Combinatorics I

Kon Yi

December 6, 2025

Abstract

The lecture note of Combinatorics I by Shagnik Das, where the NTU cool site is https://cool.ntu.edu.tw/courses/55532/.

Contents

1 Chatting **2**
1.1 Prime Numbers . . . 2

2 Elementary Counting Principles **4**
2.1 Sum Rule . . . 4
2.2 Product Rule . . . 6
2.3 Double-Counting argument . . . 7
2.4 Permutations . . . 7
2.5 Binomial Theorem . . . 10
2.6 Divisor Function . . . 13

3 Partitions **15**
3.1 Number of nonnegative integer solution to $x_1 + \cdots + x_k = n$ . . . 15
3.2 Stirling numbers of the first kind . . . 18
3.3 The twelvefold way of Counting . . . 26

4 Generating Functions **28**
4.1 Dictionary for operations . . . 31
4.2 Recurrence relation . . . 32
4.3 Generating function operation . . . 34
4.4 Products of Generating Functions . . . 35
4.5 Catalan Numbers . . . 36
4.6 Exponential Generating Functions . . . 43

5 Asymptotics **51**
5.1 Factorial . . . 53
5.2 Binomial coefficients . . . 55
5.3 Partition Function . . . 57

6 Partially ordered sets **61**
6.1 Inclusion-exclusion principle . . . 61
6.2 Partially ordered sets (Posets) . . . 65
6.3 Mobius Inversion . . . 72

Chapter 1

Chatting

Lecture 1

1.1 Prime Numbers

2 Sep. 15:30

Theorem 1.1.1 (Euclid $\approx$ 300 BCE)**.** There are infinitely many primes.

proof. (Saidak, 2006).

- Every natural number has at least one prime factor.
- No prime divides n and $n+1$, for any $n \in \mathbb{N}$.

Consider a sequence of pronic number

$$p_1 = 2,\ p_{n+1} = p_n(p_n + 1).$$

Then the number of prime factors of p_n is strictly increasing in n: p_{n+1} has all the factors of p_n together with the (disstinct) ones of $p_n + 1$.

Example 1.1.1. $p_1 = 2, p_2 = 6, p_3 = 42, p_4 = 1806$, where the prime factors of them are $\{2\}$, $\{2,3\}, \{2,3,7\}, \{2,3,7,43\}$.

■

1.1.1 How many prime numbers are there?

Definition 1.1.1. We define

$$\pi(n) = |\{p : 1 \leq p \leq n : p \text{ is prime}\}|.$$

Note 1.1.1. By Saidak's proof, we know $\pi(p_n) \geq n$. In fact, $\pi(p_n) \geq \log_2 n$.

Theorem 1.1.2 (Legendre, $\approx$ 1800 LE)**.**

$$\pi(n) \approx \frac{n}{\ln n} \Leftrightarrow \lim_{n\to\infty} \frac{\pi(n)\ln n}{n} = 1$$

Note 1.1.2. Proven by Hadamard and independently de la Vallée Poussin(1896).

Theorem 1.1.3 (Better Approximation).
Dirichlet: $\pi(x) \approx Li(x) = \int_2^x \frac{1}{\ln t} dt$.
Known: $\pi(n) = Li(n) + O\left(ne^{-a\sqrt{\ln n}}\right)$
Believed: $\pi(n) = Li(n) + O\left(\sqrt{n}\ln n\right)$

Chapter 2

Elementary Counting Principles

Fundemental problem: Given a set S, and we want to determine $|S|$.

2.1 Sum Rule

Theorem 2.1.1 (Sum Rule). If $S = \dot\bigcup_{i=1}^{k} S_i$, then $|S| = \sum_{i=1}^{k} |S_i|$.

Note 2.1.1. $\dot\bigcup$ means disjoint union.

Example 2.1.1. A drawer contains 8 pairs of yellow socks, 5 pairs of blue socks, and 3 pairs of red socks. How many socks are there in total.

Informal proof. $2 \times (8 + 5 + 3) = 32$. ■

Proof. Let S be the set of socks in the drawer, then $S = \dot\bigcup_{p \in P} S_p$, where P is the set of pairs of socks, and S_p is the set of two socks in the pair where $p \in P$. By the sum rule,

$$|S| = \sum_{p \in P} |S_p| = \sum_{p \in P} 2 = 2|P| = 32.$$

$P = P_{\text{yellow}} \dot\cup P_{\text{blue}} \dot\cup P_{\text{red}}$. By the sum rule,

$$|P| = |P_{\text{yellow}}| + |P_{\text{blue}}| + |P_{\text{red}}| = 8 + 5 + 3 = 16.$$

■

Note 2.1.2. Sum rule is the basis for case analysis arguments. It needs two requirements:

- Cover each case.
- Cover each case exactly once.

Example 2.1.2. Counting subset of a general set.

Notation. If X is a set, and $k \in \mathbb{N} \cup \{0\}$, then

$$\binom{X}{k} = \{T : \; T \subseteq X, \; |T| = k\}.$$

We define the binomial coefficient as

$$\binom{|X|}{k} = \left|\binom{X}{k}\right|.$$

i.e. Given $n \geq k \geq 0$, $\binom{n}{k}$ is the number of k-element subsets of a set of size n. ■

Proposition 2.1.1 (Pascal's relation)**.** If $n \geq k \geq 1$, then

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

Proof. Proof. Let X be an n-element set (e.g. $X = [n] = \{1, 2, \ldots, n\}$), and let $S = \binom{X}{k} = \{T \subseteq X : |T| = k\}$. Then, by definition, $\binom{n}{k} = |S|$. For each k-element subset, we can ask: "Do you contain n?" Let

$$S_0 = \{T : T \subseteq X, n \notin T, |T| = k\},$$

and

$$S_1 = \{T : T \subseteq X, n \in T, |T| = k\}.$$

Then, $S = S_0 \uplus S_1$. By the sum rule, $|S| = |S_0| + |S_1|$. Observe that

$$\begin{aligned} S_0 &= \{T \subseteq [n], n \notin T, |T| = k\} \\ &= \{T \subseteq [n-1], |T| = k\}, \end{aligned}$$

so by definition,

$$|S_0| = \binom{|[n-1]|}{k} = \binom{n-1}{k}.$$

$$S_1 = \{T \subseteq [n], n \in T, |T| = k\}.$$

Let

$$S_1' = \{T' \subseteq [n-1], |T'| = k-1\},$$

then we know a bijection from S_1 to S_1':

$$T \in S_1 \longleftrightarrow T \setminus \{n\} \in S_1'.$$

Theorem 2.1.2 (bijection rule)**.** Given two sets S and S', if there is a bijection $f : S \to S'$, then $|S| = |S'|$.

By this rule, we know

$$|S_1| = |S_1'| = \binom{|[n-1]|}{k-1} = \binom{n-1}{k-1}.$$

Hence,

$$\binom{n}{k} = |S| = |S_0| + |S_1| = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

■

2.1.1 Pascal's Triangle

We can use Pascal's relation to compute $\binom{n}{k}$.

Note 2.1.3. Boundary case: $\binom{n}{0} = 1$, $\binom{n}{n} = 1$. Also, $\binom{n}{k} = 0$ for $k = -1, n+1$.

$$
\begin{array}{ccccccc}
 & & & \binom{0}{0} & & & \\
 & & \binom{1}{0} & & \binom{1}{1} & & \\
 & & & \searrow \ \swarrow & & & \\
 & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & \\
 & & \searrow \ \swarrow & & \searrow \ \swarrow & & \\
\binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3}
\end{array}
$$

2.2 Product Rule

Theorem 2.2.1. If $S = S_1 \times S_2 \times \cdots \times S_k = \{(x_1, x_2.., x_k), x_i \in S_i\}$, then $|S| = \prod_{i=1}^{k} |S_i|$.

Proof. Induction on k:
Base case: $k = 1$, trivial.
Induction step: seperate into cases bases on choice of $x_{k+1} \in S_{k+1}$. Let

$$S(x) = \{(x_1, \dots, x_{k+1}) \in S, x_{k+1} = x \in S_{k+1}\},$$

then

$$S = \dot{\bigcup_{x \in S_{k+1}}} S(x) \to |S| = \sum_{x \in S_{k+1}} |S(x)|.$$

But $S(x) = S_1 \times S_2 \times \cdots \times \{x\}$, which is in bijection with $S_1 \times S_2 \times \cdots \times S_k$. By induction rule,

$$|S(x)| = |S_1 \times S_2 \times \cdots \times S_k| \quad \forall x$$

Hence,

$$
\begin{aligned}
|S| = \sum_{x \in S_{k+1}} |S(x)| &= \sum_{x \in S_{k+1}} |S_1 \times S_2 \times \dots S_k| \\
&= |S_1 \times S_2 \times \cdots \times S_k| \times |S_{k+1}| = |S_1| \times |S_2| \times \cdots \times |S_{k+1}|.
\end{aligned}
$$

■

Example 2.2.1. Consider binary strings of length n.

Proof.

$$S = \{0,1\}^n \Rightarrow |S| = |\{0,1\}^n| = |\{0,1\}|^n = 2^n.$$

■

Definition 2.2.1 (Power Set). Given a finite set X, let 2^X denote the set of all subsets of X (also denoted $\mathcal{P}(x)$), which is called the power set.

Corollary 2.2.1. $\left|2^X\right| = 2^{|X|}$.

Proof. Without lose of generality, $X = [n]$. We build a bijection between $2^{[n]}$ and the set of binary string of length n. Suppose for every $T \in 2^{[n]}$, we have $\chi_T = (x_1, x_2, \dots, x_n)$, where

$$x_i = \begin{cases} 1, & \text{if } i \in T; \\ 0, & \text{if } i \notin T. \end{cases}$$

Then,

$$\left|2^{[n]}\right| = |\{0,1\}^n| = 2^n.$$

■

2.3 Double-Counting argument

If we count a set in two different ways, the answer should be equal.

Example 2.3.1. Count $2^{[n]}$.

Proof.

1. Product rule $\to 2^n$.
2. Use the sum rule, split the subsets by size.

$$2^{[n]} = \binom{[n]}{0} \cup \binom{[n]}{1} \cup \ldots \cup \binom{[n]}{n}$$

Hence, we have the following proposition:

Proposition 2.3.1. For all $n \geq 0$,

$$2^n = \sum_{k=0}^{n} \binom{n}{k}.$$

⊛

2.4 Permutations

Lecture 2

5 Sep. 13:10

As previously seen. Instead of choosing the subsets all at once, we could pick one element at a time, then we can try to use product rule.

Example 2.4.1. Consider

$$\binom{[10]}{3}.$$

Proof. At the choice of the first element, we have 10 choices, the second one has 9 choisce, while the third one has 8 choice, but we didn't consider the order of each picked elements. ⊛

Definition 2.4.1. Given a set X and $k \in \mathbb{N} \cup \{0\}$, a k-permutation of X is

- an ordered choice of k distinct elements from X.
- a k-tuple $(x_1, x_2, \ldots, x_k)$ with $x_i \in X$ and $x_i \neq x_j$ for each $i \neq j$.
- an injection $f : [k] \to X$.

where these 3 statements are equivalent.

Notation. $X^{\underline{k}} = \{k\text{-permutation of } X\} \subseteq X^k$ where $X^k = X \times X \times \cdots \times X$ allows repitition of the elements but $X^{\underline{k}}$ don't allow repitition.

Note 2.4.1. If $|X| = n$, then

$$n^{\underline{k}} = \left|X^{\underline{k}}\right|.$$

Definition 2.4.2.

- a n-permutation is a n-permutation of $[n]$.
- a X-permutation is a $|X|$-permutation of X.

Theorem 2.4.1 (Generalized Product Rule)**.** Suppose we are enumerating S, and can uniquely determine an element $s \in S$ through a series of k questions, if i-th problem always has n_i possible outcomes, independently to the permutation, then

$$|S| = n_1 \times n_2 \times \cdots \times n_k = \prod_{i=1}^{k} n_i$$

Proof. Can make a bijection from S to

$$[n_1] \times [n_2] \times \cdots \times [n_k].$$

Map each element in S to the index of its answer in the series of answer.

Our moral is when counting we don't care about what the options are but only how many options. ∎

Proposition 2.4.1.

$$\begin{aligned} n^{\underline{k}} &= n(n-1)\dots(n-(k-1)) \\ &= \prod_{i=0}^{k-1}(n-i) = \frac{n!}{(n-k)!}. \end{aligned}$$

Proof. Use the generalized product rule.

Question i: What is the i-th element in the k-permutation of $[n]$?

We can choose anything except what we're alreafy chosen, so there are $i-1$ forbidden choices and thus there are $n-(i-1)$ possible choices. ∎

Proposition 2.4.2. For all $0 \leq k \leq n$,

$$\binom{n}{k} = \frac{n^{\underline{k}}}{k^{\underline{k}}} = \frac{\left(\frac{n!}{(n-k)!}\right)}{k!} = \frac{n!}{k!(n-k)!}.$$

Proof. Double-count $[n]^{\underline{k}}$ i.e. k-permutation of $[n]$.

- Direct counting $\left|[n]^{\underline{k}}\right| = n^{\underline{k}}$.
- First choose the k elements to appear in the k-permutation, $\binom{n}{k}$ options, then choose the order in which they appear, $k^{\underline{k}}$ options.

  Then, by the generalized product rule, the number of k-permutation of $[n]$ is $\binom{n}{k} \cdot k^{\underline{k}}$.

Hence,

$$n^{\underline{k}} = \left|[n]^{\underline{k}}\right| = \binom{n}{k} \cdot k^{\underline{k}}.$$

∎

Corollary 2.4.1. We can then use this result to reprove Pascal's Property again.

Proof. ∎

Exercise 2.4.1. 6 players at the tennis club want to have three matches involving all the players? How many ways can we arrange the games.

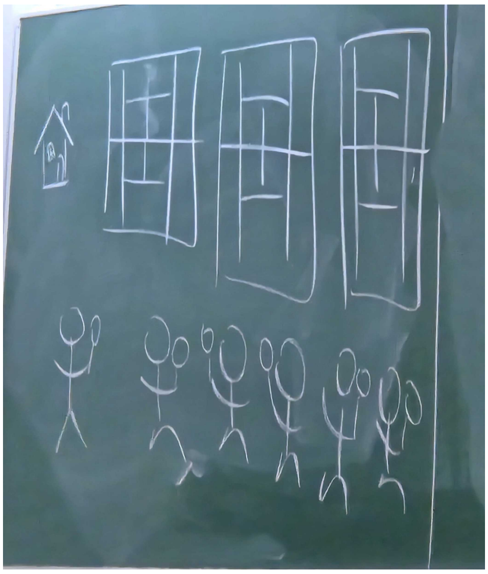


Figure 2.1: Tennis Games

Proof. We only care about who plays against whom, not about which court or who versus first, e.t.c.

The arrangement of games is a set of three disjoint pairs of players.

$$\{\{1,2\},\{3,4\},\{5,6\}\} \neq \{\{1,3\},\{2,4\},\{5,6\}\}.$$

Double-count the arrangements of games where counts do matter.

- Choose a pair of players for Court A: $\binom{6}{2}$
- Choose a pair of players for Court B: $\binom{4}{2}$
- Choose a pair of players for Court C: $\binom{2}{2}$

Generalized product rule tells

$$\text{number of choices} = \binom{6}{2}\binom{4}{2}\binom{2}{2} = 90.$$

Second count: First gets a set of 3 pairs, say there are x possibilities , and assign the three pairs to 3 courts, so there are 3! , so $x \cdot 3! = 90$, and thus $x = \frac{90}{3!} = 15$. ■

Lecture 3

Actually we have an alternitive prove: 9 Sep. 15:30

proof by direct computation.

- Q1: Who's the opponent for the 1-st player? There are 5 choices.
- Q2: Who plays the next lowest numbered player? There are 3 choices.

The left 2 players are the opponents to each other. Hence, there are $3 \times 5 = 15$possible pairings. ■

More generally, if we have $n = 2k$ players to pair up, then the first proof gives there are

$$\frac{\binom{n}{2}\binom{n-2}{2}\cdots\binom{2}{2}}{\left(\frac{n}{2}\right)!}$$

possible pairings, while the second proof gives that there are

$$(n-1)\cdot(n-3)\cdot(n-5)\ldots := (n-1)!! \neq ((n-1)!)!.$$

By this, we know these two numbers must be equal, or more rigorously, we can write

$$\frac{\binom{n}{2}\binom{n-2}{2}\cdots\binom{2}{2}}{\left(\frac{n}{2}\right)!} = 2^n \cdot \frac{\frac{n(n-1)}{2}\,\frac{(n-2)(n-3)}{2}\cdots}{n(n-2)(n-4)\ldots 2} = (n-1)\cdot(n-3)\cdot\ldots.$$

Example 2.4.2. How mant shortest routes on the grid are there from $(0,0)$ to (n,m)?

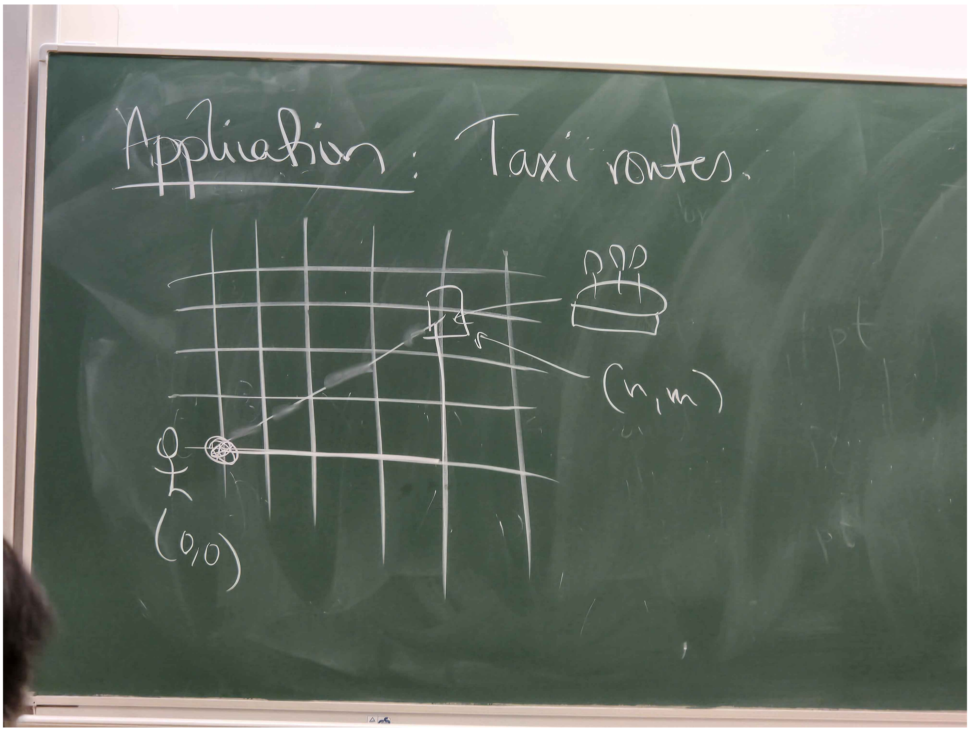


Figure 2.2: Taxi routes

Proof. Shortest route is of length $n+m$,m up-steps and n right-steps. We can think of a shortest route to be a binary string of length $n+m$ with n 1s and m 0s, so we want to count how many such binary strings are there. Choose n of them to be 1s, while the other are 0s. Hence, there are $\binom{n+m}{n}$ possibilities. ⊛

2.5 Binomial Theorem

Theorem 2.5.1 (Binomial Theorem)**.** For any $n \in \mathbb{N} \cup \{0\}$, and $x, y \in \mathbb{R}$, we have

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Example 2.5.1. $(x+y)^0 = 1 = \sum_{k=0}^{0} x^k y^{0-k}$.

