5.2. **Dyck paths.** First we discussed random walks. This is a "stochastic process" which you learn about in Math 56a. Then we converted this into a Dyck paths. Then I showed the bijection from binary trees to Dyck paths. I asked the class to come up with the inverse construction (Dyck paths $\Rightarrow$ binary tree).

5.2.1. *random walk.* By a (fair) random walk we mean a sequence of integers $X_0, X_1, X_2, \cdots$ so that, given X_t, $X_{t+1} = X_t + 1$ or $X_t - 1$ with equal probability. If you flip a coin and get 1$ for each head and lose 1$ for each tail then X_t represents the amount of money you have after t tosses of the coin.

Question 5.12. *Suppose $X_0 = 0$. Then what is the probability that $X_{2n} = 0$? In terms of equations this is:*

$$\mathbb{P}(X_{2n} = 0 \,|\, X_0 = 0) = ?$$

The formula for probability, given that each outcome has equal probability, is

$$\mathbb{P}(A) = \frac{\#\text{ways that } A \text{ can occur}}{\text{total } \# \text{ of possibilities}}$$

(If the coin were not fair, the equation is more complicated.)

Since we take $2n$ steps going either left or right with equal probability, there are 2^{2n} total number of possibilities. This goes in the denominator. To end up where we started we need the same number of left steps and right steps. This number is $2n$ choose n. So, students figured out that the answer is:

$$\mathbb{P}(X_{2n} = 0 \,|\, X_0 = 0) = \frac{\binom{2n}{n}}{2^{2n}}$$

Question 5.13. *Suppose that $X_0 = 0$. Then what is the probability that, not only is $X_{2n} = 0$ but $X_k \geq 0$ for $k = 0, 1, 2, \cdots, 2n$?*

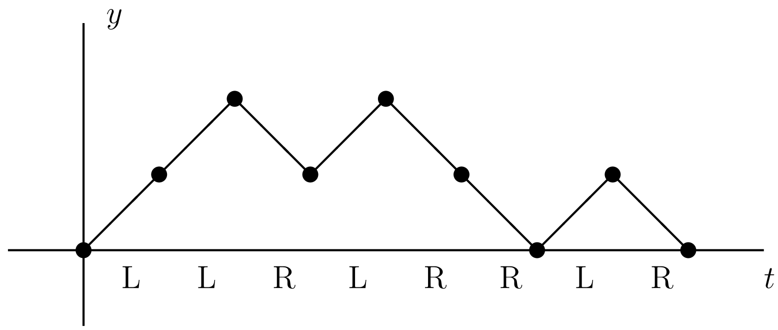
This is something that we did in Math 23b. The answer is:

$$\mathbb{P}(X_{2n} = 0, X_k \geq 0, 0 \leq k \leq 2n \,|\, X_0 = 0) = \frac{\frac{1}{n+1}\binom{2n}{n}}{2^{2n}} = \frac{C(n)}{2^{2n}}$$

The difference between these two equations is the factor of $1/(n+1)$. This factor represents the *conditional probability*:

$$\mathbb{P}(X_k \geq 0, 0 \leq k \leq 2n \,|\, X_0 = 0, X_{2n} = 0) = \frac{1}{n+1}$$

I drew the following picture.



Although we didn't prove it yet (It's torture!) this formula says that the number of paths X_t which go up and down the same number of times and always stay above the y axis. Imagining that you are walking left and sometimes stepping left and sometimes right, every step in the positive y direction is a left step and every step in the negative direction is a right step. So the random walk is represented by a sequence of L's and R's.

5.2.2. *definition of Dyck path.*

Definition 5.14. A *Dyck path* is a word with $2n$ letters: n L's and n R's with the property that, if the word is broken into two parts, the first part (the one on the left) has at least as many L's as R's.

As I explained in class, this definition has the advantage of being completely rigorous and not relying on a picture.

Theorem 5.15. *The number of Dyck paths is at most* $\binom{2n}{n}$.

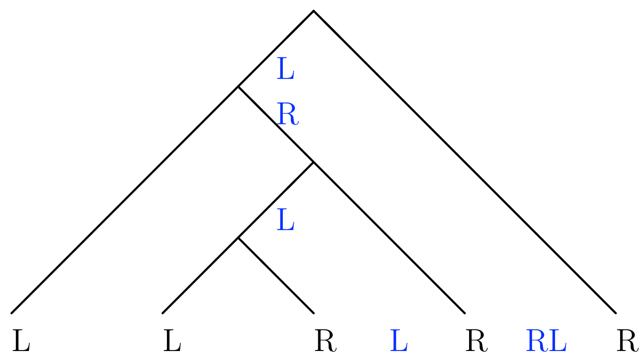
Proof. This binomial is the total number of words made out of n L's and n R's. □

5.2.3. *bijection with set of binary trees.*

Theorem 5.16. *There is a bijection between the set of rooted binary trees with n nodes and the set of Dyck paths with $2n$ letters.*

In order to prove this we need to show how every binary tree gives a Dyck path and, conversely, how every Dyck path gives a binary tree. This will give a bijection. This bijection should also preserve the "rank." I described the first mapping and challenged the class to find an algorithm for the inverse, i.e., which binary tree gives the given Dyck path?

In words, the mapping from binary trees to Dyck paths goes as follows. As Liz pointed out, for each binary tree with $2n$ nodes, there are n edges going to the left and n edges going to the right. However, these edges are not in a row. Some of them are hanging in the air. We bring these letters down to the level of the leaves by sliding them to the right.



But, what is the inverse construction? What is the formula for the binary tree.

5.2.4. *rank.* I also pointed out that the bijection should preserve "rank." The equation:

$$Catalan\# = \text{sum of } Narayana\#'s:$$

$$C(n) = \sum_{k=1}^{n} N(n,k)$$

is usually interpreted in the following way. $C(n)$ is the number of objects that we are considering. They might be Dyck paths of length $2n$ or binary trees with n nodes. Each object has a "rank" k between 1 and n. The number of items in our set with rank k is $N(n,k)$. This explains why the Narayana numbers add up to the Catalan number.

For Dyck paths the rank is the number of peaks in the graph of the path. So, $k = 3$ in our example. For binary trees it is the number of rows that the nodes come in.

(1) Find the unique binary tree with rank 1.
(2) How can you tell from a Dyck path where the peaks are?