MATH 465 - INTRODUCTION TO COMBINATORICS
LECTURE 9

1. Catalan numbers

The *Catalan numbers* C_n are defined by

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

n	0	1	2	3	4	5	6	7	8	9
C_n	1	1	2	5	14	42	132	429	1430	4862

These numbers enumerate many families of combinatorial objects. Note that it is not clear from the definition we have given that these are integers.

1.1. **Circular arrangemets of ± 1s.**

Suppose $\alpha = a_1 a_2 \cdots a_{2n+1}, a_i \in \{1, -1\}$ is a circular arrangement of $n+1$ 1s and n -1s. We often write $-$ for -1. Such a representation of α is of course not unique. Any cyclic shift $a_{k+1} \cdots a_{2n+1} a_1 \cdots a_k$ represents the same circular arrangement. In general, the cyclic shifts of a circular arrangement of 1s and -1s are not distinct. For example, if we have $1-1-$ and cyclically shift it two units, we get $1-1-$ again (writing $-$ for -1). However:

Lemma 1.1. *If we have $n+1$ 1s and n -1s, all cyclic shifts of α are distinct.*

Proof. Suppose α is invariant under cyclic shift by k. Write α k times in a line

$$\alpha\alpha\cdots\alpha = a_1 \cdots a_{2n+1} a_1 \cdots a_{2n+1} \cdots a_1 \cdots a_{2n+1}.$$

There are $(n+1)k$ 1s and nk -1s in this sequence. Let $\beta = a_1 a_2 \cdots a_k$. Then, since α is invariant under cyclic shift by k,

$$\alpha\alpha\cdots\alpha \ (k \text{ times}) = \beta\beta\cdots\beta \ (2n+1 \text{ times}).$$

Now we count the number of 1s in two different ways:

(1) α contains $n+1$ 1s, so $\alpha\alpha\cdots\alpha$ (k times) contains $k(n+1)$ 1s.
(2) Suppose β contains l 1s. Then, $\beta\beta\cdots\beta$ ($2n+1$ times) contains $(2n+1)l$ 1s.

Therefore, we must have $k(n+1) = (2n+1)l$. Since $n+1$ and $2n+1$ are coprime, k is divisible by $2n+1$, so that cyclic shift by k is a number of full rotations of the circle. □

Theorem 1.2. *The number of circular arrangements of $n+1$ 1s and n -1s, up to rotation, is the Catalan number C_n.*

Proof. There are $\binom{2n+1}{n}$ ways to place $n+1$ 1s and n -1s in a line, and for each circular arrangement, there are $2n+1$ linear arrangements related by cyclic shifts. By the division principle, we get $\frac{1}{2n+1}\binom{2n+1}{n}$. □

Corollary 1.3. *C_n is an integer.*

1.2. **Ballot sequences.** A *ballot sequence* of length $2n$ is a sequence of n 1s and n -1s in which each initial segment contains at least as many 1s as -1s. More precisely, a ballot sequence is a sequence $a_1, \ldots, a_{2n}$ such that:

(1) $a_i \in \{\pm 1\}$;
(2) $a_1 + \cdots + a_{2n} = 0$;
(3) Each partial sum $a_1 + \cdots + a_k, 1 \leq k \leq 2n$, is nonnegative.

Example 1.4 ($n = 3$.)**.** There are five ballot sequences:

$$111 - - - \quad 11 - 1 - - \quad 11 - -1 - \quad 1 - 1 - 1 - \quad 1 - -11 - .$$

Theorem 1.5. *The number of ballot sequences of length $2n$ is C_n.*

The proof will rely on the following lemma.

Lemma 1.6. *In any circular arrangement of ± 1s whose sum is equal to 1, there is exactly one location starting from which all clockwise partial sums are positive.*

Proof. Suppose $\alpha = a_1 a_2 \cdots a_{2n+1}$ is a circular arrangement of ± 1s such that $a_1 + a_2 + \cdots + a_{2n+1} = 1$.

(1) Uniqueness: Suppose there are two such locations $1 \leq k < l \leq 2n+1$. The circle breaks into two segments $a_k a_{k+1} \cdots a_{l-1}$ and $a_l a_{l+1} \cdots a_{2n+1} a_1 \cdots a_{k-1}$ such that the sums of each segment are ≥ 1. The total sum is therefore ≥ 2, a contradiction.
(2) Existence: Suppose there is no location starting from which all clockwise partial sums are positive. Start someplace and find a nonpositive clockwise partial sum. Start at the end and find a nonpositive clockwise partial sum and so on. Eventually we will revisit a place, and the sum of all entries we have seen is ≤ 0, a contradiction.

