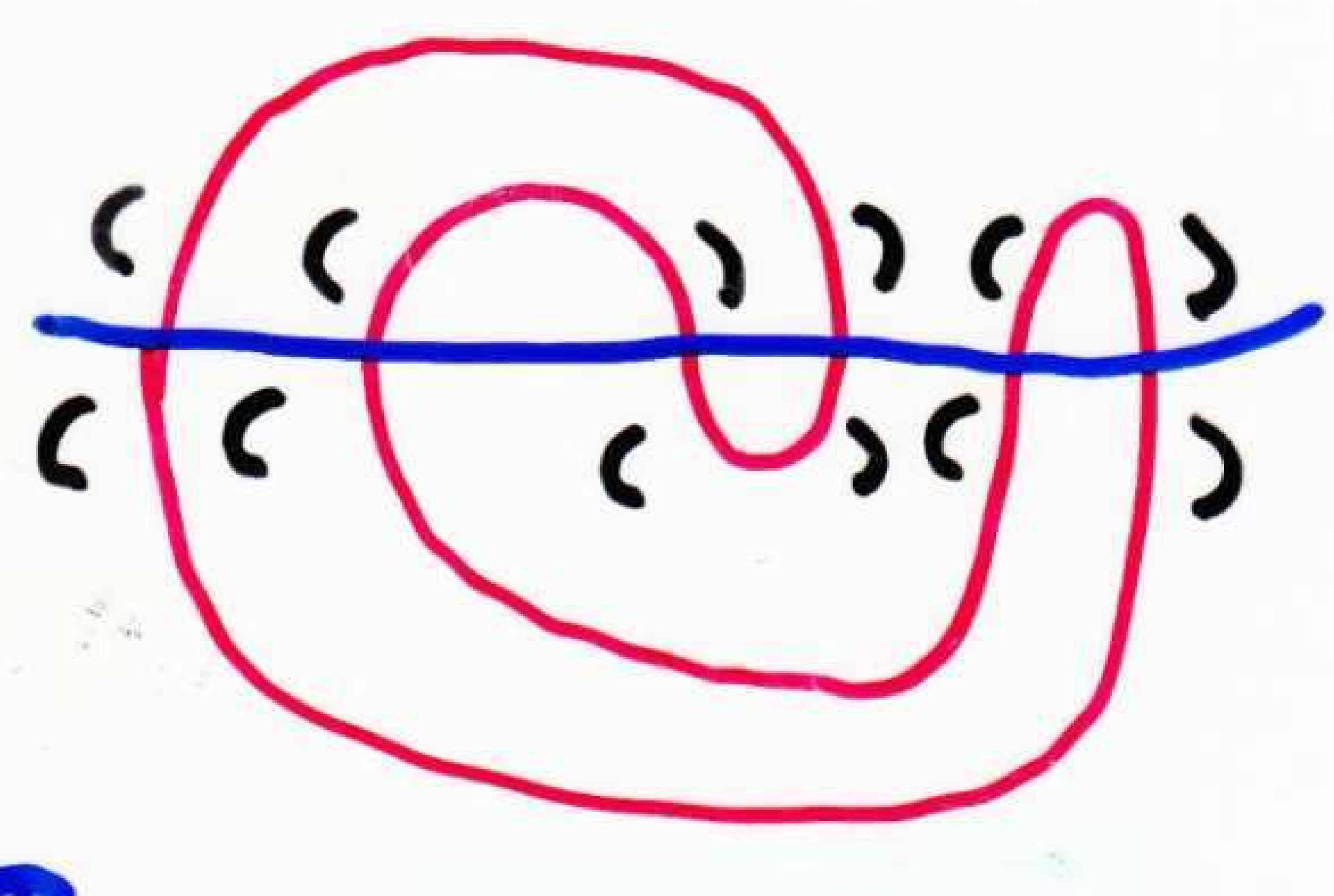
$J(2) = 2$

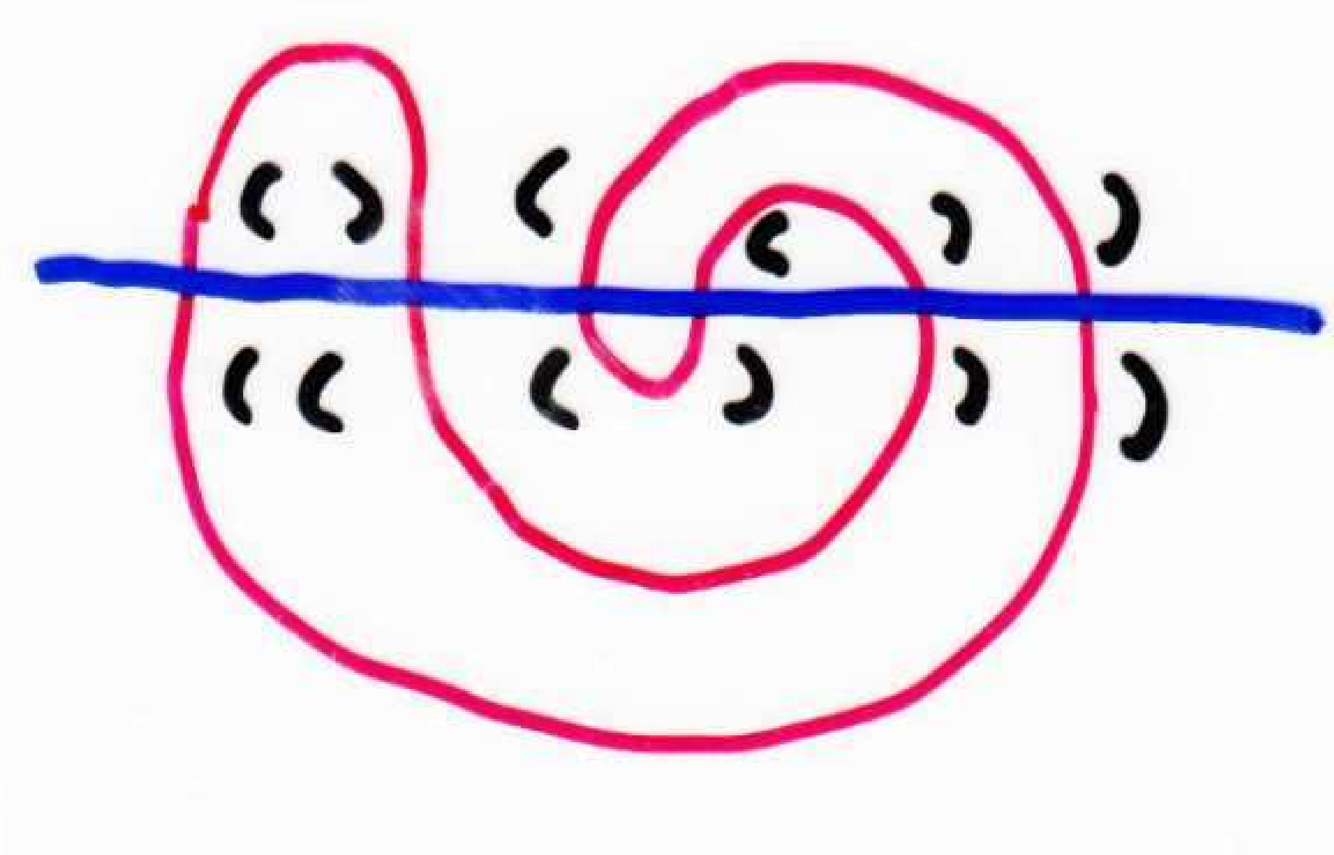


$J(3) = 8$

A Jordan curve is a closed, non-intersecting curve.

Both the region above and below the line characterizes a parenthesization:





Thus $C(n) \le J(n) \le C(n)^2$

Bounded by the Catalan numbers!

Combinatorial Proof that $C_N = \binom{2N}{N}\frac{1}{N+1}$

Consider a sequence of $N+1$ 1's & N -1's $(1,-1,1,-1,1,1,-1)$ which totals up to 1.