Example 2.5.2. $(x+y)^1 = x+y$, while

$$\sum_{k=0}^{1} \binom{1}{k} x^k y^{1-k} = \binom{1}{0} x^0 y^1 + \binom{1}{1} x^1 y^0 = y + x.$$

proof of binomial theorem.

$$(x+y)^n = \underbrace{(x+y)(x+y)(x+y)\dots(x+y)}_{n \text{ factors}}$$

From each factor, we pick a term x or y, multuply chosen factors together. If we choose k x's, then we must choose $n-k$ y's, so the monomial is $x^k y^{n-k}$, where the coefficient of $x^k y^{n-k}$ is the number of ways of choosing k x's. Also, the possible monomials are $x^k y^{n-k}$ for $k = 0, 1, 2, \dots, n$. Hence, we have

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

■

We can use this formula to derive identities for the binomial coefficients, by plugging in values for x and y.

Example 2.5.3. $x = 1, y = 1$.

Proof.

$$2^n = (x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \sum_{k=0}^{n} \binom{n}{k}.$$

⊛

Example 2.5.4. $y = -1, x = 1$.

Proof.

$$(x+y)^n = (-1+1)^n = 0^n = \begin{cases} 1, & \text{if } n = 0; \\ 0, & \text{if } n \geq 1. \end{cases}$$

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (-1)^k = \sum_{2|k} \binom{n}{k} - \sum_{2\nmid k} \binom{n}{k}$$

⊛

Corollary 2.5.1.

$$\sum_{2|k} \binom{n}{k} = \sum_{2\nmid k} \binom{n}{k}$$

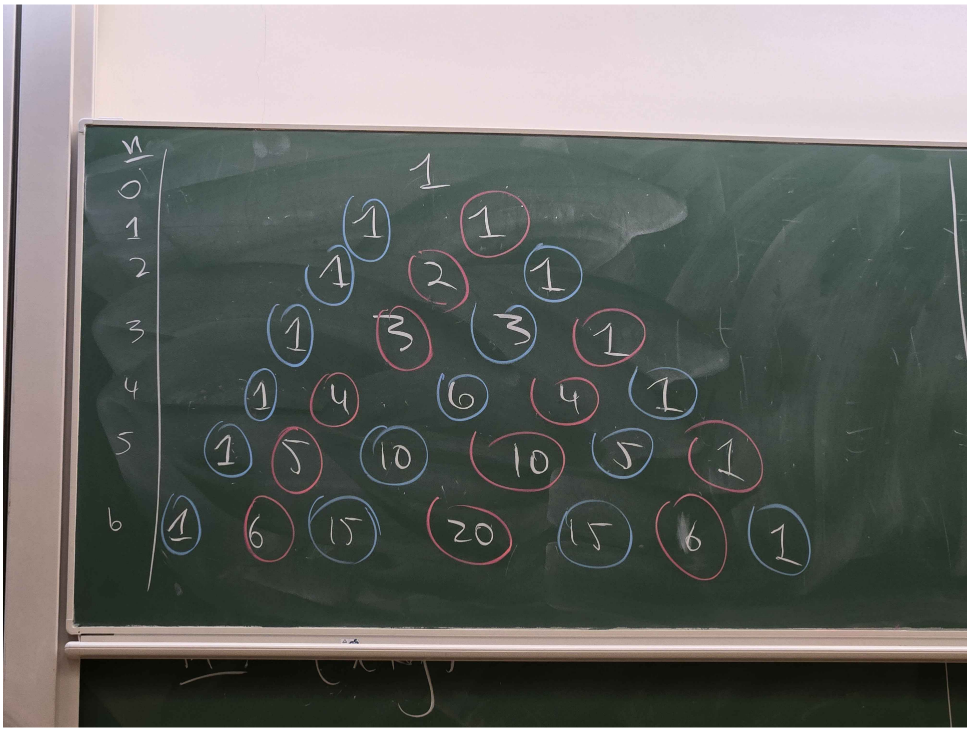


Figure 2.3: The sum of even terms is equal to the sum of odd terms.

Theorem 2.5.2. $\forall n \geq k$, we have
$$\binom{n}{k} = \binom{n}{n-k}.$$

Proof.
$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n!}{(n-k)!k!} = \frac{n!}{(n-(n-k))!} = \binom{n}{n-k}.$$

Remark 2.5.1. Choosing a subset of k elements from n is equivalent to choose $n-k$ elements to discard, and we can build a bijection between these two methods.

■

For n even.
Consider the bijection
$$S \mapsto S \triangle \{n\} = \begin{cases} S - \{n\}, & \text{if } n \in S; \\ S \cup \{n\}, & \text{if } n \notin S. \end{cases}$$

Hence,
$$|S \triangle \{n\}| \in \{|S|-1, |S|+1\},$$
so if $|S|$ is odd, then $S \triangle \{n\}$ is even, and vice versa. We know this is a bijection (self-inverse), so we have odd-sized sets to even-sized set. Hence, $\sum_{2|k} \binom{n}{k} = \sum_{2\nmid k} \binom{n}{k}$.

Example 2.5.5. $x = 2, y = 1$.

Proof.
$$(2+1)^n = 3^n = \sum_{k=0}^{n} \binom{n}{k} 2^k$$

Counting partitions $[n] = A \cup B \cup C$, each element has a choice of 3 sets to go into. Hence, the product rule says there are 3^n partitions, while RHS uses sum rule bases on $k = |A \cup B|$. ⊛

2.6 Divisor Function

Definition 2.6.1 (Divisor Functions). Given a natural number $n \in \mathbb{N}$, let $d(n)$ count the number of divisors of n.

Example 2.6.1.

$$\begin{aligned}
d(1) &= 1 = |\{1\}| \\
d(2) &= 2 = |\{1,2\}| \\
d(3) &= 2 = |\{1,3\}| \\
d(4) &= 3 = |\{1,2,4\}| \\
d(5) &= 2 = |\{1,5\}|.
\end{aligned}$$

Corollary 2.6.1. $d(n) = 2$ if and only if n is a prime.

Now we want to compute the average value of $d(n)$.

Definition 2.6.2.

$$\bar{d}(n) = \frac{\sum_{i=1}^n d(i)}{n}.$$

We can use double-counting. First, notice that

$$d(i) = \sum_{\substack{j\in[i]\\ j\mid i}} 1.$$

Hence,

$$\sum_{i=1}^n d(i) = \sum_{i=1}^n \sum_{\substack{j\in[i]\\ j\mid i}} 1.$$

We can exchange the order of summation:

$$n\bar{d}(n) = \sum_{i=1}^n d(i) = \sum_{i=1}^n \sum_{j:j\mid i} 1 = \sum_{j=1}^n \sum_{\substack{i\in[n]\\ j\mid i}} 1.$$

For fixed j, we know

$$\sum_{\substack{i\in[n]\\ j\mid i}} 1 = \left\lfloor \frac{n}{j} \right\rfloor.$$

Hence, we have

$$n\bar{d}(n) = \sum_{j=1}^n \left\lfloor \frac{n}{j} \right\rfloor,$$

which is equivalent to

$$\bar{d}(n) = \frac{1}{n}\sum_{j=1}^n \left\lfloor \frac{n}{j} \right\rfloor.$$

Observe that

$$\frac{n}{j} - 1 \le \left\lfloor \frac{n}{j} \right\rfloor \le \frac{n}{j},$$

so

$$H_n - 1 = \frac{1}{n}\sum_{j=1}^{n}\left(\frac{n}{j} - 1\right) \leq \overline{d}(n) \leq \frac{1}{n}\sum_{j=1}^{n}\frac{n}{j} = \sum_{j=1}^{n}\frac{1}{j} = H_n \approx \ln n.$$

Hence,

$$H_n - 1 \leq \overline{d}(n) \leq H_n,$$

which gives $\overline{d}(n) \sim \ln n$.

Chapter 3

Partitions

How many ways can we divide n items into k groups? Need to specify details to get well-posed questions.

1. Items distinguishable or not?
2. Groups distinguishable or not?
3. Can we have empty groups? Can we have group with more than one item?

Example 3.0.1. Professor has 49 students, to distribute 3000% between the students.

Proof. Indistinguishable items: percentage points.
Distinguishable groups: students $k = 49$. No restriction on sizes of groups. Formally, we are enumerating

$$S = \left\{ (x_1, x_2, \ldots, x_{49}) \mid x_i \geq 0, x_i \in \mathbb{Z}, \sum_{i=1}^{49} x_i = 3000 \right\}$$

⊛

Lecture 4

3.1 Number of nonnegative integer solution to $x_1 + \cdots + x_k = n$ 12 Sep. 12:20

We can represent solutions using a "stars and bar" diagaram:

- n stars represent the items
- $k - 1$ bars to divides the groups

Example 3.1.1. $x_1 = 3, x_2 = 1, x_3 = 0, x_4 = 5$. $(k = 4, n = 9)$

Proof.

$$\underbrace{***}_{x_1} \mid \underbrace{*}_{x_2} \mid\mid \underbrace{*****}_{x_3}$$

⊛

Hence, we can use a projection between solution and diagrams with $k - 1$ bars and n stars.

Each diagram consists of $n + k - 1$ symbols. Once we know which are the bars, we know the full diagram.

$$\text{number of diagrams} = \binom{n+k-1}{k-1} = \binom{n+k-1}{n}$$

Proposition 3.1.1. The number of non-negative integer solutions to $x_1 + \cdots + x_k = n$ is $\binom{n+k-1}{k-1}$.

Now we have a new problem.

Question. How many solutions are there to $x_1 + \cdots + x_k = n$ with $x_i \geq 1$ for all i?

We can let $y_i = x_i - 1$, then $y_i \geq 0$ and $y_1 + \cdots + y_k = n - k$. Hence, the answer is

$$\binom{(n-k)+(k-1)}{k-1} = \binom{n-1}{k-1}.$$

Definition 3.1.1 (Multisets). An unordered collection of elements with repretition allowed.

$$\{\{1,1,1,2,3\}\} \neq \{\{1,2,3\}\}$$

can be represented as an ordered tuple in increasing order.

Example 3.1.2. How many multisets of size n are there from a set of size k?

Proof. Let x_i be the multiplicites of the i-th element in the multiset. Then $x_i \geq 0$ and

$$x_1 + \cdots + x_k = n.$$

Hence, the number of multisets is

$$\binom{n+k-1}{k-1}.$$

⊛

Alternatively, multisets are $(a_1, \ldots, a_n)$ with $1 \leq a_1 \leq \cdots \leq a_n \leq k$. Now if we let $b_i = a_i + i - 1$, then

$$(b_1, \ldots, b_n) = (a_1, a_2 + 1, \ldots, a_n + n - 1) \text{ with } 1 \leq b_1 < b_2 < \cdots < b_n \leq n + k - 1.$$

Note that there is a bijection between $\{(a_1, \ldots, a_n)\}$ and $\{(b_1, \ldots, b_n)\}$. This shows the number of multisets of size n from $[k]$ is the number of subsets of $[n+k-1]$ of size n, which is

$$\binom{n+k-1}{n} = \binom{n+k-1}{k-1}.$$

Now we add some new setting.

- Distinguishable items
- Indistinguishable groups
- Groups non-empty.

The objects we are counting is

$$\{S_1, S_2, \ldots, S_k\}$$

with $S_1 \cup S_2 \cup \ldots \cup S_k = [n]$ and $S_i \neq \varnothing$ for all i.

Definition 3.1.2 (The Stirling Number of the second kind). $S(n,k)$ is defined to be number of partitions of n distinct items into k indistinguishable non-empty groups.

Example 3.1.3. $S(n,1) = 1$ for all $n \geq 1$. $S(n,n) = 1$ for all n. $S(n, n-1) = \binom{n}{2}$ for all $n \geq 2$. $S(n,2) = 2^{n-1} - 1$.

Proof. We just talk about the $S(n,2)$ one. Since we can choose any subset of $[n]$, so there are 2^n possibilities, but each partition is counted twice, so we have to divide it by 2, and subtract the partition that includes empty group, so it is $2^{n-1} - 1$. ⊛

Proposition 3.1.2. For all n, k,

$$S(n,k) = S(n-1,k-1) + kS(n-1,k).$$

Proof. Case analysis:

- Case 1: $\{n\}$ is a group.
  This means the remaining $n-1$ elements are partitioned into $k-1$ groups, so there are $S(n-1,k-1)$ possibilities.
- Case 2: $\{n\}$ is not a group.
  $n-1$ left elments is first partitioned into k groups, then we can distribute the n-th element into each group, so there are $kS(n-1,k)$ possibilities.

By sum rule, we know

$$S(n,k) = S(n-1,k-1) + kS(n-1,k).$$

■

Example 3.1.4. Using induction to prove

$$S(n,n-1) = \binom{n}{2}.$$

Proof.

$$\begin{aligned} S(n,n-1) = S(n-1,n-2) + (n-1)S(n-1,n-1) &= S(n-1,n-2) + (n-1) \\ = \cdots = 1 + 2 + \cdots + n - 1 &= \binom{n}{2}. \end{aligned}$$

⊛

Now what if the groups are distinguishable? Also, we have

- items distinguishable
- groups distinguishable
- groups non-empty.

Short answer: $S(n,k)k!$.

Lecture 5

We can observe that the number of ways of partitioning n distinct items into k distinct nonempty groups is $S(n,k)k!$. 16 Sep. 15:30

Question. How many ways can we partition n distinct items into l distinct groups (not necessarily nonempty)?

Answer. l^n: product rule, each element has l choice for which group to go to. ⊛

Alternative method. Count by the number of nonempty groups (k), and then use sum rule. Partition elements into k nonempty indistinguishable groups, which has $S(n,k)$ choices, and then map the k sets to the l groups injectively, so there are $l^{\underline{k}} = l(l-1)\ldots(l-k+1)$ choices. Hence, the total number of partition is

$$\sum_{k=0}^{l} S(n,k)l^{\underline{k}}.$$

By double counting, we know

$$l^n = \sum_{k=0}^{n} S(n,k) l^{\underline{k}} = \sum_{k=0}^{n} S(n,k) l^{\underline{k}}.$$

■

Proposition 3.1.3. For any field F, and $x \in F$, $n \in \mathbb{N} \cup \{0\}$, then

$$x^n = \sum_{k=0}^{n} S(n,k) x^{\underline{k}}.$$

(We define $x^{\underline{k}} = x(x-1)\dots(x-(k-1))$.)

Proof. There are polynomials of degree $\leq n$ that agree for all $x \in \mathbb{N}$, so they must agree everywhere. ■

We can observe that $\{x^n \mid n \in \mathbb{N} \cup \{0\}\}$ forms a basis for

$$F[x] = \left\{ \sum_{k=0}^{n} a_k x^k : a_k \in F \right\}.$$

Since x^n is a linear combination of $\{x^{\underline{n}} \mid n \in \mathbb{N} \cup \{0\}\}$, that means this is also a basis for $F[x]$. And the proposition shows that the change of basis matrix is the matrix of Stirling numbers of the second kind:

$$\begin{pmatrix} 1 & & & & & 0 & 0 \\ & 1 & & & & & 0 \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ S(n,k) & & & & & & 1 \end{pmatrix} \begin{pmatrix} x^{\underline{0}} \\ x^{\underline{1}} \\ x^{\underline{2}} \\ \vdots \\ x^{\underline{k}} \\ \vdots \end{pmatrix} = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ \vdots \\ x^k \\ \vdots \end{pmatrix}.$$

3.2 Stirling numbers of the first kind

Recall the permutation π is a bijection from $[n]$ to $[n]$.

Example 3.2.1. $\pi = 32154$, then $\pi(1) = 3, \pi(2) = 2, \pi(3) = 1, \pi(4) = 5, \pi(5) = 4$.

Example 3.2.2. $\pi_1 = 312, \pi_2 = 213$, then $\pi_2 \circ \pi_1 = 321$ and $\pi_1 \circ \pi_2 = 132$.

Claim 3.2.1. $\forall \pi \in S_n$, $\forall x \in [n]$, $\exists i \in [n]$ s.t. $\pi^i(x) = x$.

Proof. Consider $\pi^1(x), \pi^2(x), \dots, \pi^n(x) \in [n]$, if any are equal to x, then we're done. Otherwise, there are only $n-1$ possible values, which are $[n] \setminus \{x\}$. Hence, there are some $j_1, j_2 \in [n]$ with $j_1 > j_2$ and $\pi^{j_1}(x) = \pi^{j_2}(x)$ by Pigeonhole principle. Applying π^{-1} for j_2 times, we get

$$\pi^{j_1 - j_2}(x) = x \quad \text{with } 1 \leq j_1 - j_2 \leq n,$$

which is a contradiction. ■

Definition 3.2.1 (cycle). For the smallest i, $1 \leq i \leq n$ with $\pi^i(x) = x$, we say

$$(x \ \pi(x) \ \pi^2(x) \ \dots \ \pi^{i-1}(x))$$

is the cycle of x.

It follows that every permutation is a union of disjoint cycles. Hence, we have cycle representation of π.

Example 3.2.3. $\pi = 32154$, the cycle form is $(13)(2)(45)$.

Definition 3.2.2 (fixed point and transposition)**.** A fixed point of a permutation is a cycle of length 1 i.e. an element x with $\pi(x) = x$. A transposition is a cycle of length 2. A permutation is cyclic if it has a single cycle (of length n).

Question. How many cyclic permutations of $[n]$ are there?

Answer. $(n-1)!$. We can first fix the head of the cycle to be 1, then for $\pi(1)$, we have $n-1$ choices, and for $\pi^2(1)$, we have $n-2$ choices, and so on, so we have $(n-1)!$ cyclic permutations.

> **Note 3.2.1.** Who is in the head of the cycle is not important.

⊛

Definition 3.2.3 (The Stirling numbers of the first kind)**.** $s_{n,k}$ (or $[s(n,k)]$) enumerate the permutation in S_n with exactly k cycles.

Example 3.2.4. $s_{n,1} = (n-1)!, s_{n,n} = 1, s_{n,n-1} = \binom{n}{2}, s_{n,2} =$not so obvious.

Proof.

$$s_{n,2} = \frac{1}{2}\sum_{k=1}^{n-1}\binom{n}{k}(k-1)!(n-k-1)!$$

Note that we multiply it by $\frac{1}{2}$ since we count each cycle-pair twice. Also, we know that a cycle of length n has $(n-1)!$ choices if we fix all n members in the cycle.

Alternatively, say the "first" cycle is the one containing 1 together with $0 \le k \le n-2$ other elements. Hence, we have

$$\begin{aligned}
s_{n,2} &= \sum_{k=0}^{n-2}\binom{n-1}{k}(k!)(n-k-2)! \\
&= \sum_{k=0}^{n-2}\frac{(n-1)!}{k!(n-k-1)!}k!(n-k-2)! = (n-1)!\sum_{k=0}^{n-2}\frac{1}{n-1-k} \\
&= (n-1)!\sum_{k=1}^{n-1}\frac{1}{k} \\
&= (n-1)!H_{n-1} \approx (n-1)!\ln n.
\end{aligned}$$

⊛

Proposition 3.2.1. $\forall n, k \ge 1$,

$$s_{n,k} = s_{n-1,k-1} + (n-1)s_{n-1,k}$$

Proof. Case analysis: is n a fixed point?

