LECTURE 4.

ABSTRACT. Lattice paths and Catalan numbers.

Consider a rectangle $m \times n$ tiled by squares 1×1; boundaries of squares form a lattice. A peg standing in the lower left corner is moving to the upper right corner; only moves "up" and "right" are allowed. The motion of the peg is called a lattice path; its total length is $m+n$ moves.

Theorem 1. *There are $\binom{m+n}{m} = \binom{m+n}{n}$ different lattice paths.*

Proof. The lattice path is uniquely determined by telling which m of the moves are "up" (the remaining n are "right"). □

Let f be a lattice path. Denote by $a_i(f)$, $0 \le i \le m$, the length of the path segment on the i-th horizontal. One has $0 \le a_i(f) \le n$ for each i, and $a_0(f) + \dots + a_m(f) = n$. Apparently, for any numbers $x_0, \dots, x_m$ with these properties there exists exactly one lattice path f such that $a_i(f) = x_i$ for all $i = 1, \dots, m$.

Corollary 1. *The equation $x_0 + \dots + x_m = n$ has $\binom{m+n}{m}$ nonnegative integer solutions.*

The solutions are called compositions of n containing $m+1$ parts.

For a lattice path f consider a sequence $0, \dots, 0, 1, \dots, 1, \dots, m, \dots, m$ where each value i is repeated $a_i(f)$ times (if $a_i(f) = 0$ then i is skipped). The sequence obtained $b_1(f), \dots, b_n(f)$ contains n terms, is increasing (not strictly) and has the property $0 \le b_i(f) \le m$. Increasing sequences may be interpreted as multisets: a multiset is a finite sequence $y_1, \dots, y_n$ considered up to permutation of its terms. Thus $\{\{1,1,2\}\}$ and $\{\{1,2,1\}\}$ are the same multiset but $\{\{1,1,2\}\}$ and $\{\{1,2,2\}\}$ are not the same. Elements of a multiset can be uniquely arranged to form an increasing sequence.

Corollary 2. *The total number of multisets of n elements x such that $0 \le x \le m$ is equal to $\binom{m+n}{m}$.*

Take now $m = n$. The lattice path in the square $n \times n$ is called a Dyck path if it lies above (or on) the diagonal joining lower left corner with the upper right corner. The number of Dyck paths is called the n-th Catalan number and denoted (usually) by c_n. It is convenient to take $c_0 = 1$ by definition.

Theorem 2. *The Catalan numbers satisfy the identity $c_n = \sum_{k=1}^{n} c_{k-1} c_{n-k}$ for any $n = 1, 2, \dots$.*

Proof. A Dyck path f starts on the diagonal at the point $(0,0)$; let k be the smallest positive integer such that the point (k,k) lies on the path f (one has $1 \le k \le n$). Count the number of Dyck paths f with a fixed k.

The segment of the path f between the points $(0,0)$ and (k,k) starts with a vertical move and finishes by a horizontal one. Between these moves there lies a Dyck path of $2(k-1)$ moves, shifted one position up. The number of such paths is c_{k-1} ($c_0 = 1$ by definition). The segment between the points (k,k) and (n,n) is a Dyck path of $2(n-k)$ moves; there are c_{n-k} of them. Thus a total number of Dyck path with a fixed k is $c_{k-1}c_{n-k}$. □

Theorem 3. *One has $C \stackrel{def}{=} \sum_{n=0}^{\infty} c_n t^n = \frac{1-\sqrt{1-4t}}{2t}$.*

Proof. Multiply the identity from Theorem 2 by t^n and sum it over all $n = 1, 2, \dots$. Since $c_0 = 1$, one has

$$C - 1 = \sum_{n=1}^{\infty} c_n t^n = \sum_{n=1}^{\infty} \sum_{k=1}^{n} c_{k-1} c_{n-k} t^n = t \sum_{n=1}^{\infty} \sum_{k=1}^{n} c_{k-1} t^{k-1} \cdot c_{n-k} t^{n-k}$$

changing variables $p \stackrel{\text{def}}{=} k-1$ and $q \stackrel{\text{def}}{=} n-k$

$$= t\Big(\sum_{p=0}^{\infty} c_p t^p\Big) \cdot \Big(\sum_{q=0}^{\infty} c_q t^q\Big) = tC^2.$$

The quadratic equation obtained has two solutions: $C = (1+\sqrt{1-4t})/(2t)$ and $C = (1-\sqrt{1-4t})/(2t)$, but only the second one is equal to 1 at $t = 0$. □

Corollary 3. $c_n = \frac{1}{n+1}\binom{2n}{n} = \frac{(2n)!}{n!(n+1)!}$.

Proof. By Newton's binomial formula for the exponent $1/2$ one has

$$\begin{aligned}\sqrt{1-4t} &= 1+\sum_{n=1}^{\infty}(-1)^n\binom{1/2}{n}4^nt^n = 1+\sum_{n=1}^{\infty}(-1)^n\frac{1}{n!}\frac{1}{2}\cdot\left(-\frac{1}{2}\right)\cdot\left(-\frac{3}{2}\right)\cdot\dots\cdot\left(-\frac{2n-3}{2}\right)4^nt^n\\ &= 1+\sum_{n=1}^{\infty}(-1)^{2n-1}\frac{1\cdot3\cdot\dots\cdot(2n-3)}{2^nn!}4^nt^n = 1-\sum_{n=1}^{\infty}\frac{(2n-2)!}{2^{2n-1}(n-1)!n!}2^{2n}t^n,\end{aligned}$$

hence, $(1-\sqrt{1-4t})/(2t) = \sum_{n=1}^{\infty}\frac{(2n-2)!}{n!(n-1)!}t^{n-1}$. Changing the summation variable $n \mapsto n+1$ one obtains the required formula. $\square$

Exercises

Exercise 1. Prove that c_n is equal to the number of balanced bracket structures containing n pairs of brackets.

For example, for one pair or brackets the only structure is `()`; for two pairs there are two: `()()` and `(())`, etc.

Exercise 2. Prove that c_n is equal to the number of triangulations of a convex $(n+2)$-gon by diagonals. Triangulation is a splitting of the polygon into triangles; some of the triangles may have common sides and/or vertices, but have otherwise no intersections. The vertices of the polygon are numbered $1, 2, \dots, (n+2)$ counterclockwise.

For example, a square with the vertices 1, 2, 3, 4 has $c_2 = 2$ triangulations: draw either a diagonal $1, 3$ or a diagonal $2, 4$. For a pentagon there are $c_3 = 5$ triangulations; if the pentagon is regular they are mapped one to the other by rotations.

Exercise 3. Let $a_0, \dots, a_n$ be variables and $*$ be the symbol of a binary operation. Prove that c_n is equal to the number of ways to put brackets correctly in the expression $a_0 * a_1 * \dots * a_n$.

For example, for $n = 2$ there are three terms $a_0 * a_1 * a_2$ and $c_2 = 2$ ways to put brackets: $(a_0 * a_1) * a_2$ and $a_0 * (a_1 * a_2)$. Note that Exercise 3 is not the same as Exercise 1!