Any ballot sequence with an extra one is such a sequence

Any such sequence has one circular shift which has the property that every prefix is positive!

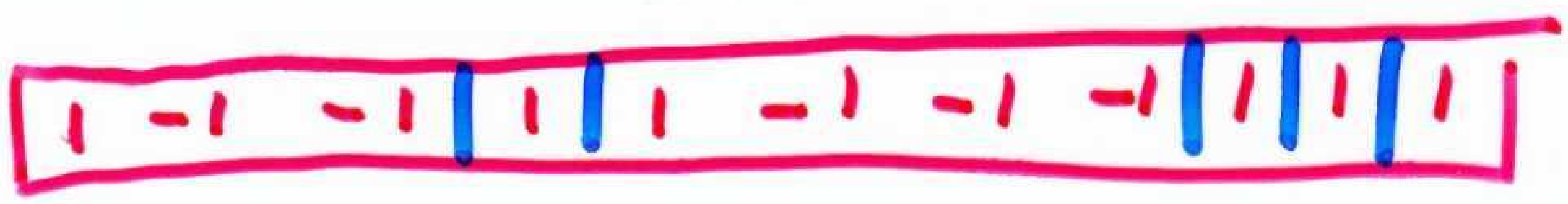
There are $\binom{2N+1}{N}$ ways to pick $N+1$ 1s and N -1's. Exactly $\frac{1}{2N+1}$ of them is the appropriate positive prefix. So

$$C_N = \binom{2N+1}{N}\frac{1}{2N+1} = \binom{2N}{N}\frac{1}{N+1}$$

$(1,-1,1,-1,1,1,-1) \rightarrow (1,1,-1,1,-1,1,-1)$

Why must every such sequence have exactly 1 all-prefix positive shift?

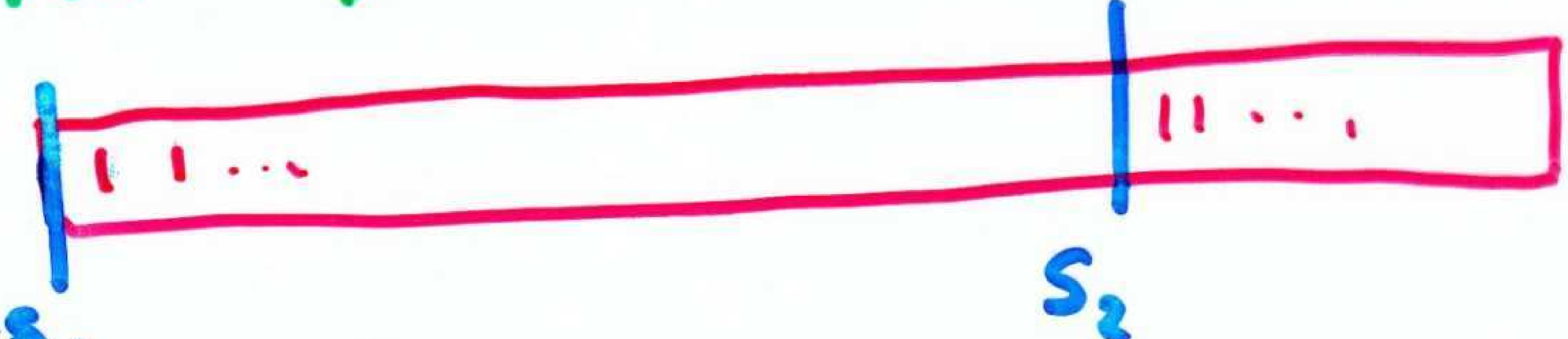
Any sequence of $N+1$ 1s + N -1s must have at least one such shift



Why? Partition into blocks by +1s, and move any negative prefix to the rear. Each moved prefix is now positive, since it is behind enough positive 1s.

Any sequence of $N+1$ 1s + N -1s can only have one such shift.

Proof by contradiction. If $\exists$ 2 starting points:



Total sum $1 \ldots N = 1$

Sum from $S_2 \ldots N \geq 1$

Sum for $S_1 \cdots S_2 - 1 \leq 0$ ← not all prefix positive