- Case 1: Yes. Removing it, and then the left $n-1$ elements can be permutated with $k-1$ cycles. Hence, there are $s_{n-1,k-1}$ choices.

- Case 2: No. We remove n from a cycle to get a permutation of $[n-1]$ with k cycles. Now, we have $n-1$ place to insert n inside. For example, we if $n=7$, and we have $(13)(2)(456)$, then we have $7-1=6$ places to insert 7 inside since (7456) and (4567) are same cycles.

To create a permutation $\pi \in S_n$ with k cycles where n is not a fixed point, we can take a permutation $\pi' \in S_{n-1}$ with k cycles, which has $s_{n-1,k}$ choices, and insert n before any element, so there are $n-1$ ways, so the number of such permutation is $(n-1)s_{n-1,k}$. By sum rule, we have

$$s_{n,k} = s_{n-1,k-1} + (n-1)s_{n-1,k}.$$

■

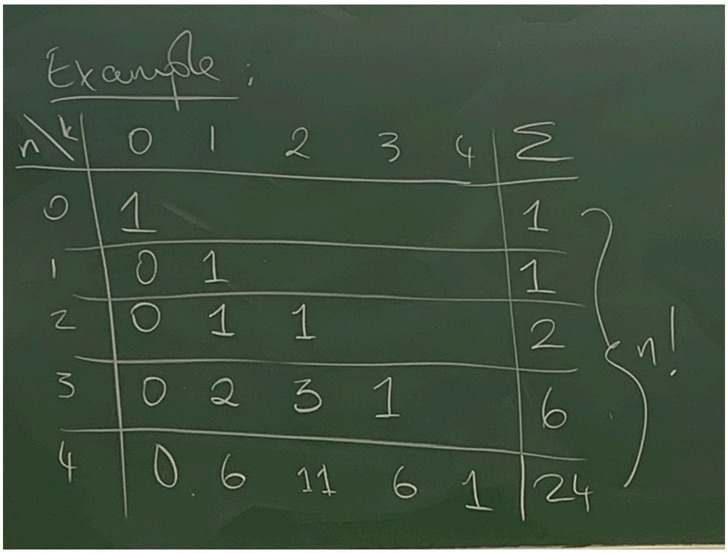


Figure 3.1: table of $s_{n,k}$

Corollary 3.2.1. $\forall n$, we have

$$\sum_{k=0}^{n} s_{n,k} = n!.$$

Proof. The number of permutations are $n!$, and every permutation consists of i cycles where $1 \leq i \leq n$, and then apply the sum rule. ■

Notation. Given $x \in F$, and $k \in \mathbb{N} \cup \{0\}$, we have

- $x^{\underline{k}} = x(x-1)\ldots(x-(k-1))$
- $x^{\overline{k}} = x(x+1)\ldots(x+(k-1)) = (x+k-1)^{\underline{k}}$.

Proposition 3.2.2. For all $x \in F$, $n \in \mathbb{N} \cup \{0\}$,

$$x^{\overline{n}} = \sum_{k=0}^{n} s_{n,k} x^k.$$

Proof. Induction on n. We know it is true for $n = 0, 1$. Note that

$$\begin{aligned} x^{\overline{n}} &= x^{\overline{n-1}}(x+n-1) \\ &= (x+n-1)\sum_{k=0}^{n-1} s_{n-1,k}x^k \\ &= x\sum_{k=0}^{n-1} s_{n-1,k}x^k + (n-1)\sum_{k=0}^{n-1} s_{n-1,k}x^k \\ &= \sum_{k=0}^{n-1} s_{n-1,k}x^{k+1} + \sum_{k=0}^{n-1}(n-1)s_{n-1,k}x^k \\ &= \sum_{k=1}^{n} s_{n-1,k-1}x^k + \sum_{k=0}^{n-1}(n-1)s_{n-1,k}x^k \\ &= \sum_{k=0}^{n} \left(s_{n-1,k-1} + (n-1)s_{n-1,k}\right)x^k \\ &= \sum_{k=0}^{n} s_{n,k}x^k. \end{aligned}$$

■

Corollary 3.2.2.

$$x^{\underline{n}} = \sum_{k=0}^{n} \underbrace{(-1)^{n-k}s_{n,k}}_{\substack{\text{signed Stirling numbers} \\ \text{of the first kind}}} x^k.$$

Proof.

$$\begin{aligned} x^{\underline{n}} &= x(x-1)\dots(x-(n-1)) \\ &= (-1)^n(-x)(-x+1)\dots(-x+(n-1)) \\ &= (-1)^n(-x)^{\overline{n}} \\ &= (-1)^n\sum_{k=0}^{n} s_{n,k}(-x)^k \\ &= \sum_{k=0}^{n}(-1)^{n-k}s_{n,k}x^k. \end{aligned}$$

■

Lecture 6

19 Sep. 12:20

Corollary 3.2.3.

$$\sum_{k=j}^{i}(-1)^{k-j}S(i,k)s_{k,j} = \delta_{i,j} = \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{if } i \neq j. \end{cases}$$

Proof. By Proposition 3.1.3, we have

$$\begin{aligned} x^i &= \sum_{k=0}^{i} S(i,k)x^{\underline{k}} = \sum_{k=0}^{i} S(i,k)\left[\sum_{j=0}^{k}(-1)^{k-j}s_{k,j}x^k\right] \\ &= \sum_{k=0}^{i}\sum_{j=0}^{k}(-1)^{k-j}S(i,k)s_{k,j}x^j \\ &= \sum_{j=0}^{i}\left(\sum_{k=j}^{i}(-1)^{k-j}S(i,k)s_{k,j}\right)x^j = x^i. \end{aligned}$$

Since $\{x^0, x^1, x^2, \dots\}$ is a basis of $F[x]$, the coefficient of x^j is 1 if $i = j$ and is 0 if $i \neq j$. ■

Question. How many ways can we distribute \$100000 of prize money to six players in the tournaments?

- Whole dollars only.
- Nonnegative prices.

It is an arbitrary partition, and there are $k = 6$ distinct groups(players). Hence, there are $\binom{1000005}{5}$ ways of distribution? However, this is not what we want, since in a tournament a better player should get more money. Actually, in this scenario, groups are indistinguishable since largest prize is for first place, and so on. Thus, our goal is to dividing n indistinguishable items into k indistinguishable (non-empty) groups.

Definition 3.2.4 (number partition). A number partition is a decomposition of n and a sum of k unordered natural numbers.

$$\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k) \text{ s.t. } \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k, \; \sum_{i=1}^{k}\lambda_i = n \text{ with } \lambda_i \in \mathbb{N}.$$

We write $\lambda \vdash n$. We define

$$p(n,k) = |\{\lambda = (\lambda_1, \dots, \lambda_k) : \lambda \vdash n\}|.$$

We also define

$$p(n, \leq k) = \sum_{i=0}^{k} p(n,i)$$

$$p(n) = p(n, \leq n) = \sum_{i=0}^{n} p(n,i).$$

Observe that

- $$p(n,0) = \begin{cases} 1, & \text{if } n = 0; \\ 0, & \text{if } n \geq 1. \end{cases}$$
- $p(n,n) = 1$
- $p(n,n-1) = 1 = |\{2,1,1,\dots\}|$
- $p(n,1) = 1$.
- $p(n,2) = \lfloor \frac{n}{2} \rfloor$.

Proposition 3.2.3. $\forall n \geq k \geq 1$,

$$p(n, k) = p(n-1, k-1) + p(n-k, k).$$

Proof. Case analysis based on size of smallest part:

- Case 1: $\lambda_k = 1$.
  Then remove the last part to get a partition of $n-1$ into $k-1$ nonempty parts. (bijective, can add part of size 1 to the end of a partition), so there are $p(n-1, k-1)$ such cases.

- Case 2: $\lambda_k \geq 2$.
  Consider $\lambda' = (\lambda_1 - 1, \lambda_2 - 1, \ldots, \lambda_k - 1)$, then $\lambda' \vdash n - k$, and this is a bijection, so there are $p(n-k, k)$ such cases.

∎

Lecture 7

Definition 3.2.5 (Ferrers diagram)**.** Visual representation of $\lambda \vdash n$. Each λ_i pirctured as a row of λ_i dots. 23 Sep. 15:30

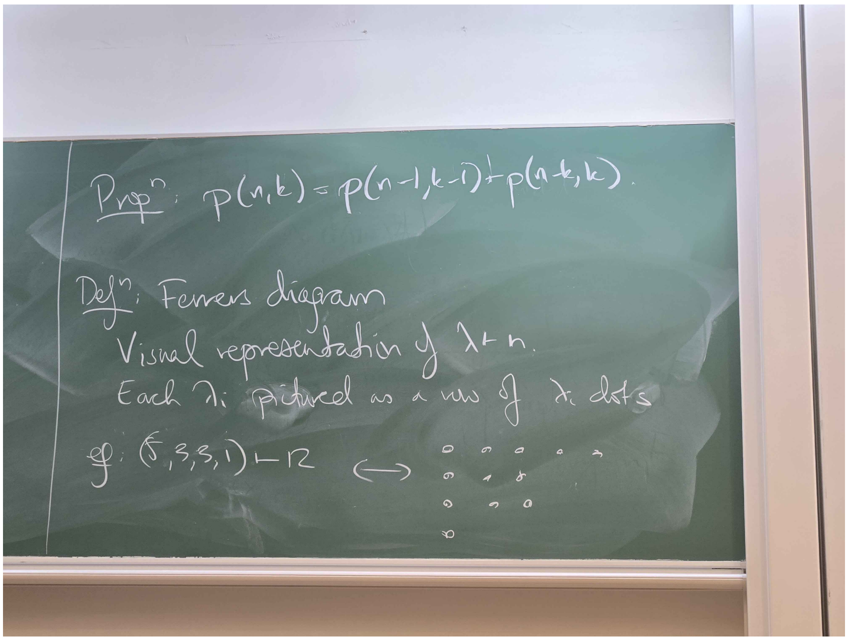


Figure 3.2: Ferrers diagram

Note 3.2.2. If we see the Ferrers diagram from the columns, then note that the number of dots in the columns is decreasing.

Definition 3.2.6. Given a parition $\lambda \vdash n$, the conjugate partition $\lambda^* \vdash n$ is given by

$$\lambda_j^* = |\{i : \lambda_i \geq j\}|.$$

Visually, λ^* is the partition obtained by reflecting λ in the diagonal $y = -x$.

Observe that λ^* is indeed a partition of n:

$$\lambda_1^* \geq \lambda_2^* \geq \ldots$$

is obvious from the definition, and

$$\sum_j \lambda_j^* = \sum_j |\{i : \lambda_i \geq j\}| = \sum_i \lambda_i = n.$$

Also, note that $(\lambda^*)^* = \lambda$.

Proposition 3.2.4. The number of partition of n into at most k parts = The number of partitions of n into parts of size $\leq k$.

Proof. The largest part of λ is the number of parts in λ^*. And so conjugation gives a bijection between these two choices of partition of n. ■

Definition 3.2.7. A parition $\lambda \vdash n$ is called self-conjugate if $\lambda^* = \lambda$.

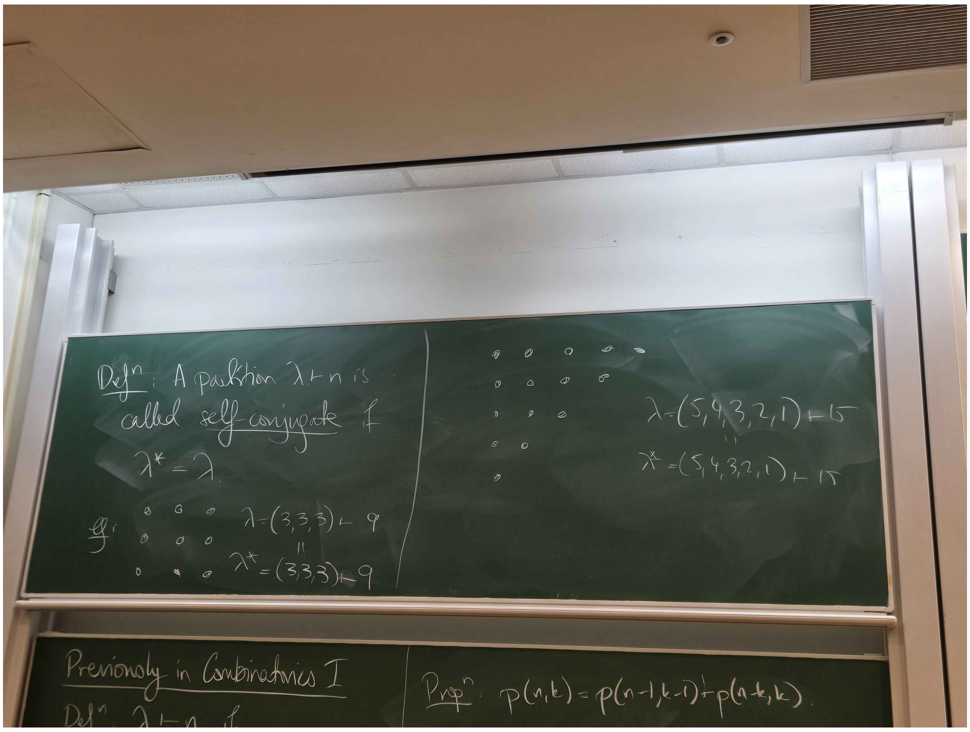


Figure 3.3: Self-conjugate

Proposition 3.2.5. The number of self-conjugate partition of n is the number of partition of n into distinct odd parts, which means

$$(\lambda_1, \lambda_2, \ldots, \lambda_k) : \lambda_1 > \lambda_2 > \cdots > \lambda_k \geq 1, \quad \forall 1 \leq i \leq k,\ \lambda_i \equiv 1 \mod 2.$$

Proof. Let λ be a self-conjugate partition. (See Figure 3.4) If we consider the dots in the first row or column (we called it a hook), since $\lambda = \lambda^*$, we have $2\lambda_1 - 1$ dots, which is an odd part. If we take the i-th part of the new partition to be the points in the i-th row or i-th column not-yer counted, then we get

$$(\lambda_i - (i-1)) + (\lambda_i - (i-1)) - 1,$$

say $\mu_i = 2(\lambda_i - (i-1)) - 1$, then $\mu \vdash n$ and

$$\begin{aligned} \mu_{i+1} &= 2\lambda_{i+1} - 2(i+1) + 1 \\ &\leq 2\lambda_i - 2(i+1) + 1 \\ &< 2\lambda_i - 2i + 1 = \mu_i, \end{aligned}$$

so μ has distinct parts and clearly μ_i is odd for all i. Hence, we have mapped our self-conjugate λ into a partition μ with distinct odd parts. This is indeed a bijeciton.

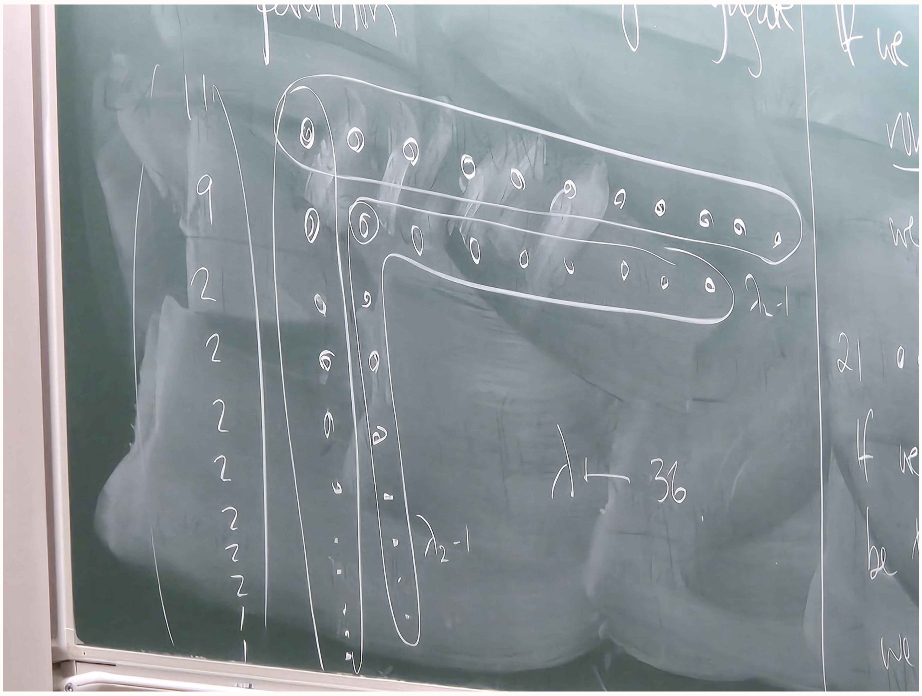


Figure 3.4: Use hook to obtain bijection

■

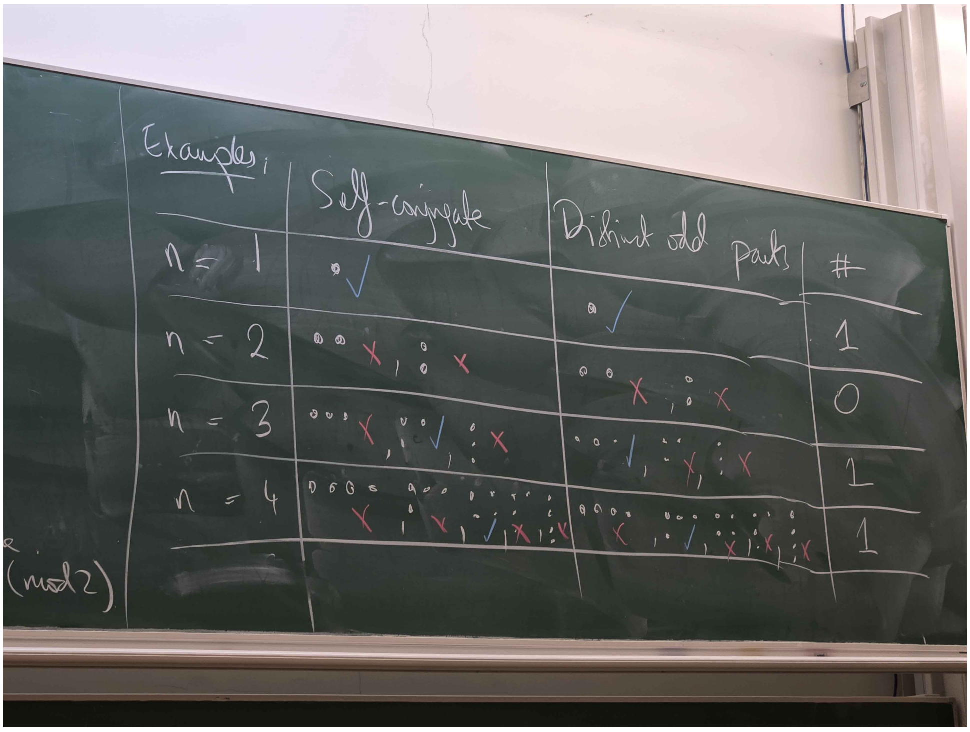


Figure 3.5: Some cases of small n.

Example 3.2.5. Square partition $\lambda = \underbrace{(k, k, \dots, k)}_{k \text{ parts}} \vdash k^2$ are self conjugate.

Corollary 3.2.4. The sum of the first k odd numbers is k^2.

Proof. By drawing hooks, it is trivial.

