5. Generating functions. Binary trees.

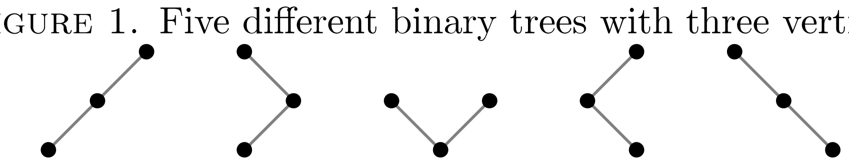
To read: [3] Chapter 12.4.

5.1. **Binary trees.**

Definition 5.1. An inductive definition of a *binary tree* can be given as follows: a binary tree either is empty (it has no vertex), or consists of one distinguished vertex called the root, plus an ordered pair of binary trees called the left subtree and right subtree.

Let b_n denote the number of binary trees with n vertices. Our goal is to find a formula for b_n.
Example. By definition we have $b_0 = 1$ and there is one empty tree. We have $b_1 = 1$, $b_2 = 2$, $b_3 = 5$.

Figure 1. Five different binary trees with three vertices.



The inductive definition of a binary tree implies the following recursive formula for b_n:

$$b_n = b_0\, b_{n-1} + b_1\, b_{n-2} + b_2\, b_{n-3} + \ldots + b_{n-1}\, b_0, \qquad n \in \mathbb{Z}_{\geq 1}. \tag{3}$$

Let $b(x) = \sum_{n=0}^{\infty} b_n x^n$ be the generating series of the sequence $\{b_n\}_{n=0}^{\infty}$. We find

$$b(x)^2 = b_0^2 + (b_1 b_0 + b_0 b_1)x + (b_2 b_0 + b_1 b_1 + b_0 b_2)x^2 + \ldots$$

The recursive relation 3 implies

$$b(x)^2 = b_1 + b_2 x + b_3 x^2 + \ldots = \frac{1}{x}(b_0 + b_1 x + b_2 x^2 + \ldots) - \frac{b_0}{x} = \frac{1}{x} b(x) - \frac{1}{x}.$$

Therefore, the generating function $b(x)$ satisfies the quadratic equation

$$x b(x)^2 - b(x) + 1.$$

This equation has two solutions

$$\frac{1 + \sqrt{1-4x}}{2x} \qquad \text{and} \qquad \frac{1 - \sqrt{1-4x}}{2x}.$$

We observe that the first solution is not bounded around $x = 0$ and the second solution is smooth around $x = 0$ tends to 1 as x tends to 0. Consider the second solution

$$\widetilde{b}(x) = \frac{1 - \sqrt{1-4x}}{2x}.$$

It has Taylor expension around $x = 0$

$$\widetilde{b}(x) = \sum_{n=0}^{\infty} \widetilde{b}_n x^n.$$

We have computed that $\widetilde{b}_0 = \widetilde{b}(0) = 1$. Moreover, the function $\widetilde{b}(x)$ satisfies the quadratic equation

$$x\, \widetilde{b}(x)^2 - \widetilde{b}(x) + 1$$

and therefore the sequence $\{\widetilde{b}_n\}_{n=0}^{\infty}$ satisfies the recursive relation (3). Since the sequences satisfy the same initial conditions $b_0 = \widetilde{b}_0$ and the same recursive relation (3) we conclude that $b_n = \widetilde{b}_n$ for all $n \in \mathbb{Z}_{\geq 0}$. The generalized binomial theorem implies

$$\sqrt{1-4x} = \sum_{k=0}^{\infty} (-4)^k \binom{1/2}{k} x^k.$$

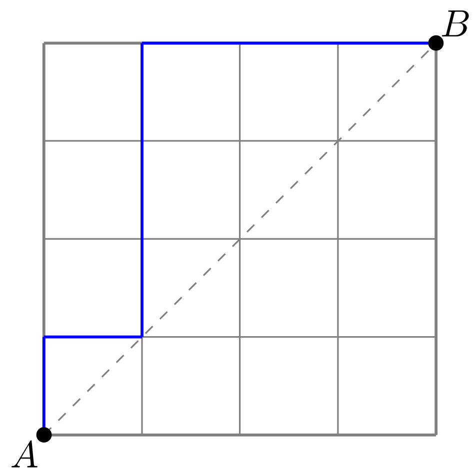
This implies $b_n = \frac{-1}{2}(-4)^{n+1}\binom{1/2}{n+1}$.

Exercise 3. Show that

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

Definition 5.2. The numbers b_n are known by the name *Catalan numbers.*

Exercise 4. Consider an $n \times n$ chessboard:



Consider the shortest paths from the corner A to the corner B following the edges of the squares (each of them consists of $2n$ edges).
(a) How many such paths are there?
(b)* Show that the number of paths that never go below the diagonal (the line AB) is exactly b_n, i.e. the Catalan number. One such path is drawn in the figure.

Acknowledgements: I thank Prof. Janos Pach for designing this course and Dr. Matthew de Courcy-Ireland for sharing his lecture notes.

References

[1] Discrete Mathematics (L. Lovasz, J. Pelikan , K. Vesztergombi);
[2] Combinatorics: Set Systems, Hypergraphs, Families of Vectors and Combinatorial Probability (B. Bollobas);
[3] Invitation to Discrete Mathematics (J. Matousek, J. Nesetril).
[4] Extremal combinatorics (S. Jukna).
[5] Combinatorial theory (M. Hall), Blaisdell publishing company, 1967.