□

Proof of Theorem 1.5. By Lemma 1.6, there exists a unique k such that $a_k a_{k+1} \cdots a_{2n+1} a_1 \cdots a_{k-1}$ has positive clockwise partial sums. Therefore $a_k = 1$ and removing it, we get a ballot sequence $a_{k+1} \cdots a_{2n+1} a_1 \cdots a_{k-1}$. □

1.3. **Dyck paths.** A *Dyck path* of length $2n$ is a lattice path $v_0 v_1 \cdots v_{2n}$ in the coordinate plane which

(1) connects the points $(0,0)$ and $(2n,0)$, i.e., $v_0 = (0,0)$ and $v_{2n} = (2n, 0)$,
(2) consists of steps $v_{i+1} - v_i \in \{(1,1), (1,-1)\}$, and
(3) is entirely contained in the upper half-plane $y \geq 0$.

Theorem 1.7. *The number of Dyck paths of length $2n$ is $C_n = \frac{1}{n+1}\binom{2n}{n}$.*

Proof. Dyck paths of length $2n$ are in bijection with ballot sequences of length $2n$:

$$v_0 v_1 \cdots v_{2n} \mapsto a_1 a_2 \cdots a_{2n},$$
$$a_i = \begin{cases} 1 & \text{if } v_i - v_{i-1} = (1,1), \\ -1 & \text{if } v_i - v_{i-1} = (1,-1). \end{cases}$$

In other words, we are recording the y-coordinates of the steps. Therefore, the y-coordinate of the point v_k is $a_1 + \cdots + a_i$. The condition that v_k is contained in $y \geq 0$ becomes the ballot condition $a_1 + \cdots + a_k \geq 0$.

□

1.4. **Recurrence for Catalan numbers.**

Lemma 1.8. *Every ballot sequence splits uniquely into a concatenation of the form*

$$1(\textit{ballot sequence}) - (\textit{ballot sequence}).$$

Proof. Suppose $a_1, \dots, a_{2n}$ is a ballot sequence. Take the first partial sum that is 0, i.e., let $2 \le k \le 2n$ be the smallest integer such that $a_1 + \cdots + a_k = 0$. Such a k must exist because $a_1 + \cdots + a_{2n} = 0$. Then $a_2 \dots a_{k-1}$ and $a_{k+1} \dots a_{2n}$ are ballot sequences, and we have

$$a_1, \dots, a_{2n} = 1, (a_2, a_3, \dots, a_{k-1}), -1, (a_{k+1}, \dots, a_{2n}).$$

Now we show uniqueness. Suppose l is such that $1, (a_2, \dots, a_{l-1}), -1, (a_{l+1}, \dots, a_{2n})$ is splitting. Since $a_2, \dots, a_{l-1}$ is a ballot sequence, $a_1 + \cdots + a_l = 0$, so $l \ge k$. If $l > k$, then since $a_1 + \cdots + a_k = 0$, we have $a_2 + \cdots + a_k = -1$, contradicting the assumption that $a_2, \dots, a_{l-1}$ is a ballot sequence. □

Theorem 1.9. *The Catalan numbers satisfy the recurrence*

$$C_n = \sum_{k+l=n-1} C_k \, C_l \, .$$

Proof. Let $\mathrm{BS}(n)$ denote the set of ballot sequences of length $2n$. By Lemma 1.8, we have a well-defined function $f : \mathrm{BS}(n) \to \bigsqcup_{k+l=n-1} \mathrm{BS}(k) \times \mathrm{BS}(l)$. f is a bijection: f^{-1} is the function that sends the pair (α, β), $\alpha \in \mathrm{BS}(k)$ and $\beta \in \mathrm{BS}(l)$, to $1\alpha - \beta \in \mathrm{BS}(n)$. □

Theorem 1.10 (Exercise)**.** *The generating function for Catalan numbers is*

$$\sum_{n=0}^{\infty} C_n \, x^n = \frac{1 - \sqrt{1-4x}}{2x}.$$