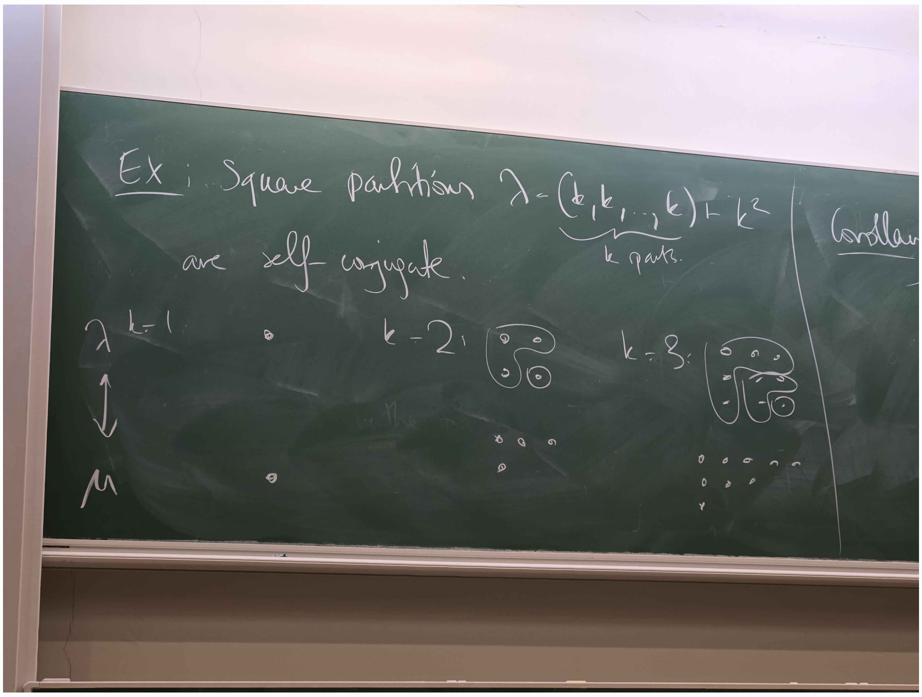


Figure 3.6: Drawing hooks to get the first k odd numbers from a square

⊛

3.3 The twelvefold way of Counting

Question. How many ways can we partition n items into k groups?

Itmes	Groups	Partition
numbered indistinguisable	numbered indistinguishable	injective(group of size ≤ 1) surjective(group of size ≥ 1) arbitrary

Table 3.1: All types of partition problem.

	Injective	Surjective	Arbitrary
Items, groups numbered	$k^{\underline{n}}$	$S(n,k)\cdot k!$	k^n
Items numbered, groups not	$\begin{cases} 1, & \text{if } k \geq n; \\ 0, & \text{if } k < n. \end{cases}$	$S(n,k)$	$\sum_{j=0}^{k} S(n,j)$
Items not, groups numbered	$\binom{k}{n}$	$\binom{n-1}{k-1}$	$\binom{n+k-1}{k-1}$
Items, groups not numbered	$\begin{cases} 1, & \text{if } k \geq n; \\ 0, & \text{if } k < n. \end{cases}$	$p(n,k)$	$\sum_{j=0}^{k} p(n,j)$

Table 3.2: All solution to all kinds of partition problem

Chapter 4

Generating Functions

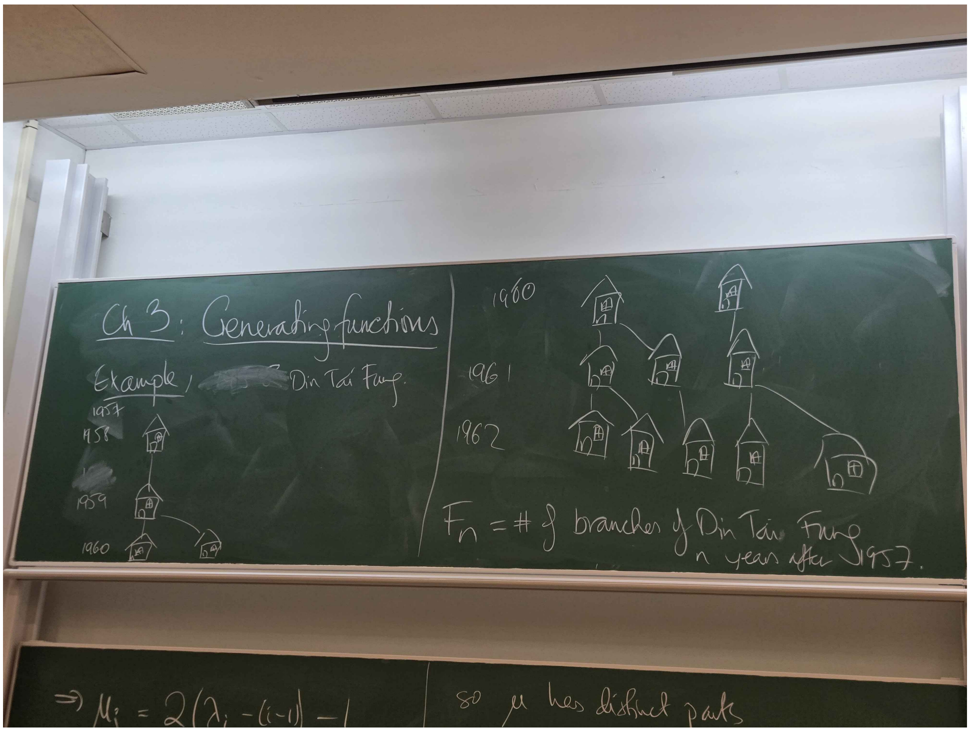


Figure 4.1: Din Tai Fung branches number

We have a recurrence relation: $\forall n \geq 2$

$$F_n = F_{n-1} + F_{n-2}$$

Example 4.0.1. If

$$F'_n = F'_{n-1} + F'_{n-1},$$

then $F'_n = 2^n F'_0$.

Suppose $\{F_n\}_{n=0}^{\infty}$ is a recurring sequence with $F_0 = 0, F_1 = 1$, then we can define a power series as

$$F(x) = F_0 + F_1 x + F_2 x^2 + \cdots = \sum_{n=0}^{\infty} F_n x^n.$$

Thus, we have

$$xF(x) = F_0 x + F_1 x^2 + \cdots = \sum_{n=0}^{\infty} F_n x^{n+1} = \sum_{n=1}^{\infty} F_{n-1} x^n.$$

If we do it again, then we can get

$$x^2F(x) = F_0x^2 + F_1x^3 + \cdots = \sum_{n=0}^{\infty} F_n x^{n+2} = \sum_{n=2}^{\infty} F_{n-2}x^n.$$

Now we have

$$F(x) - xF(x) - x^2F(x) = F_0x^0 + F_1x^1 - F_0x^1 + \sum_{n=2}^{\infty} \underbrace{(F_n - F_{n-1} - F_{n-2})}_{=0} x^n = x.$$

Hence, $(1-x-x^2)F(x) = x$, and thus

$$F(x) = \frac{x}{1-x-x^2} = \frac{A}{1-\alpha_1 x} + \frac{B}{1-\alpha_2 x}.$$

Now we solve the A, B, α_1, α_2.

$$\begin{aligned}\frac{A}{1-\alpha_1} + \frac{B}{1-\alpha_2} &= \frac{A(1-\alpha_2 x) + B(1-\alpha_1 x)}{(1-\alpha_1 x)(1-\alpha_2 x)} \\ &= \frac{(A+B) - (A\alpha_2 + B\alpha_1)x}{1-(\alpha_1+\alpha_2)x + \alpha_1\alpha_2 x^2} = \frac{x}{1-x-x^2}.\end{aligned}$$

Hence, we want

$$\begin{cases} A + B = 0 \\ A\alpha_2 + B\alpha_1 = -1 \\ \alpha_1 + \alpha_2 = 1 \\ \alpha_1\alpha_2 = -1 \end{cases},$$

by solving α_1, α_2 first, we can get $\alpha_1 = \frac{1+\sqrt{5}}{2}$ and $\alpha_2 = \frac{1-\sqrt{5}}{2}$, and thus we can solve $A = \frac{1}{\sqrt{5}}$ and $B = -\frac{1}{\sqrt{5}}$. Hence, we have

$$F(x) = \frac{x}{1-x-x^2} = \frac{\frac{1}{\sqrt{5}}}{1-\left(\frac{1+\sqrt{5}}{2}\right)x} - \frac{\frac{1}{\sqrt{5}}}{1-\left(\frac{1-\sqrt{5}}{2}\right)x}.$$

Now since we know

$$\frac{1}{1-\alpha} = 1 + \alpha + \alpha^2 + \dots,$$

so we can get

$$\begin{aligned}F(x) &= \frac{1}{\sqrt{5}}\left(\left(1 + \left(\frac{1+\sqrt{5}}{2}\right)x + \left(\left(\frac{1+\sqrt{5}}{2}\right)x\right)^2 + \dots\right) - \left(1 + \left(\frac{1-\sqrt{5}}{2}x + \left(\left(\frac{1-\sqrt{5}}{2}\right)x\right)^2 + \dots\right)\right)\right) \\ &= \frac{1}{\sqrt{5}}\sum_{n=0}^{\infty}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right)x^n = \sum_{n=0}^{\infty} F_n x^n.\end{aligned}$$

Hence, we have

$$F_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right).$$

Lecture 8

Observe that 26 Sep. 12:20

$$\left|\frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n\right| < \frac{1}{2}.$$

Hence, F_n is the integer closed to

$$\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n.$$

The idea is to encode a sequence of numbers

$$a_0, a_1, a_2, \ldots$$

as coefficients in a power series

$$A(x) = \sum_{n=0}^{\infty} a_n x^n.$$

Proposition 4.0.1. Let $(a_0, a_1, \ldots)$ be a sequence of real numbers. If $|a_n| < K^n$ for all $n \in \mathbb{N}$, then

$$\forall x \in \left(-\frac{1}{K}, \frac{1}{K}\right), \text{ we have } A(x) = \sum_{n=0}^{\infty} a_n x^n$$

converges absolutely.

Proof. Suppose $x \in \left(-\frac{1}{K}, \frac{1}{K}\right)$, then

$$A(x) = \sum_{n=0}^{\infty} |a_n x^n| \leq \sum_{n=0}^{\infty} |K^n x^n| = \sum_{n=0}^{\infty} (|Kx|)^n,$$

which is a geometric series, and since $|Kx| < 1$, so it converges. ∎

$A(x)$ has derivatives of all orders at $x = 0$, and for all $n \geq 0$,

$$A^{(n)}(0) = a_n n!.$$

In particular, the values of $A(x)$ around the origin determine this sequence (a_n) uniquely. We treat $A(x)$ as a formal power series. Thus, we can usually eaily verigy results using induction.

Definition 4.0.1. Given a sequence $(a_0, a_1, \ldots)$ of real numbers, the generating function of the sequence is the (formal) power series

$$\sum_{n=0}^{\infty} a_n x^n.$$

Example 4.0.2. Suppose we have a sequence $(1, 1, 1, \ldots)$, then

$$A(x) = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

converges for $|x| < 1$.

Example 4.0.3. Suppose we have a sequence $\left(0, 1, \frac{1}{2}, \ldots\right)$, then

$$A(x) = \sum_{n=0}^{\infty} \frac{x^n}{n} = -\ln(1-x)$$

converges for $|x| < 1$.

Example 4.0.4. Suppose we have a sequence $\left(1, 1, \frac{1}{2}, \ldots, \frac{1}{n!}, \ldots\right)$, then

$$A(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$$

converges for all $x \in \mathbb{R}$.

Example 4.0.5. Suppose r is a fixed number and we have a sequence

$$\left(\binom{r}{0}, \binom{r}{1}, \dots\right),$$

then

$$A(x) = \sum_{n=0}^{\infty} \binom{r}{n} x^n = (1+x)^r.$$

converges for $|x| < 1$.

Remark 4.0.1. The special case:

$$\begin{aligned}\frac{1}{(1-x)^t} = (1-x)^{-t} &= \sum_{n=0}^{\infty} \binom{-t}{n} (-x)^n = \sum_{n=0}^{\infty} \binom{-t}{n} (-1)^n x^n \\ &= \sum_{n=0}^{\infty} \binom{t+n-1}{n} x^n\end{aligned}$$

since

$$(-1)^n \binom{-t}{n} = (-1)^n \frac{(-t)(-t-1)\dots(-t-n+1)}{n!} = \frac{t(t+1)\dots(t+n-1)}{n!} = \binom{t+n-1}{n}.$$

4.1 Dictionary for operations

- Sum:

$$\begin{aligned}A(x) &\sim (a_0, a_1, \dots) \\ B(x) &\sim (b_0, b_1, \dots) \\ A(x) + B(x) &\sim (a_0 + b_0, a_1 + b_1, \dots)\end{aligned}$$

- Scalar multiplication:

$$\begin{aligned}A(x) &\sim (a_0, a_1, \dots) \\ \lambda A(x) &\sim (\lambda a_0, \lambda a_1, \dots) \quad \forall \lambda > 0.\end{aligned}$$

- Shifting to the right:

$$\begin{aligned}(a_0, a_1, \dots) &\sim \sum_{n=0}^{\infty} a_n x^n \\ (0, a_0, a_1, \dots) &\sim \sum_{n=1}^{\infty} a_{n-1} x^n = x \sum_{n=0}^{\infty} a_n x^n \\ A(x) &\to xA(x)\end{aligned}$$

Note 4.1.1. By repeating shifting to the right, we can get

$$x^k A(x) \sim (\underbrace{0, 0, \dots, 0}_{k}, a_0, a_1, \dots).$$

- Shifting to the left:

$$(a_0, a_1, \dots) \sim \sum_{n=0}^{\infty} a_n x^n$$
$$(a_1, a_2 \dots) \sim \sum_{n=1}^{\infty} a_n x^{n-1} = \frac{A(x) - a_0}{x}.$$

> **Note 4.1.2.** By repeating
> $$\frac{A(x) - a_0 - a_1 x - \cdots - a_{k-1} x^{k-1}}{x^k},$$
> we can shift to the left by k terms.

- Substituting λx for x with some $\lambda \in \mathbb{R}$.

$$A(\lambda x) = \sum_{n=0}^{\infty} a_n (\lambda x)^n = \sum_{n=0}^{\infty} (a_n \lambda^n) x^n$$

and it corresponds to $(a_0, \lambda a_1, \lambda^2 a_2, \dots)$.

> **Example 4.1.1.** Suppose we want $(1, \lambda, \lambda^2, \dots)$, then taking $(1, 1, \dots)$ and substituting x by λx, so we will change $\frac{1}{1-x}$ to $\frac{1}{1-\lambda x}$, and this means change $(1, 1, \dots)$ to $(1, \lambda, \lambda^2, \dots)$.

Lecture 9

3 Oct. 12:20

4.2 Recurrence relation

4.2.1 Linear homogeneous constant-coefficient recurrence relations

Suppose

$$a_n = \alpha_{k-1} a_{n-1} + \alpha_{k-2} a_{n-2} + \cdots + \alpha_1 a_{n-k+1} + \alpha_0 a_{n-k} \tag{4.1}$$

holds for all $n \geq k$ and we have initial conditions $a_0, a_1, \dots, a_{k-1}$. Then if we define the generating function:

$$A(x) = \sum_{n=0}^{\infty} a_n x^n,$$

then we have

$$\begin{aligned}
\alpha_{k-1} x A(x) &= \sum_{n=1}^{\infty} \alpha_{k-1} a_{n-1} x^n \\
\alpha_{k-2} x^2 A(x) &= \sum_{n=2}^{\infty} \alpha_{k-2} a_{n-2} x^n \\
&\vdots \\
\alpha_0 x^k A(x) &= \sum_{n=k}^{\infty} \alpha_0 a_{n-k} x^n,
\end{aligned}$$

so we have

$$\begin{aligned}
A(x) \left[1 - \alpha_{k-1} x - \alpha_{k-2} x^2 - \cdots - \alpha_0 x^k\right] &= \sum_{n=k}^{\infty} (a_n - \alpha_{k-1} a_{n-1} - \cdots - \alpha_0 a_{n-k}) x^n + R(x) \\
&= R(x),
\end{aligned}$$

where $R(x)$ is a polynomial of degree $k-1$ depending on coefficient α_i and the initial terms $a_0, a_1, \ldots, a_{k-1}$. Hence, we have

$$A(x) = \frac{R(x)}{1 - \alpha_{k-1}x - \alpha_{k-2}x^2 - \cdots - \alpha_0 x^k}.$$

If

$$1 - \alpha_{k-1}x - \alpha_{k-2}x^2 - \cdots - \alpha_0 x^k = (1-\lambda_1 x)(1-\lambda_2 x)\ldots(1-\lambda_k x),$$

then we have

$$A(x) = \frac{A_1}{1-\lambda_1 x} + \frac{A_2}{1-\lambda_2 x} + \cdots + \frac{A_k}{1-\lambda_k x}.$$

for some constants $A_1, A_2, \ldots, A_k$, which means

$$a_n = A_1\lambda_1^n + A_2\lambda_2^n + \cdots + A_k\lambda_k^n$$

by comparing the n-th coefficient of $A(x)$ and R.H.S.

Definition 4.2.1. Given the recurrence relation Equation 4.1, then the characteristic polynomial is

$$p(z) = z^k - \alpha_{k-1}z^{k-1} - \alpha_{k-2}z^{k-2} - \cdots - \alpha_1 z - \alpha_0.$$

If we let $z = \frac{1}{x}$, then multiplying

$$1 - \alpha_{k-1}x - \alpha_{k-2}x^2 - \cdots - \alpha_{k-1}x^{k-1} - \alpha_0 x^k$$

by z^k, we have

$$z^k - \alpha_{k-1}z^{k-1} - \alpha_{k-2}z^{k-2} - \cdots - \alpha_1 z - \alpha_0.$$

Hence, $(1-\lambda_1 x)(1-\lambda_2 x)\ldots(1-\lambda_k x)$ becomes $(z-\lambda_1)(z-\lambda_2)\ldots(z-\lambda_k)$ and thus

$$\{\lambda_i : 1 \leq i \leq k\}$$

are the roots of $p(z)$.

Note 4.2.1. This method is only true when $\lambda_i \neq \lambda_j$ for all $i \neq j$.

Question. What if there is repeated root?

For example, if

$$p(z) = (z-\lambda_1)(z-\lambda_2)^2,$$

then

$$A(x) = \frac{A_1}{1-\lambda_1 x} + \frac{A_2 + A_3 x}{(1-\lambda_2 x)^2}.$$

Theorem 4.2.1. Suppose a sequence is defined by

$$a_n = \alpha_{k-1}a_{n-1} + \cdots + \alpha_0 a_{n-k} \quad \forall n \geq k$$

with initial conditions $a_0, a_1, \ldots, a_{k-1}$. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the roots of the characteristic polynomial $p(z)$.

(1) If the roots are distinct, then

$$a_n = \sum_{i=1}^{k} A_i \lambda_i^n$$

for constants $A_1, A_2, \ldots, A_k$ determined by $a_0, \ldots, a_{k-1}$.

(2) If we have repeated roots, say

$$p(z) = (z-\lambda_1)^{k_1}(z-\lambda_2)^{k_2}\ldots(z-x_q)^{k_q},$$

then

$$a_n = \sum_{i=1}^{q} \left(\sum_{j=0}^{k_i - 1} C_{ij} n^j \right) \lambda_i^n .$$

Lecture 10

4.3 Generating function operation

7 Oct. 15:30

- Substituting x^k for x, then

$$(a_0, a_1, \dots) \to \left(a_0, \underbrace{0, 0, \dots, 0}_{k}, \dots, \begin{cases} 0, & \text{if } k \nmid n; \\ a_{\frac{n}{k}}, & \text{if } k \mid n. \end{cases} \right)$$

since for $A(x) = \sum_{i=0}^{\infty} a_i x^i$, we have

$$A\left(x^k\right) = \sum_{i=0}^{\infty} a_i \left(x^k\right)^i = \sum_{i=0}^{\infty} a_i x^{ki}.$$

- Differentiation:

$$(a_0, a_1, \dots) \to (a_1, 2a_2, 3a_3, \dots).$$

- Integration:

$$(a_0, a_1, \dots) \to \left(0, \frac{a_1}{1}, \frac{a_2}{2}, \dots\right)$$

since

$$\int_0^x A(t)\,\mathrm{d}t = \sum_{n=0}^{\infty} \int_0^x a_n t^n \,\mathrm{d}t = \sum_{n=0}^{\infty} a_n \left[\frac{t^{n+1}}{n+1}\right]_0^x = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1}.$$

Example 4.3.1. Find the generating functions for the sequences:

(i) $a_n = 2^{\left\lfloor \frac{n}{2} \right\rfloor}$

(ii) $a_n = (n+1)^2$.

Proof.

(i) Note that $(a_i)_{i=0}^{\infty} = (1, 1, 2, 2, 4, 4, 8, 8, 16, 16, \dots)$, and we can write it as

$$(1, 0, 2, 0, 4, 0, 8, 0, 16, 0, \dots) + (0, 1, 0, 2, 0, 4, 0, 8, 0, 16, \dots),$$

where the first term is $B(x) = (1, 2, 4, 8, 16, \dots)$ spread by 2, which is $B\left(x^2\right)$, and the second term is $xB\left(x^2\right)$. Note that $B(x) = \frac{1}{1-2x}$.

(ii) If $b_n = n$ corresponds to $B(x)$, then

$$A(x) = \frac{\mathrm{d}B(x)}{\mathrm{d}x}$$

has $a_n = (n+1)b_{n+1} = (n+1)^2$. Also, if the sequence $c_n = 1$ has generating function $C(x) = \frac{1}{1-x}$, then

$$\frac{\mathrm{d}}{\mathrm{d}x} C(x) \sim (1, 2, 3, \dots),$$

so

$$B(x) = x \frac{\mathrm{d}}{\mathrm{d}x} C(x),$$

and we have

$$A(x)=\left[x\left(\frac{1}{1-x}\right)'\right]'=\frac{2}{(1-x)^3}-\frac{1}{(1-x)^2}.$$

⊛

4.4 Products of Generating Functions

Suppose $A(x)=\sum_{n=0}^{\infty}a_nx^n$ and $B(x)=\sum_{n=0}^{\infty}b_nx^n$,

> **Question.** What can we say about $C(x)=A(x)B(x)$?

$$\begin{aligned}
A(x)B(x)&=\left(\sum_{n=0}^{\infty}a_nx^n\right)\left(\sum_{n=0}^{\infty}b_nx^n\right)\\
&=\sum_{n,m\geq 0}a_nb_mx^{n+m}\\
&=\sum_{r\geq 0}\left(\sum_{\substack{n,m\geq 0\\ n+m=r}}a_nb_m\right)x^r=\sum_{r\geq 0}\left(\sum_{n=0}^{r}a_nb_{r-n}\right)x^r\\
&=\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}a_kb_{n-k}\right)x^n
\end{aligned}$$

i.e. $C(x)$ corresponds to the sequence $c_n=\sum_{k=0}^{n}a_kb_{n-k}$, and we call it a convolution.

Combinatorial interpretation

Suppose (a_n) represents the number of ways of completing task 1 with a budget of \$$n$, and (b_n) represents the number of ways of completing task 2 with a budget of \$$n$. Then the convolution $c_n=\sum_{k=0}^{n}a_kb_{n-k}$ represents the number of ways of completing both task 1 and 2 with a combinatorial budget of \$$n$.

Example 4.4.1. Designing a physics course for n days, which has theoretical part including one midterm exam and it is followed by a practical part, which includes two experiments. This is a convolution of the sequences

$$(a_n)=\#\text{ of ways of planning theory and }(b_n)=\#\text{ of ways of planning practical.}$$

Note that $a_n=n$ and $b_n=\binom{n}{2}$, so

$$A(x)=x\left(\frac{1}{1-x}\right)'=\frac{x}{(1-x)^2}\qquad B(x)=\frac{x^2}{(1-x)^3}$$

since $b_n=\binom{n}{2}=\frac{n(n-1)}{2}$. Hence, $C(x)=\frac{x^3}{(1-x)^5}$ is the generating function for the sequence (c_n) where c_n is the number of ways of designing an n-day course. Now since

$$C(x)=\frac{x^3}{(1-x)^5}=x^3(1-x)^{-5}$$

and

$$(1-x)^{-5}=\sum_{n=0}^{\infty}\binom{-5}{n}(-x)^n,$$

and

$$\binom{-5}{n}=\frac{(-5)(-5-1)(-5-2)\ldots(-5-(n-1))}{n!},$$

so we have

$$\binom{-5}{n}(-x)^n = \frac{5(5+1)\dots(5+(n-1))}{n!}x^n = \binom{n+4}{n}x^n,$$

and we have to shift it by 3, so

$$c_n = \binom{n+1}{n-3} = \binom{n+1}{4}.$$

Remark 4.4.1. If we think of putting 4 bars in $n+1$ space, then it can be easily thought that $c_n = \binom{n+1}{4}$.

Note 4.4.1. We can think of there are $n+1$ spaces, and the boundary between the theoretical part and the practical part is $|$, and the midterm of theoretical part is a, while the experiments of practical part are b_1, b_2, and the other symbols are $*$, which are some normal course days, then we can first choose 4 spaces out of $n+1$ spaces, and put on $a, |, b_1, b_2$ in order, then use $*$ to fill all the other spaces. Note that this method corresponds to a way of designing the courses, so $c_n = \binom{n+1}{4}$.

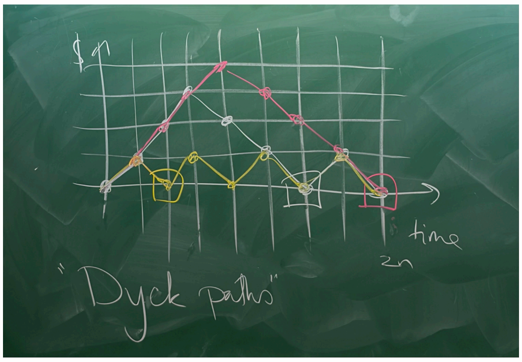
4.5 Catalan Numbers

Example 4.5.1. Bank balance goes up by \$1 or down by \$1, and bank balance should always be ≥ 0.

Question. Start with \$0. How many ways can we have \$0 after $2n$ days?

Answer. Define c_n = number of ways of having \$0 after $2n$ days. Consider the first timewe have \$0. Suppose it happens on Day $2k$, $k \geq 1$, then c_{n-k} ways proceeding from Day $2k$ to Day $2n$, where $c_0 = 1$. For the initial period: Since we know during Day 1 to Day $2k-1$, we've never been to \$0, and in Day 1, we should have $+1$, and in Day $2k$, we must have -1, so in between, we must have at least \$1, so it is Dyck path from Day 1 to Day $2k-1$ with axis $y = 1$. Hence, by the sum rule,

$$c_n = \sum_{k=1}^{n} c_{k-1}c_{n-k} \quad \forall n \geq 1 \text{ with } c_0 = 1.$$



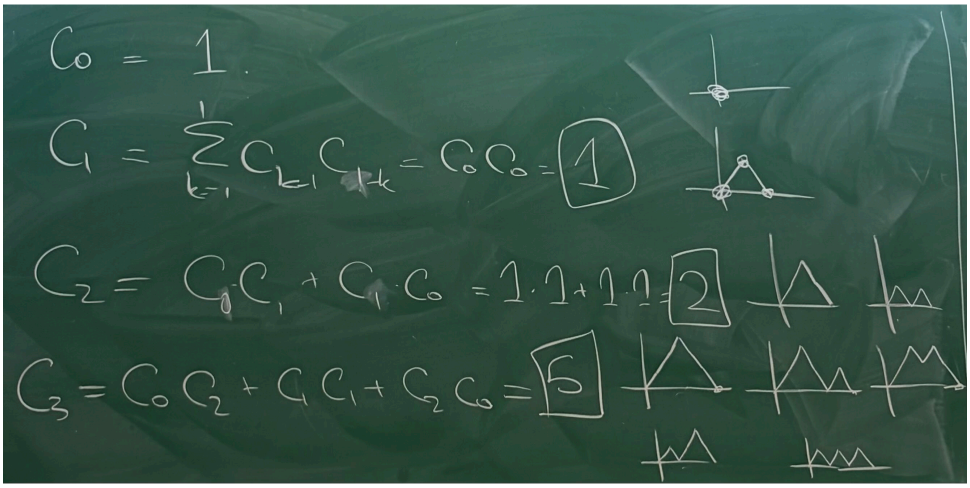


Figure 4.2: Example of c_i.

Lecture 11

14 Oct. 15:30

As previously seen. Suppose

$$A(x) = \sum_{n=0}^{\infty} a_n x^n, \quad B(x) = \sum_{n=0}^{\infty} b_n x^n,$$

and

$$C(x) = A(x)B(x) = \sum_{n=0}^{\infty} c_n x^n,$$

then $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. We call this a convolution.

As previously seen. Catalan numbers

$$c_n = \left| \left\{ (b_1, b_2, \ldots, b_{2n}) \in \{-1,1\}^{2n} : \sum_{i=1}^{2n} b_i = 0, \sum_{j=1}^{i} b_j \geq 0 \quad \forall j \right\} \right|.$$

If we count recursively, based on first returned to 0 (On Day $2k$). Then, from Day $2k$ to Day $2n$, we start at \$0 and end at \$0, and never drop below to \$0, so there are c_{n-k} ways. Also, in the first $2k$ days, it starts at \$0 and ends at \$0. Hence, from Day 1 to Day $2k-1$, it always has at least \$1, and thus there is a bijection, which means there are c_{k-1} probabilities. Therefore, by the product and sum rules,

$$c_n = \sum_{k=1}^{n} c_{k-1} c_{n-k},$$

where $c_0 = 1$.

Define v_n to be the number of very Catalan sequences, which is the sequence of length $2n$ that start with $+1$ and end at \$0, and never drop below \$1 in between. Then,

$$v_n = \begin{cases} 0, & \text{if } n = 0; \\ c_{n-1}, & \text{if } n \geq 1. \end{cases}$$

Then,

$$c_n = \sum_{k=1}^{n} c_{k-1} c_{n-k} = \sum_{k=1}^{n} v_k c_{n-k} = \sum_{k=0}^{n} v_k c_{n-k},$$

so c_n is the convolution of (v_n) and (c_n). Thus, if

$$C(x) = \sum_{n=0}^{\infty} c_n x^n, \quad V(x) = \sum_{n=0}^{n} v_n x^n,$$

then

$$C(x)V(x) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} v_k c_{n-k} x^n = \sum_{n=0}^{\infty} c_n x^n - 1 = C(x) - 1.$$

since

$$\sum_{k=0}^{n} v_k c_{n-k} = \begin{cases} c_0 - 1, & \text{if } n = 0; \\ c_n, & \text{if } n \geq 1. \end{cases}$$

Since

$$v_n = \begin{cases} 0, & \text{if } n = 0; \\ c_{n-1}, & \text{if } n \geq 1, \end{cases}$$

so $V(x) = xC(x)$. Hence,

$$(xC(x))\, C(x) = C(x) - 1 \Rightarrow xC(x)^2 - C(x) + 1 = 0.$$

Hence, we have

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

However, which one is correct? Note that

$$\lim_{x \to 0} C(x) = C(0) = 1,$$

and

$$\lim_{x \to 0} \frac{1 + \sqrt{1-4x}}{2x} = \text{D.N.E.} \quad \lim_{x \to 0} \frac{1 - \sqrt{1-4x}}{2x} = 1.$$

Besides,

$$\sqrt{1-4x} = (1-4x)^{\frac{1}{2}} = \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (-4x)^n = \underbrace{1}_{n=0} + \underbrace{\frac{1}{2}(-4)x}_{n=1} + \underbrace{\left(-\frac{1}{8}\right)(16)x^2}_{n=2} + \ldots.$$

Thus, the coefficient of x^n in $(1-4x)^{\frac{1}{2}}$ is

$$\begin{aligned}\binom{\frac{1}{2}}{n}(-4)^n &= \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\left(\frac{1}{2}-2\right)\dots\left(\frac{1}{2}-(n-1)\right)}{n!}(-4)^n \\ &= \frac{\frac{1}{2}\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\dots\left(-\frac{2n-3}{2}\right)}{n!}(-4)^n \\ &= \frac{-(2n-3)!!}{n!}2^n \\ &= -\frac{(2n-2)!}{n!(n-1)!2^{n-1}}2^n \\ &= (-2)\frac{(2n-2)!}{n!(n-1)!}.\end{aligned}$$

Thus, coefficient of x^n in $C(x)$ $(n \geq 1)$ is

$$-\frac{1}{2} \times \left\{\text{coefficient of } x^{n+1} \text{ in } (1-4x)^{\frac{1}{2}}\right\},$$

which means

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

Question. How many ways are there of starting from \$0 and goinf $+/-$ \$1 each day and ending at \$0 after $2n$ days if we can go below \$0 in between?

Answer. $\binom{2n}{n}$ since we just have to make sure $+\$1$ is as many as $-\$1$. ⊛

4.5.1 k-wise products and compositions of generating functions

Suppose we have sequences

$$a_0^{(j)}, a_1^{(j)}, \dots, a_n^{(j)}, \dots$$

enumerating the number of ways we can carrry out Task j with a budget of \$$n$ with $1 \leq j \leq k$. Let

$$A^{(j)}(x) = \sum_{n=0}^{\infty} a_n^{(j)} x^n$$

be the generating functions. Then, what does

$$A(x) = \prod_{j=1}^{k} A^{(j)}(x)$$

represent?

The answer is: If $A(x) = \sum_{n=0}^{\infty} a_n x^n$, then a_n is the number of ways of taking a budget of \$$n$ splitting it between the k tasks, and carrying out each task with its budget since

$$a_n = \sum_{\substack{l_1, l_2, \dots, l_k \\ l_j \geq 0, \sum_{j=1}^{k} l_j = n}} \prod_{j=1}^{k} a_{l_j}^{(j)}.$$

Composition of generating functions

If $A(x) = \sum_{n=0}^{\infty} a_n x^n$ and $B(x) = \sum_{n=0}^{\infty} b_n x^n$, then what does $B(A(x))$ represent? Suppose

$$B(A(x)) = \sum_{n=0}^{\infty} b_n \left(A(x)\right)^n = \sum_{n=0}^{\infty} c_n x^n,$$

then

$$c_0 = \underbrace{b_0}_{n=0} + \underbrace{b_1 a_0}_{n=1} + \underbrace{b_2 a_0^2}_{n=2} + \cdots = \sum_{n=0}^{\infty} b_n a_0^n.$$

In order for this to be finite, we require $a_0 = 0$. If $a_0 = 0$, then computing any coefficient c_n in $C(x) = B(A(x))$ is a finite computation, so it is a well-defined power series.

Example 4.5.2. $A(x) = \lambda x$ or $A(x) = x^k$.

A special case is that if $B(x) = \frac{1}{1-x}$, then $b_n = 1$ for all n, and

$$B(A(x)) = \sum_{n=0}^{\infty} A(x)^n.$$

Claim 4.5.1. If a_n is the number of ways of carrying out a task with a budget of \$$n$ with $a_0 = 0$, then $\frac{1}{1-A(x)}$ is the generating function for the number of ways of carrying out the task any number of times with a total budeget of \$$n$.

Example 4.5.3. Suppose we have an army with n (identical) soldiers. How many ways can we divide the soldiers into units, and pick a captain of each unit?

Proof. Suppose a_n is the number of ways of picking a captain from a unit of n soldiers, then we know $a_n = n$, so we know $A(x)$ can be obtained by differentiating $\frac{1}{1-x}$ and shifting it to the right by 1 unit, so $A(x) = \frac{x}{(1-x)^2}$. Now if $F(x)$ is the generating function for forming units and then picking a captain for each, then

$$F(x) = \frac{1}{1-A(x)} = \frac{1-2x+x^2}{1-3x+x^2} = 1 + \frac{x}{1-3x+x^2} = 1 + \frac{\alpha}{1-\lambda_1 x} + \frac{\beta}{1-\lambda_2 x},$$

where $(1-\lambda_1 x)(1-\lambda_2 x) = 1-3x+x^2$,and α, β are chosen appropriately. By solving it, we know

$$F(x) = 1 + \frac{\frac{1}{\sqrt{5}}}{1-\left(\frac{\sqrt{5}+3}{2}\right)x} + \frac{-\frac{1}{\sqrt{5}}}{1-\left(\frac{-\sqrt{5}+3}{2}\right)x},$$

so

$$f_n = \frac{1}{\sqrt{5}}\left(\left(\frac{3+\sqrt{5}}{2}\right)^n - \left(\frac{3-\sqrt{5}}{2}\right)^n\right)$$

for all $n \geq 1$ and $f_0 = 1$. ⊛

Lecture 12

17 Oct. 12:20

As previously seen. If $A_i(x)$ is the generating function for the number of ways of completing Task i with a budget of size n, then $\prod_{i=1}^{k} A_i(x)$ is the generating function for the number of ways of taking a budget of size n and

(1) splitting it arbitrary into k parts.

(2) using the i-th part to carry out Task i.

There is one spaecial case $A(x)^k$, which is the number of ways that carrying a task k times with total budget of size n.

Note that

$$B(A(x)) = \sum_{k=0}^{\infty} b_k A(x)^k = \sum_{k=0}^{\infty} b_k \left(\sum_{n=0}^{\infty} a_n x^n\right)^k,$$

and we need $a_0 = 0$ for this to be well-defined. There is a special case: $b_n = 1$, which means $B(x) = \frac{1}{1-x}$, so

$$B(A(x)) = \frac{1}{1 - A(x)} = \sum_{k=0}^{\infty} A(x)^k,$$

where the coefficient of x^n is the number of ways of carrying out a task on arbitrary number of times with a budget of n.

The general case: $B(A(x)) = \sum_{k=0}^{\infty} b_k A(x)^k$, where a_n is the number of ways carrying out a task with a budget of n, while b_k is the number of ways carrying out a second task on a set of k items. Thus, the coefficient of x^n in $B(A(x))$ counts the number of ways to take a budget of n, and

(1) carry out the first task an arbitrary number of times, and

(2) carry out the second task on the outcomes from (1).

Example 4.5.4 (Designing Labubus)**.** We have n minutes to design new Labubus, and then decide which new models to discard, sell, or sell as premium models.

Proof. We have two types of tasks.

(1) $A(x)$: Designing of new models in n minutes.

(2) $B(x)$: Classification of designs as goodbye/okay/premium.

Thus,

$$A(x) = \sum_{n=0}^{\infty} a_n x^n,$$

where a_n is the number of ways of designing a new Labubus in n minutes. Every minute we can change a different feature or not. Hence, there are 2^n options in n minutes but cannot have no changes, so $a_n = 2^n - 1$, where $a_0 = 0$, so it is well-defined. Note that $2^n \leftrightarrow \frac{1}{1-2x}$ and $1 \leftrightarrow \frac{1}{1-x}$, so

$$2^n - 1 \leftrightarrow \frac{1}{1-2x} - \frac{1}{1-x} = A(x) \Rightarrow A(x) = \frac{x}{1 - 3x + 2x^2}.$$

Also, we know $B(x) = \sum_{k=0}^{\infty} b_k x^k$, so $b_k = 3^k$ and thus $B(x) = \frac{1}{1-3x}$, so

$$B(A(x)) = \frac{1}{1 - 3A(x)} = 1 + \frac{3x}{1 - 6x + 2x^2} = 1 + \frac{\alpha}{1 - \lambda_1 x} + \frac{\beta}{1 - \lambda_2 x},$$

we can solve that $\lambda_{1,2} = 3 \pm \sqrt{7}$ since λ_1, λ_2 are the roots of the characteristic polynomial $z^2 - 6z + 2$. By solving α, β, we can get

$$B(A(x)) = 1 + \frac{3}{2\sqrt{7}} \left(\frac{1}{1 - (3 + \sqrt{7})x} - \frac{1}{1 - (3 - \sqrt{7})x} \right),$$

so the coefficient of x^n is

$$f_n := \begin{cases} 1, & \text{if } n = 0; \\ \frac{3}{2\sqrt{7}} \left((3 + \sqrt{7})^n - (3 - \sqrt{7})^n \right), & \text{if } n \geq 1. \end{cases}$$

⊛

Lecture 13

28 Oct. 15:30

As previously seen. If $A_i = \sum_{n=0}^{\infty} a_n^{(i)} x^n$ is the generating function for performing task i with a budget of n, then $A(x) = \prod_{i=1}^{k} A_i(x)$ is the generating function for the number of ways of completing tasks $1, 2, \dots, k$ with total budget of n.

One of the application is to calculate $P(x) = \sum_{n=0}^{\infty} p(n)x^n$, where $p(n)$ is the number of ways of writing n as an unordered sum of natural numbers.

- Task 1: Write a number n as a sum of 1s, and define $a_n^{(1)}$ is the number of ways of doint this, then $A_1(x) = \frac{1}{1-x}$ since the generating function of $a_n^{(1)}$ is the genrating function of $(1, 1, 1, \dots)$.

- Task 2: Write a number n as a sum of 2s, and define $a_n^{(2)}$ to be the number of ways of doing this, then we know
$$a_n^{(2)} = \begin{cases} 1, & \text{if } 2 \mid n; \\ 0, & \text{if } 2 \nmid n. \end{cases},$$
so $\left(a_1^{(2)}, a_2^{(2)}, \dots\right) = (1, 0, 1, 0, 1, 0, 1, 0, 1, \dots)$, so $A_2(x) = \frac{1}{1-x^2}$.

- Task i, and similarly we can define $a_n^{(i)}$ and we know
$$a_n^{(i)} = \begin{cases} 1, & \text{if } i \mid n; \\ 0, & \text{if } i \nmid n. \end{cases},$$
and we know $A_i(x) = \frac{1}{1-x^i}$.

Hence, we know
$$P(x) = \sum_{n=0}^{\infty} p(n)x^n = \prod_{i=1}^{\infty} A_i(x) = \prod_{i=1}^{\infty} \frac{1}{1-x^i} = \prod_{i=1}^{\infty} \left(1 + x^i + x^{2i} + x^{3i} + \dots\right).$$

Example 4.5.5. How to compute $p(n, k)$?

Proof. Since we know
$$\begin{aligned} p(n,k) &= \#\text{ of partitions of } n \text{ into } k \text{ non-empty parts} \\ &= \#\text{ of partitions of } n \text{ with largest part of size } k, \end{aligned}$$
so we have
$$P_k(x) = \sum_{n=0}^{\infty} p(n,k)x^n = \left(1 + x^{1\cdot 1} + x^{2\cdot 1} + \dots\right)\left(1 + x^{1\cdot 2} + x^{2\cdot 2} + \dots\right)\dots\left(x^{1\cdot k} + x^{2\cdot k} + \dots\right),$$
so $P_k(x) = x^k \prod_{i=1}^{k} \frac{1}{1-x^i}$. ⊛

Example 4.5.6. Suppose $o(n)$ is the number of partitions of n into odd parts, then
$$O(x) = \sum_{n=0}^{\infty} o(n)x^n = \prod_{i=1}^{\infty} \frac{1}{1-x^{2i-1}}.$$

Example 4.5.7. Suppose $d(n)$ is the number of partitions of n into distinct parts, then
$$D(x) = \sum_{n=0}^{\infty} d(n)x^n = \prod_{i=1}^{\infty} \left(1 + x^i\right).$$

Example 4.5.8. Suppose $q(n)$ is the number of self-conjugate partition, then we know $q(n)$ is equal to number of partitions of n into distinct odd parts, so

$$Q(x) = \sum_{n=0}^{\infty} q(n)x^n = \prod_{n=1}^{\infty} \left(1 + x^{2i-1}\right).$$

Theorem 4.5.1. For all n, the number of partitions of n into odd parts is equal to the number of partitions of n into distinct parts i.e. $o(n) = d(n)$.

Proof. Since we know

$$\begin{aligned} O(x) = \sum_{n=0}^{\infty} o(n)x^n &= \prod_{i=1}^{\infty} \frac{1}{1 - x^{2i-1}} \\ &= \frac{\prod_{i=1}^{\infty} \left(1 - x^{2i}\right)}{\prod_{i=1}^{\infty} \left(1 - x^{i}\right)} = \prod_{i=1}^{\infty} \frac{\left(1 - x^{2i}\right)}{\left(1 - x^{i}\right)} = \prod_{i=1}^{\infty} \left(1 + x^{i}\right) \\ &= D(x) = \sum_{n=0}^{\infty} d(n)x^n, \end{aligned}$$

so this is true. ∎

4.6 Exponential Generating Functions

Generally, generating functions work well when sequences are at most exponential, since if $|a_n| \le c\lambda^n$ for all n, then $A(x) = \sum_{n=0}^{\infty} a_n x^n$ converges absolutely for $|x| < \frac{1}{\lambda}$. However, combinatorial sequences often grows faster, especially when order is included.

Example 4.6.1. $n! \ge \left(\frac{n}{e}\right)^n$, so $A(x) = \sum_{n=0}^{\infty} n! x^n$ only converges for $x = 0$.

Definition 4.6.1. Given a sequence $(a_n)_{n\ge 0}$, its exponential generating function is the (formal) power series $\hat{A}(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$.

> **Note 4.6.1.** If $|a_n| \le C\lambda^n n!$, then $\hat{A}(x)$ converges absolutely for $|x| < \frac{1}{\lambda}$.

Example 4.6.2. If $a_n = 1$ for all $n \in \mathbb{N} \cup \{0\}$, then we know the exponential generating function of $(a_n)_{n\ge 0}$ is

$$\hat{A}(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x.$$

If $(a_n)_{n\ge 0} \leftrightarrow \hat{A}(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$ and $(b_n)_{n\ge 0} \leftrightarrow \hat{B}(x) = \sum_{n=0}^{\infty} b_n \frac{x^n}{n!}$, then we have

$$(\lambda a_n + \mu b_n)_{n\ge 0} \leftrightarrow \sum_{n=0}^{\infty} (\lambda a_n + \mu b_n) \frac{x^n}{n!} = \lambda\hat{A}(x) + \mu\hat{B}(x).$$

Also, if we multiplying by x, then

$$x\hat{A}(x) = \sum_{n=0}^{\infty} a_n \frac{x^{n+1}}{n!} = \sum_{n=1}^{\infty} a_{n-1} \frac{x^n}{(n-1)!} = \sum_{n=1}^{\infty} n a_{n-1} \frac{x^n}{n!} \leftrightarrow (n a_{n-1})_{n\ge 0}.$$

If we divide by x, then we have

$$\frac{\hat{A}(x) - a_0}{x} = \sum_{n=1}^{\infty} a_n \frac{x^{n-1}}{n!} = \sum_{n=0}^{\infty} a_{n+1} \frac{x^n}{(n+1)!} = \sum_{n=0}^{\infty} \frac{a_{n+1}}{n+1} \frac{x^n}{n!} \leftrightarrow \left(\frac{a_{n+1}}{n+1}\right)_{n\ge 0}.$$

Example 4.6.3. Consider the sequence given by $a_0 = 1$, and $\forall n \geq 1$, we have

$$a_{n+1} = (n+1)a_n - n^2 + 1.$$

Let $\hat{A}(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$, then we know

$$\begin{aligned}
\hat{A}(x) &= a_0 + \sum_{n=0}^{\infty} a_{n+1} \frac{x^{n+1}}{(n+1)!} \\
&= 1 + \sum_{n=0}^{\infty} \left((n+1)a_n - n^2 + 1\right) \frac{x^{n+1}}{(n+1)!} \\
&= 1 + \sum_{n=0}^{\infty} a_n \frac{x^{n+1}}{n!} - \sum_{n=0}^{\infty} \left(n^2 - 1\right) \frac{x^{n+1}}{(n+1)!} \\
&= 1 + x\hat{A}(x) - \sum_{n=0}^{\infty} (n-1) \frac{x^{n+1}}{n!} \\
&= 1 + x\hat{A}(x) - \sum_{n=1}^{\infty} n \frac{x^{n+1}}{n!} + \sum_{n=0}^{\infty} \frac{x^{n+1}}{n!} \\
&= 1 + x\hat{A}(x) - x^2 e^x + x e^x,
\end{aligned}$$

so we have

$$\hat{A}(x) = \frac{1}{1-x} + x e^x = \sum_{n=0}^{\infty} (n! + n) \frac{x^n}{n!} = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!},$$

so $a_n = n! + n$.

Products of exponential generating functions

Suppose $(a_n)_{n\geq 0} \leftrightarrow \hat{A}(x)$ and $(b_n)_{n\geq 0} \leftrightarrow \hat{B}(x)$, then

Question. How do we interpret $\hat{A}(x)\hat{B}(x)$ as an exponential generating function?

Note that

$$\hat{A}(x)\hat{B}(x) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \frac{a_k}{k!} \frac{b_{n-k}}{(n-k)!} \right) x^n = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \frac{n!}{k!(n-k)!} a_k b_{n-k} \right) \frac{x^n}{n!} = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k} \right) \frac{x^n}{n!},$$

so if we define $\hat{C}(x) = \hat{A}(x)\hat{B}(x)$ to be the exponential generating function for the sequences $(c_n)_{n\geq 0}$, then

$$c_n = \sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k}.$$

Its combinatorial interpretation: Suppose a_n is the number of ways of doint Task 1 on a set of n elements, and b_n is the number of ways of doint Task 2 on a set of n elements, then $c_n = \sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k}$ is the number of ways to take a set $[n]$ of n elements and

(1) Choose a partition $[n] = A \cup B$.

(2) Perform Task 1 on A.

(3) Perform Task 2 on B.

Example 4.6.4. A professor and TA are grading n exams. We should divide exams arbitrarily: professor grades his and TA grades hers. Also the professor orders the exams to grade one at a time and decide if a student passes or not, while TA orders her exams and passes everyone.

Proof. Suppose a_n is the number of ways professor can grade a set of n exams, then $a_n = n!2^n$, and thus $\hat{A}(x) = \sum_{n=0}^{\infty} n!2^n \frac{x^n}{n!} = \sum_{n=0}^{\infty} (2x)^n = \frac{1}{1-2x}$. Also, suppose b_n is the number of ways TA can grade a set of n exams, then $b_n = n!$, so we know

$$\hat{B}(x) = \sum_{n=0}^{\infty} b_n \frac{x^n}{n!} = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x},$$

so if c_n is the number of ways the professor and TA can grade a set of n exams together, then

$$\hat{C}(x) = \sum_{n=0}^{\infty} c_n x^n = \hat{A}(x)\hat{B}(x) = \frac{1}{(1-2x)(1-x)} = \frac{2}{1-2x} + \frac{-1}{1-x} = \sum_{n=0}^{\infty} \left(2^{n+1} - 1\right) x^n,$$

so $c_n = \left(2^{n+1} - 1\right) n!$. ⊛

Lecture 14

Continuing the last example, here we want to give a combinatorial proof. We can first order the n exams, and choose the first k to give to the professor to grade, so there are $n!2^k$ options, so there are 31 Oct. 12:20

$$n! \sum_{k=0}^{n} 2^k = n!(2^{n+1} - 1)$$

ways of grading.

There is another combinatorial interpretation: Suppose we permutate those n exams, then there are 2^n subsets of these n exams for a fixed permutation, and for every permutation, every subset corresponds to a method that will be graded by professor and has same result(pass/fail) as the biggest element of this subset, so there are $2^{n+1} - 1$ ways of choosing the result given by professor since for every subset, there are 2 possible results(pass/fail), but if we pick empty set, then choose pass or fail does not matter since it means the professor does not grade any exam and thus only has one possibility, so the answer is $n!\left(2^{n+1} - 1\right)$.

Differentiation

Now if we differentiate $\hat{A}(x)$, then we know

$$\frac{\mathrm{d}}{\mathrm{d}x}\hat{A}(x) = \sum_{n=1}^{\infty} a_n \frac{nx^{n-1}}{n!} = \sum_{n=1}^{\infty} a_n \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} a_{n+1} \frac{x^n}{n!} \leftrightarrow (a_{n+1})_{n\geq 0}.$$

Thus, this shift the sequence to the left.

Integration

If we integrate $\hat{A}(x)$, then we have

$$\int_0^x \hat{A}(t)\,\mathrm{d}t = \sum_{n=0}^{\infty} \int_0^x a_n \frac{t^n}{n!}\,\mathrm{d}t = \sum_{n=0}^{\infty} a_n \frac{x^{n+1}}{(n+1)\cdot n!} = \sum_{n=0}^{\infty} a_n \frac{x^{n+1}}{(n+1)!} = \sum_{n=1}^{\infty} a_{n-1} \frac{x^n}{n!} \leftrightarrow (0, a_0, \dots),$$

which is equivalent to shift the sequence to the right.

Bell numbers

Example 4.6.5 (Bell numbers)**.** There are n different pieces of candy. An unspecified number of children will ring your bell, take a subset of the candy until you run out, then how many ways can the candy be distributed?

Proof. Mathematically, we are asking how many ways can a set $[n]$ be partitioned into (indistinguishable) nonempty subsets? Thus, the answer is

$$\sum_{k=0}^{n} S(n,k).$$

Note 4.6.2. The problem statements may be a little confusing, so just follow the mathematical version of the problem.

⊛

Definition 4.6.2 (Bell numbers)**.** We define the bell number to be

$$B(n) := \sum_{k=0}^{n} S(n,k),$$

which is the number of ways to partition $[n]$ into non-empty subsets.

Example 4.6.6. $B(0)=1, B(1)=1, B(2)=2, B(3)=5$.

Theorem 4.6.1. For all $n \geq 0$, we have

$$B(n+1) = \sum_{k=0}^{n} \binom{n}{k} B(k).$$

Proof. We can analyze which subset should $n+1$ belong to, so we can count on the number of other elements that are in the same subset as $n+1$. If the number of elements that are in the same subset as $n+1$ is k, then there are

$$\binom{n}{k} B(n-k) = \binom{n}{n-k} B(n-k)$$

choices, so there are totally

$$\sum_{k=0}^{n} \binom{n}{n-k} B(n-k) = \sum_{k=0}^{n} \binom{n}{k} B(k)$$

ways of partitioning $[n+1]$ into non-empty subsets. ∎

Now let $\hat{B}(x)$ be the exponential function for $B(n)$, then

$$\hat{B}(x) = \sum_{n=0}^{\infty} B(n) \frac{x^n}{n!}.$$

Define

$$\hat{C}(x) = \sum_{n=0}^{\infty} B(n+1) \frac{x^n}{n!} = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \binom{n}{k} B(k) \right) \frac{x^n}{n!} = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \binom{n}{k} B(k) \cdot 1 \right) \frac{x^n}{n!},$$

then note that this is the convolution:

$$\left[\sum_{n=0}^{\infty} B(n) \frac{x^n}{n!} \right] \left[\sum_{n=0}^{\infty} 1 \cdot \frac{x^n}{n!} \right] = \hat{B}(x) e^x.$$

Note that $\hat{C}(x) = \frac{\mathrm{d}}{\mathrm{d}x} \hat{B}(x)$, so if we differentiate it, we have

$$\frac{\mathrm{d}}{\mathrm{d}x} \hat{B}(x) = \hat{B}(x) \cdot e^x.$$

Hence,
$$\frac{\frac{\mathrm{d}}{\mathrm{d}x}\hat{B}(x)}{\hat{B}(x)} = e^x \Leftrightarrow \frac{\mathrm{d}}{\mathrm{d}x}\left[\ln \hat{B}(x)\right] = e^x,$$
so we have
$$\hat{B}(x) = e^{e^x - 1}.$$

Lecture 15

Multiple products

4 Nov. 15:30

Let $a_n^{(i)}$ be the number of ways of carrying out Task i on a set of size n, and suppose
$$\hat{A}^{(i)}(x) = \sum_{n=0}^{\infty} a_n^{(i)} \frac{x^n}{n!},$$
then for $\hat{A}(x) = \prod_{i=1}^{k} \hat{A}^{(i)}(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$, we know a_n counts the number of ways of partitioning $[n] = S_1 \uplus S_2 \uplus \ldots \uplus S_k$ and carrying out Task i on S_i for each $i \in [k]$. Equivalently,
$$a_n = \sum_{l_1, l_2, \ldots, l_k \ l_i \geq 0, \sum l_i = n} \binom{n}{l_1, l_2, \ldots, l_k} a_{l_1}^{(1)} a_{l_2}^{(2)} \ldots a_{l_k}^{(k)}, \text{ where } \binom{n}{l_1, l_2, \ldots, l_k} = \frac{n!}{(l_1)!(l_2)!\ldots(l_k)!}.$$

Exercise 4.6.1. Let a_n counts the number of ways of eating n xiaolongbow. Let
$$\hat{A}(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}.$$
Suppose we have k friends who want to share n xiaolongbows. What is the exponential generating function of this?

Answer. $\left[\hat{A}(x)\right]^k$. ⊛

Exercise 4.6.2. Suppose we have k strangers, what then?

Composition of exponential generating function

Question. What does $\hat{B}\left(\hat{A}(x)\right)$ represent?

$$\hat{B}\left(\hat{A}(x)\right) = \sum_{k=0}^{\infty} b_k \frac{\left(\hat{A}(x)\right)^k}{k!} = \sum_{k=0}^{\infty} b_k \frac{\left(\sum_{n=0}^{\infty} a_n \frac{x^n}{n!}\right)^k}{k!}.$$

Note that if $a_0 \neq 0$, then it is not well-defined since every k term contributes to the constant term and thus there is an infinite sum for coefficient. If $a_0 = 0$, then contribution to x^n always comes from $0 \leq k \leq n$, which is a finite computation, and thus well-defined.

For special case that $b_k \equiv 1 \leftrightarrow \hat{B}(x) = e^x$, then
$$\hat{B}(\hat{A}(x)) = e^{\hat{A}(x)} = \sum_{k=0}^{\infty} \left[\frac{\left(\hat{A}(x)\right)^k}{k!}\right].$$

As argued previously, $\left[\frac{\left(\hat{A}(x)\right)^k}{k!}\right]$ enumerates the number of ways of partitioning $[n]$ into k non-empty subsets and performing the same task on each. If we summing over k, then it means the partitions of $[n]$ into an arbitrary number of nonempty subsets.

Remark 4.6.1. Since the k tasks here are all $\hat{A}(x)$, which are all same, so the order of the subsets do not matters, so we have to divide $k!$ to calculate the number of ways of partitioning $[n]$ into k subsets and then perform this task on each subset.

Proposition 4.6.1. If $\hat{A}(x)$ is the exponential generating function for performing a task on a set of size n, with $a_0 = 0$, then $e^{\hat{A}(x)}$ is the exponential generating function for the number of ways of partitioning a set of size n into arbitrarily many non-empty subsets, and performing the task on each.

Example 4.6.7. If the task is making a non-empty set containing all n elements, then

$$a_n = \begin{cases} 0, & \text{if } n = 0; \\ 1, & \text{if } n \geq 1. \end{cases}$$

Hence, we have $\hat{A}(x) = e^x - 1$, and thus $e^{\hat{A}(x)} = e^{e^x - 1}$. This is because $e^{\hat{A}(x)}$ is the exponential generating function of number of ways of taking $[n]$ and partitioning it into arbitralily many non-empty subsets.

Remark 4.6.2. e^{e^x} is not a sensible generating function since if you expand it, then its x^0 term is

$$\sum_{k=0}^{\infty} \frac{1}{k!},$$

which is an infinite sum.

Example 4.6.8. There are n people coming to dinner and we want to seat them at arbitralily many round table. How many ways can we arrange the people around the tables?

Proof. What matters is who sits at each table, and the circular under at the table, so we know

$$\sum_{k=0}^{n} s_{n,k} = n!$$

is the answer. Or, we can argue that there is a bijection between S_n and seating arrangement, since cycles in a permutation can corresponds to tables. Or, for the generating function approach, we can suppose the task is to sit people around a single non-empty table, so $a_n = \frac{n!}{n} = (n-1)!$. Hence,

$$\hat{A}(x) = \sum_{n=1}^{\infty} a_n \frac{x^n}{n!} = \sum_{n=1}^{\infty} (n-1)! \frac{x^n}{n!} = \sum_{n=1}^{\infty} \frac{x^n}{n} = \ln\left(\frac{1}{1-x}\right).$$

Hence, we know $e^{\hat{A}(x)}$ is the exponential generating function of partitioning n people into arbitrarily many tables and sitting each table, which is what we want, and

$$e^{\hat{A}(x)} = e^{\ln\left(\frac{1}{1-x}\right)} = \frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = \sum_{n=0}^{\infty} n! \frac{x^n}{n!},$$

so the number is $n!$. ⊛

Example 4.6.9. How many ways can we partition $[n]$ into subsets of size $3, 4$, and 7?

Proof. We can let the individual task to be the number of ways of take n elements and put them into a single subset of size $3, 4, 7$, so we know

$$a_n = \begin{cases} 1, & \text{if } n = 3, 4, \text{ and } 7; \\ 0, & \text{otherwise.} \end{cases}$$

Note that

$$\hat{A}(x) = \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^7}{7!}.$$

Now we can repeat an arbitrary number of times to partition all elements, so the exponential generating function for this problem is

$$e^{\hat{A}(x)} = e^{\frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^7}{7!}} = e^{\frac{x^3}{3!}} e^{\frac{x^4}{4!}} e^{\frac{x^7}{7!}}.$$

⊛

General Composition

Suppose $a_0 = 0$, then

$$\hat{B}\left(\hat{A}(x)\right) = \sum_{k=0}^{\infty} b_k \frac{\left(\hat{A}(x)\right)^k}{k!},$$

and we know $\frac{\hat{A}(x)^k}{k!}$ is the number of ways of splitting n elements into k non-empty subsets and performing Task 1 on each, and b_k is the number of performing Task 2 on a set of k subsets. Hence, the coefficient of x^n of $\hat{B}\left(\hat{A}(x)\right)$ means the number of ways of taking a set of size n and

(1) Partitioning into arbitrarily many non-empty subsets.

(2) Performing Task 1 ($\hat{A}$) on each subset.

(3) Performing Task 2 ($\hat{B}$) on the set of subsets.

Example 4.6.10. Seating chart for a wedding, and we want each table has even number of people. Hence, we need to decide

(1) Who sits together at a table?

(2) How do they sit around the table? (Permutation, not just cyclic order, and we want even number of people)

(3) Which table do they sit at?

Proof. The first task is to seat n people around a non-empty table such that there are even number of people:

$$a_n = \begin{cases} 0, & \text{if } 2 \nmid n; \\ n!, & \text{if } 2 \mid n. \end{cases}$$

Hence,

$$\hat{A}(x) = \sum_{n=1}^{\infty} a_n \frac{x^n}{n!} = \sum_{2|n, n\geq 2} n! \frac{x^n}{n!} = \sum_{2|n, n\geq 2} x^n = \sum_{k=1}^{\infty} x^{2k} = \frac{x^2}{1-x^2}.$$

The second task is to assign groups to tables, so $b_k = k!$, so

$$\hat{B}(x) = \sum_{k=0}^{\infty} k! \frac{x^k}{k!} = \sum_{k=0}^{\infty} x^k = \frac{1}{1-x}.$$

Hence, the exponential generating function is

$$\begin{aligned}
\hat{C}(x) = \hat{B}\left(\hat{A}(x)\right) &= \frac{1}{1-\hat{A}(x)} \\
&= \frac{1}{1-\frac{x^2}{1-x^2}} = \frac{1-x^2}{1-2x^2} = 1 + \frac{x^2}{1-2x^2} \\
&= 1 + x^2 \sum_{k=0}^{\infty} \left(2x^2\right)^k = 1 + x^2 \sum_{k=0}^{\infty} 2^k x^{2k} \& = 1 + \sum_{k=0}^{\infty} 2^k x^{2k+2} = 1 + \sum_{2|n, n\geq 2} 2^{\frac{n}{2}-1} x^n \\
&= 1 + \sum_{2|n, n\geq 0} 2^{\frac{n}{2}-1} n! \frac{x^n}{n!},
\end{aligned}$$

so we know

$$c_n = \begin{cases} 0, & \text{if } n \text{ odd;} \\ 1, & \text{if } n = 0; \\ 2^{\frac{n}{2}-1} n!, & \text{if } n \text{ even and } n \geq 2. \end{cases}$$

⊛

Chapter 5

Asymptotics

Example 5.0.1.

(1) Number of ways of partitioning \$$3n$ between 5 people.

(2) Number of ways of distributing n distinct artworks between 5 people.

Answer.

(1) $\binom{3n+4}{5}$.

(2) 5^n.

⊛

Lecture 16

Now we focus on the growth of a sequence when $n \to \infty$. 7 Nov. 12:20

Example 5.0.2.

$$a_n = 0.01n \left(\frac{3+\sqrt{5}}{2}\right)^n + 10^{100} \left(\frac{3+\sqrt{5}}{2}\right)^n + 10 \cdot 2^n + 1000 \left(\frac{3-\sqrt{5}}{2}\right)^n,$$

the first term grows fastest.

Example 5.0.3. How many arithmetic operations do we need to multiply two $n \times n$ matrices?

Answer. $(2n-1)n^2 = 2n^3 - n^2$. ⊛

However, actually we have better algorithm! It takes only $Cn^{2.371339}$ times of arithmetic operations (2024).

Definition 5.0.1. Given function $f, g : \mathbb{N} \to \mathbb{R}$, we say $f = O(g)$ if $\exists C \in \mathbb{R}$ and $n_0 \in \mathbb{N}$ s.t. $|f(n)| \leq C\,|g(n)|$ for all $n \geq n_0$.

Example 5.0.4. For any $k \in \mathbb{N}$, $\lambda > 1$, $n^k = O\left(\lambda^n\right)$.

Definition 5.0.2. We say $f = \Omega(g)$ iff there exists $c > 0$ and $n_0 \in \mathbb{N}$ s.t. $f(n) \geq c\,|g(n)|$ for all $n \geq n_0$.

Corollary 5.0.1. $f = O(g)$ iff $g = \Omega(f)$ for $f, g \geq 0$.

Definition 5.0.3. $f = o(g)$ iff

$$\lim_{n\to\infty} \frac{|f(n)|}{|g(n)|} = 0.$$

Definition 5.0.4. $f = \omega(g)$ iff

$$\lim_{n\to\infty} \frac{f(n)}{|g(n)|} = +\infty.$$

Definition 5.0.5. $f = o(g)$ iff $g = \omega(f)$ if $f, g > 0$.

Alternatively, we say $f \ll g$ iff $f = o(g)$ and $f \gg g$ iff $f = \omega(g)$, but sometimes $f \ll g$ means $f = O(g)$ and $f \gg g$ iff $f = \Omega(g)$. Also, we say $f \approx g$ iff

$$f(n) = (1 + o(1))g(n) \Leftrightarrow \lim_{n\to\infty} \frac{f(n)}{g(n)} = 1.$$

Example 5.0.5 (Again)**.**

$$a_n = 0.01n \left(\frac{3+\sqrt{5}}{2}\right)^n + 10^{100} \left(\frac{3+\sqrt{5}}{2}\right)^n + 10 \cdot 2^n + 1000 \left(\frac{3-\sqrt{5}}{2}\right)^n,$$

then $a_n = O\left(n\left(\frac{3+\sqrt{5}}{2}\right)^n\right)$, which means

$$a_n \approx 0.01n \left(\frac{3+\sqrt{5}}{2}\right)^n = (0.01 + o(1))n \left(\frac{3+\sqrt{5}}{2}\right)^n.$$

Also, we know

$$a_n = 0.01n \left(\frac{3+\sqrt{5}}{2}\right)^n + O\left(\left(\frac{3+\sqrt{5}}{2}\right)^n\right).$$

Asymptotic arithmetic

If $f_1 = O(g_1)$ and $f_2 = O(g_2)$, then $f_1 + f_2 = O(g_1 + g_2)$. If $g_2 = o(g_1)$, then in addition, $f_1 + f_2 = O(g_1)$.

If $f_1 = O(g_1)$ and $f_2 = O(g_2)$, then $f_1 f_2 = O(g_1 g_2)$.

Example 5.0.6.

$$(101n^2 - 57n + 90)(n^2 - 55n + 101) = O(n^2)O(n^2) = O(n^4) = 101n^4 + O(n^3).$$

Example 5.0.7. What is $\sum_{i=1}^n i^3$?

Proof. For upper bound, we know

$$\sum_{i=1}^n i^3 \leq \sum_{i=1}^n n^3 = n^4,$$

so $\sum_{i=1}^n i^3 = O(n^4)$.

For lower bound, we know

$$\sum_{i=1}^{n} i^3 \geq \sum_{i=\frac{n}{2}}^{n} i^3 \geq \sum_{i=\frac{n}{2}}^{n} \left(\frac{n}{2}\right)^3 \geq \left(\frac{n}{2}\right)^4 = \Omega\left(n^4\right).$$

Hence, we know $\sum_{i=1}^{n} i^3 = \Theta\left(n^4\right)$. ⊛

Definition 5.0.6. $f = \Theta(g)$ iff $f = O(g)$ and $f = \Omega(g)$.

Example 5.0.8.

$$f(n) = \begin{cases} 100 \cdot 2^n, & \text{if } 2 \mid n; \\ \frac{1}{100} 2^n, & \text{if } 2 \nmid n, \end{cases}$$

then $f(n) = \Theta\left(2^n\right)$.

Lecture 17

11 Nov. 15:30

Theorem 5.0.1.

$$\sum_{i=1}^{n} \binom{i}{3} = \binom{n+1}{4}.$$

Proof. By induction on n.

- Base case: $n \leq 3$ is trivial.
- Induction step:

$$\sum_{i=1}^{n} \binom{i}{3} = \sum_{i=1}^{n-1} \binom{i}{3} + \binom{n}{3} = \binom{n}{4} + \binom{n}{3} = \binom{n+1}{4}.$$

■

Now since

$$\binom{n}{3} = \frac{i(i-1)(i-2)}{6} = \frac{i^3}{6} + O\left(i^2\right),$$

so we have

$$\begin{aligned}\sum_{i=1}^{n} i^3 &= \sum_{i=1}^{n} 6\binom{i}{3} + \sum_{i=1}^{n} \left[i^3 - 6\binom{n}{3}\right] = 6\binom{n+1}{4} + \sum_{i=1}^{n} O\left(i^2\right) \\ &= 6\binom{n+1}{4} + O\left(n^3\right) = \frac{n^4}{4} + O\left(n^3\right).\end{aligned}$$

Suppose

$$\sum_{i=1}^{n} i^3 = \left(C + O(1)\right) n^4$$

for some constant C, then we could want

$$\left(C + O(1)\right)(n+1)^4 = \sum_{i=1}^{n+1} i^3 = \sum_{i=1}^{n} i^3 + (n+1)^3 = (C + O(1))n^4 + n^3 + O\left(n^2\right).$$

5.1 Factorial

> **Question.** How quickly does the factorial $n!$ goes?

We know

$$n! = 1 \times 2 \times 3 \times \cdots \times n \begin{cases} \leq n \times n \times \cdots \times n = n^n \\ \geq 1 \times 1 \times \cdots \times 1 = 1 \\ \geq 2 \times 2 \times \cdots \times 2 = 2^{n-1}. \end{cases}$$

On the other hand, we have

$$n! = 1 \times 2 \times \cdots \times \frac{n}{2} \times \left(\frac{n}{2}+1\right) \times \cdots \times n$$

$$\begin{cases} \leq \left(\frac{n}{2}\right)^{\frac{n}{2}} \times n^{\frac{n}{2}} = \frac{n^n}{2^{\frac{n}{2}}} = \left(\frac{n}{\sqrt{2}}\right)^n \\ \geq 2^{\frac{n}{2}-1} \times \left(\frac{n}{2}\right)^{\frac{n}{2}} = \frac{n^{\frac{n}{2}}}{2} = \frac{1}{2}\left(\sqrt{n}\right)^n. \end{cases}$$

Also, we know

$$\ln(n!) = \sum_{i=1}^{n} \ln i = S,$$

where

$$S \leq \int_1^{n+1} \ln t \, \mathrm{d}t = [t \ln t - t]_{t=1}^{n+1} = (n+1)\ln(n+1) - n,$$

Figure 5.1: The inequality comes from the area between the curve and the x-axis

so we know

$$n! \leq \frac{(n+1)^{n+1}}{e^n},$$

and

$$(n+1)^{n+1} = (n+1)(n+1)^n = (n+1)n^n\left(\frac{n+1}{n}\right)^n = (n+1)n^n\left(1+\frac{1}{n}\right)^n \leq (n+1)n^n e,$$

so we have

$$n! \leq e(n+1)\left(\frac{n}{e}\right)^n.$$

On the other hand, we have

$$\ln(n!) \geq \int_1^t \ln t \, \mathrm{d}t = [t \ln t - t]_1^n = n \ln n - n + 1,$$

so we have
$$n! \geq \left(\frac{n}{e}\right)^n e.$$

Hence, we have
$$e\left(\frac{n}{e}\right)^n \leq n! \leq (n+1)e\left(\frac{n}{e}\right)^n.$$

Stirling's Approximation
$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

5.2 Binomial coefficients

$$\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k!} \leq \frac{n^k}{k!}$$
$$\binom{n}{k} = \frac{n}{k}\cdot\frac{n-1}{k-1}\dots\frac{n-k+1}{1} \geq \left(\frac{n}{k}\right)^k.$$

Hence, we know for all $n \geq k \geq 0$,
$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \frac{n^k}{k!}.$$

Observe that if $k = k(n)$ is constant as $n \to \infty$, then the upper bound is asymptotically tight, i.e.
$$\binom{n}{k} \approx \frac{n^k}{k!} \text{ for } k = O(1).$$

Rigorously, this is because
$$\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k!} = \frac{n\cdot n\left(1-\frac{1}{n}\right)\cdot\dots\cdot n\left(1-\frac{k-1}{n}\right)}{k!} = \frac{n^k}{k!}\left(1 - O\left(\frac{k^2}{n}\right)\right).$$

If $k = \omega(1)$, then
$$\binom{n}{k} \leq \frac{n^k}{k!} \leq \frac{n^k}{\left(\frac{k}{e}\right)^k} = \left(\frac{ne}{k}\right)^k,$$

so we have
$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k.$$

If $k = o(n)$, i.e. $\frac{n}{k} \to \infty$, then $\left(\frac{ne}{k}\right)^k$ is a good approximation. If $k = \Theta(n)$, then the approximation is not so good.

We have
$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$

and if $n, k \to \infty$, we may assume $n - k \to \infty$, otherwise use $\binom{n}{k} = \binom{n}{n-k} \approx \frac{n^{n-k}}{(n-k)!}$. By Stirling's approximation,
$$\begin{aligned}
\binom{n}{k} &\approx \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{\sqrt{2\pi k}\left(\frac{k}{e}\right)^k\sqrt{2\pi(n-k)}\left(\frac{n-k}{e}\right)^{n-k}} \\
&= \sqrt{\frac{n}{2\pi k(n-k)}}\frac{n^n}{k^k(n-k)^{n-k}} = \sqrt{\frac{n}{2\pi k(n-k)}}\cdot\left(\frac{n}{k}\right)^k\cdot\left(\frac{n}{n-k}\right)^{n-k}.
\end{aligned}$$

Hence,
$$\begin{aligned}
\log_2\binom{n}{k} &\approx \log_2\left(\sqrt{\frac{n}{2\pi k(n-k)}}\right) + k\log_2\left(\frac{n}{k}\right) + (n-k)\log_2\left(\frac{n}{n-k}\right) \\
&= \left[-\frac{k}{n}\log_2\left(\frac{k}{n}\right) - \left(1-\frac{k}{n}\right)\log_2\left(1-\frac{k}{n}\right)\right]n + o(n) \\
&= H\left(\frac{k}{n}\right)n + o(n),
\end{aligned}$$

where

$$H(x) = -x \log_2 x - (1-x)\log_2(1-x)$$

is the binary entropy function and thus

$$\binom{n}{k} = 2^{H\left(\frac{k}{n}\right)n + o(n)}.$$

Hence, we know

$$\binom{n}{k} \begin{cases} \approx \dfrac{n^k}{k!} \text{ for } k = O(1) \\ \leq \left(\dfrac{ne}{k}\right)^k \text{ for } \omega(1) = k = o(n) \\ = 2^{H\left(\frac{k}{n}\right)n+o(n)} \text{ for } k = \Theta(n). \end{cases}$$

> **Question.** How big is the largest binomial coefficient?

Note that

$$\binom{n}{k} \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$$

since

$$\binom{n}{k+1} = \frac{n!}{(k+1)!(n-k-1)!} = \frac{n-k}{k+1}\frac{n!}{k!(n-k)!} = \frac{n-k}{k+1}\binom{n}{k},$$

and $\frac{n-k}{k+1} \leq 1$ if $k > \frac{n}{2}$ and ≥ 1 if $k \leq \frac{n}{2}$.

Hence,

$$\binom{n}{\frac{n}{2}} = 2^{H\left(\frac{1}{2}\right)n+o(n)} = 2^{n+o(n)}.$$

Trivially, we know

$$\binom{n}{\frac{n}{2}} \leq 2^n$$

since $\binom{n}{\frac{n}{2}}$ is the number of subsets of $[n]$ of size $\frac{n}{2}$. Also,

$$\binom{n}{\frac{n}{2}} \geq \left(\frac{n}{\left(\frac{n}{2}\right)}\right)^{\frac{n}{2}} = 2^{\frac{n}{2}},$$

so this approximation is very bad. But,

$$2^n = \sum_{k=0}^{n}\binom{n}{k} \leq \sum_{k=0}^{n}\binom{n}{\frac{n}{2}} = (n+1)\binom{n}{\frac{n}{2}},$$

so we have

$$\binom{n}{\frac{n}{2}} \geq \frac{2^n}{n+1}.$$

Thus,

$$\frac{2^n}{n+1} \leq \binom{n}{\frac{n}{2}} \leq 2^n.$$

> **Question.** What is the correct polynomial term?

Note that

$$\binom{n}{\frac{n}{2}} = \frac{n!}{\left(\frac{n}{2}\right)!\left(\frac{n}{2}\right)!} = 2^n \frac{n!}{\left[\left(\frac{n}{2}\right)!2^{\frac{n}{2}}\right]\left[\left(\frac{n}{2}\right)!2^{\frac{n}{2}}\right]},$$

and we have

$$\begin{aligned}
\frac{n!}{\left[\left(\frac{n}{2}\right)! \cdot 2^{\frac{n}{2}}\right]^2} &= \frac{1 \times 2 \times \cdots \times n}{\left[\left(1 \times 2 \times \cdots \times \frac{n}{2}\right) \times (2 \times 2 \times \cdots \times 2)\right]^2} \\
&= \frac{1 \times 3 \times 5 \times \cdots \times (n-1) \times 2 \times 4 \times 6 \times \cdots \times n}{[2 \times 4 \times 6 \times \cdots \times n]^2} \\
&= \frac{1 \times 3 \times 5 \times \cdots \times (n-1)}{2 \times 4 \times 6 \times \cdots \times n},
\end{aligned}$$

so

$$\binom{n}{\frac{n}{2}} = 2^n R,$$

where

$$R = \frac{1 \times 3 \times 5 \times \cdots \times (n-1)}{2 \times 4 \times \cdots \times n}.$$

Note that

$$\begin{aligned}
R^2 &= \frac{1 \times 1 \times 3 \times 3 \times 5 \times 5 \times \cdots \times (n-1) \times (n-1)}{2 \times 2 \times 4 \times 4 \times \cdots \times n \times n} \\
&= \frac{1}{n+1} \times \prod_{i=1}^{\frac{n}{2}} \frac{(2i-1)(2i+1)}{(2i)^2} = \frac{1}{n+1} \times \prod_{i=1}^{\frac{n}{2}} \left(1 - \frac{1}{(2i)^2}\right) \leq \frac{1}{n+1}.
\end{aligned}$$

On the other hand,

$$R^2 = \frac{1}{2 \times n} \times \prod_{i=1}^{\frac{n}{2}-1} \frac{(2i+1)^2}{(2i)(2i+2)} = \frac{1}{2n} \prod_{i=1}^{\frac{n}{2}-1} \left(1 + \frac{1}{2i(2i+2)}\right) \geq \frac{1}{2n}.$$

Hence,

$$\frac{1}{2n} \leq R^2 \leq \frac{1}{n+1},$$

so we have

$$\frac{1}{\sqrt{2n}} \leq R \leq \frac{1}{\sqrt{n+1}},$$

which means

$$\frac{2^n}{\sqrt{2n}} \leq \binom{n}{\frac{n}{2}} \leq \frac{2^n}{\sqrt{n+1}} \leq \frac{2^n}{\sqrt{n}}.$$

Using Stirling:

$$\binom{n}{\frac{n}{2}} \approx \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{\left(\sqrt{2\pi \frac{n}{2}} \cdot \left(\frac{\frac{n}{2}}{e}\right)^{\frac{n}{2}}\right)^2} = \frac{2^n}{\sqrt{\frac{2}{\pi} n}}.$$

5.3 Partition Function

Let $p(n)$ be the number of ways of writing n as an unordered sum of natural numbers.

> **Question.** How quickly does $p(n)$ goes?

For the upper bound, we can instead count ordered sums. Now since

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

and the number of ordered partition of n into k parts is $\binom{n-1}{k-1}$. Thus, the number of ordered partitions is

$$\sum_{k=1}^{n} \binom{n-1}{k-1} = 2^{n-1},$$

so $p(n) \leq 2^{n-1}$.

Lecture 18

14 Nov. 12:20

As previously seen.

$$e\left(\frac{n}{e}\right)^n \le n! \le (n+1)\left(\frac{n}{2}\right)^n, \quad n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \text{ [Stirling]}.$$

Also,

$$\binom{n}{k}\begin{cases} \approx \dfrac{n^k}{k!}, & \text{when } k = O(1).\\ \le \left(\dfrac{ne}{k}\right)^k, & \text{for all } k,\ \text{tight for } \omega(1) = k = o\left(\sqrt{n}\right).\\ = 2^{\left(H\left(\frac{k}{n}\right)+o(1)\right)n}, & \text{for } k = \Omega(n), \end{cases}$$

where $H(x) = -x\log_2(x) - (1-x)\log_2(1-x)$.

Now we continue to talk about Partition asymptotics. We know

$$\begin{aligned} p(n) &= \#\text{ of ways of writing } n \text{ as an unordered sum of positive integers}\\ &\le \#\text{ of ways of writing } n \text{ as an ordered sum of positive integers}\\ &= \sum_{k=1}^{n}\binom{n-1}{k-1} = 2^{n-1}. \end{aligned}$$

Generating functions

$$P(x) = \sum_{n=0}^{\infty} p(n)x^n = \prod_{j=1}^{\infty}\left(1 + x^{1\cdot j} + x^{2\cdot j} + \dots\right) = \prod_{j=1}^{\infty}\frac{1}{1-x^j}.$$

Now we build a bijection. If $\lambda \vdash n$, with i_j picks of size j, we choose $x^{i_j\cdot j}$ term from the j-th factor. Hence,

$$\prod_{j=1}^{\infty}\frac{1}{1-x^j} = \sum_{k\ge 0} p(k)x^k \ge p(n)x^n$$

for all $x \ge 0$. Hence,

$$p(n) \le \frac{\left(\prod_{j=1}^{\infty}\frac{1}{1-x^j}\right)}{x^n},$$

and thus

$$\begin{aligned} \ln\left(p(n)\right) &\le -n\ln x - \sum_{j=1}^{\infty}\ln\left(1-x^j\right)\\ &= -n\ln x + \sum_{j=1}^{\infty}\sum_{\ell=1}^{\infty}\frac{(x^j)^\ell}{\ell} \quad \text{by Taylor series}\\ &= -n\ln x + \sum_{\ell=1}^{\infty}\frac{1}{\ell}\cdot\sum_{j=1}^{\infty}\left(x^\ell\right)^j = -n\ln x + \sum_{\ell=1}^{\infty}\frac{1}{\ell}\cdot\frac{x^\ell}{1-x^\ell}. \end{aligned}$$

Note that

$$1 - x^\ell = (1-x)\left(1 + x + x^2 + \dots + x^{\ell-1}\right) \ge (1-x)\ell x^{\ell-1} \quad \text{for } 0 < x < 1.$$

Hence,

$$\begin{aligned} \ln\left(p(n)\right) &\le -n\ln x + \sum_{\ell=1}^{\infty}\frac{1}{\ell}\frac{x^\ell}{(1-x)\cdot\ell x^{\ell-1}}\\ &= -n\ln x + \frac{x}{1-x}\sum_{\ell=1}^{\infty}\frac{1}{\ell^2}\\ &= -n\ln x + \frac{\pi^2}{6}\frac{x}{1-x} = n\ln\frac{1}{x} + \frac{\pi^2}{6}\cdot\frac{x}{1-x} \quad \text{for all } x \in (0,1). \end{aligned}$$

Let $u = \frac{x}{1-x}$, then

$$\frac{1}{x} = 1 + \frac{1}{u} \Rightarrow \ln\left(p(n)\right) \leq n \ln\left(1 + \frac{1}{u}\right) + \frac{\pi^2}{6}u \leq \frac{n}{u} + \frac{\pi^2}{6}u,$$

where the last inequality is because

$$\ln\left(1 + \frac{1}{u}\right) \leq \frac{1}{u}.$$

Now we want to find where the maximum of $\frac{n}{u} + \frac{\pi^2}{6}u$ occurs, so suppose

$$\frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{n}{u} + \frac{\pi^2}{6}u\right) = 0,$$

we can check the maximum occurs at $u = \frac{\sqrt{6n}}{\pi}$, so

$$\ln\left(p(n)\right) \leq \frac{n}{u} + \frac{\pi^2}{6}u \leq (\pi\sqrt{n})/\sqrt{6} + \frac{\pi}{\sqrt{6}}\sqrt{n} = \frac{2\pi}{\sqrt{6}}\sqrt{n},$$

which means

$$p(n) \leq e^{\frac{2\pi}{\sqrt{6}}\sqrt{n}}.$$

For the lower bound: Recall $p(n,k)$ is the number of unordered partitions into k parts, and the number of ordered partitions into k parts is $\binom{n-1}{k-1}$. Hence,

$$p(n) \geq p(n,k) \leq \binom{n-1}{k-1}.$$

Notice that

$$\binom{n-1}{k-1} \leq p(n,k)k!$$

since each unordered partition into k parts is counted $\leq k!$ times in $\binom{n-1}{k-1}$. Hence,

$$p(n) \geq p(n,k) \geq \frac{\binom{n-1}{k-1}}{k!} \quad \forall k \geq 1.$$

Hence, we want to choose k to maximize R.H.S. to get a better lower bound. Note that

$$\frac{g(k+1)}{g(k)} = \frac{\frac{\binom{n-1}{k}}{(k+1)!}}{\frac{\binom{n-1}{k-1}}{k!}} = \frac{\binom{n-1}{k}}{\binom{n-1}{k-1}} \cdot \frac{k!}{(k+1)!} = \frac{n-k}{k} \cdot \frac{1}{k+1} \approx \frac{n}{k^2}.$$

At the optimal k, we expect

$$\frac{g(k+1)}{g(k)} \approx 1,$$

so $k \approx \sqrt{n}$, and thus we can swt $k = \lfloor\sqrt{n}\rfloor$. Hence,

$$p(n) \geq g\left(\lfloor\sqrt{n}\rfloor\right) = \frac{\binom{n-1}{\lfloor\sqrt{n}\rfloor - 1}}{(\lfloor\sqrt{n}\rfloor)!} \geq \frac{\left(\frac{n-1}{\lfloor\sqrt{n}\rfloor - 1}\right)^{\lfloor\sqrt{n}\rfloor - 1}}{\left(\frac{\lfloor\sqrt{n}\rfloor}{e}\right)^{\lfloor\sqrt{n}\rfloor}} \approx C^{\sqrt{n}}.$$

Theorem 5.3.1 (Hardy-Ramanujan, 1918).

$$p(n) \approx \frac{1}{4\sqrt{3}}\frac{1}{n}e^{\frac{2\pi}{\sqrt{6}}\sqrt{n}}.$$

Lecture 19

When we calculated the lower bound of $p(n)$, we say 18 Nov. 15:30

$$
\begin{aligned}
p(n) &\geq p(n,k) \geq \frac{\binom{n-1}{k-1}}{k!} = \frac{\frac{k}{n}\binom{n}{k}}{k!} \geq \frac{(\frac{k}{n})(n-k)^k}{(k!)^2} = \frac{k}{n}\cdot\frac{n^k\left(1-\frac{k}{n}\right)^k}{(k!)^2} \\
&\geq \frac{k}{n}\cdot\frac{n^k\left(1-\frac{k}{n}\right)^k}{\left[\left(\frac{k}{e}\right)^k(k+1)\right]^2} \geq \frac{k}{n(k+1)^2}\cdot\left(\frac{ne^2}{k^2}\right)^k\underbrace{\left(1-\frac{k}{n}\right)^k}_{\to e^{-1}} \gtrsim \left(\frac{1}{e}+o(1)\right)\frac{1}{n^{\frac{3}{2}}}e^{2+o(1)\sqrt{n}}
\end{aligned}
$$

Chapter 6

Partially ordered sets

6.1 Inclusion-exclusion principle

Recall the sum rule. If S can be partitioned as $S = S_1 \uplus S_2 \uplus \ldots \uplus S_n$, then $|S| = \sum_{i=1}^{n} |S_i|$.

Example 6.1.1. Suppose 46 friends go out to dinner to celebrate the end of yet another Combinatorics lecture. 27 people eat pork xiaolongbaos and 16 peoples eat vegetarian xiaolongbaos. Then, how many prople doesn't eat any xiaolongbaos at all?

Proof. Draw a Venn-diagram, then we know the information is not enough to solve this problem since we don't know how many people eat both pork and vegetarian xiaolongbaos, so the answer is not unique.

Now we suppose 5 people ate both types of xiaolongbaos. Let

$$\begin{aligned} F &= \{\text{friends}\} \\ X &= \{\text{xiaolongbao eaters}\} \\ H &= \{\text{xiaolongbao non-eaters}\} \\ P &= \{\text{pork xiaolongbaos eaters}\} \\ V &= \{\text{vegetarian xiaolongbaos eaters}\} \\ B &= \{\text{eaten of both}\}, \end{aligned}$$

then we're interested in $|H|$. By sum rule, we know $|F| = |X| + |H|$, so $|H| = |F| - |X| = 46 - |X|$. Note that $X = P \cup V$, and this is not a disjoint union, so we cannot apply the sum rule directly. However, $X = P \uplus (V \setminus (V \cap P)) = P \uplus (V \setminus B)$ is a disjoint union. Hence, $|X| = |P| + |V \setminus B|$. But $V = (V \setminus B) \uplus B$, so by sum rule we have $|V| = |V \setminus B| + |B|$. Hence, $|V \setminus B| = |V| - |B| = 16 - 5 = 11$, which gives $|X| = |P| + |V \setminus B| = 27 + 11 = 38$. Hence, $|H| = |F| - |X| = 46 - 38 = 8$. ⊛

Theorem 6.1.1. If A, B are two sets, then

$$|A \cup B| = |A| + |B| - |A \cap B| \leq |A| + |B|.$$

Proof. Since $A \cup B = A \uplus (B \setminus A)$ and $B = (B \setminus A) \uplus (A \cap B)$. Hence, by sum rule, we have

$$|B \setminus A| = |B| - |A \cap B|, \quad \text{and } |A \cup B| = |A| + |B \setminus A| = |A| + |B| - |A \cap B|.$$

■

Theorem 6.1.2 (Inclusion-Exclusion). Let S be a finite set, with subsets $A_1, A_2, \ldots, A_n \subseteq S$. Then

$$\underbrace{\left| S \setminus \bigcup_{i=1}^{n} A_i \right|}_{\substack{\#\text{ of elements} \\ \text{not in any } A_i}} = \sum_{I \subseteq [n]} (-1)^{|I|} \left| \bigcap_{i \in I} A_i \right|.$$

Note that $\bigcap_{i \in \varnothing} A_i = S$.

Proof. The first choice is induction. The second choice is to count how often an element is counted. Note that we have an identity:

$$\prod_{i=1}^{n} (1 + x_i) = \sum_{I \subseteq [n]} \prod_{i \in I} x_i,$$

where $\prod_{i \in \varnothing} x_i = 1$.

Definition 6.1.1. The characteristic function f_i for A_i is the function

$$f_i : S \to \{0, 1\}, \quad f_i(s) = \begin{cases} 1, & \text{if } s \in A_i; \\ 0, & \text{if } s \notin A_i. \end{cases}$$

Then, for any $s \in S$,

$$\prod_{i=1}^{n} (1 - f_i(s)) = \begin{cases} 1, & \text{if } s \in S \setminus \left(\bigcup_{i=1}^{r} A_i \right); \\ 0, & \text{if } s \in \bigcup_{i=1}^{n} A_i. \end{cases}$$

Hence,

$$\left| S \setminus \left(\bigcup_{i=1}^{n} A_i \right) \right| = \sum_{s \in S} \prod_{i=1}^{n} (1 - f_i(s)).$$

Applying the identity:

$$\begin{aligned} \left| S \setminus \left(\bigcup_{i=1}^{n} A_i \right) \right| &= \sum_{s \in S} \prod_{i=1}^{n} (1 - f_i(s)) = \sum_{s \in S} \sum_{I \subseteq [n]} \prod_{i \in I} (-f_i(s)) \\ &= \sum_{I \subseteq [n]} (-1)^{|I|} \sum_{s \in S} \prod_{i \in I} f_i(s) = \sum_{I \subseteq [n]} (-1)^{|I|} \left| \bigcap_{i \in I} A_i \right| \end{aligned}$$

since we know $\prod_{i \in I} f_i(s) = 1$ iff $s \in \bigcap_{i \in I} A_i$, otherwise it is equal to 0. ■

Corollary 6.1.1.

$$\left| \bigcup_{i=1}^{n} A_i \right| = \sum_{\substack{I \subseteq [n] \\ I \neq \varnothing}} (-1)^{|I|+1} \left| \bigcap_{i \in I} A_i \right|$$

Proof. Note that

$$S = \left(\bigcup_{i=1}^{n} A_i \right) \cup \left(S \setminus \bigcup_{i=1}^{n} A_i \right),$$

so by sum rule we have

$$\left| \bigcup_{i=1}^{n} A_i \right| = |S| - \left| S \setminus \bigcup_{i=1}^{n} A_i \right|,$$

which means

$$\left|\bigcup_{i=1}^{n} A_i\right| = |S| - \left(\sum_{I\subseteq[n]} (-1)^{|I|}\left|\bigcap_{i\in I} A_i\right|\right) = |S| - \left((-1)^{|\varnothing|}\left|\bigcap_{i\in\varnothing} A_i\right| + \sum_{\substack{I\subseteq[n]\\ I\neq\varnothing}} (-1)^{|I|}\left|\bigcap_{i\in I} A_i\right|\right)$$

$$= \sum_{\substack{I\subseteq[n]\\ I\neq\varnothing}} (-1)^{|I|+1}\left|\bigcap_{i\in I} A_i\right|.$$

■

Example 6.1.2 (Derrangements). A professor wants to return 46 midterms to his 46 students. He shuffles the exams uniformly at random and goes one to every student. Then,

$$\mathbb{P}(\text{somebody gets their own exam}) = ?$$

Proof. We do a warm-up first:

$$\mathbb{P}\left(\text{student } \#i \text{ gets his own exam}\right) = \frac{1}{46}.$$

Now we can map the distribution of exams to permutations of $[46]$. Student i gets exam $\pi(i)$, where $\pi \in S_{46}$. Hence, we can define

$$A_i = \{\# \text{ of methods s.t. student } \#i \text{ gets own exam}\} = \{\pi : \pi(i) = i\}.$$

We want to find $\left|\bigcup_{i=1}^{46} A_i\right|$. Hence, we know

$$\left|\bigcup_{i=1}^{46} A_i\right| = \sum_{\substack{I\subseteq[46]\\ I\neq\varnothing}} (-1)^{|I|+1}\left|\bigcap_{i\in I} A_i\right|,$$

where

$$\left|\bigcap_{i\in I} A_i\right| = (46-|I|)!$$

since $\bigcap_{i\in I} A_i$ counts the number of methods s.t. $\pi(i) = i$ for all $i \in I$. Hence,

$$\left|\bigcup_{i=1}^{46} A_i\right| = \sum_{\substack{I\subseteq[46]\\ I\neq\varnothing}} (-1)^{|I|+1}(46-|I|)!,$$

so

$$\mathbb{P}\left(\text{somebody gets their own exam}\right) = \frac{\left|\bigcup_{i=1}^{46} A_i\right|}{46!} = \sum_{\substack{I\subseteq[46]\\ I\neq\varnothing}} \frac{(-1)^{|I|+1}(46-|I|)!}{46!}.$$

Note that our summands only depend on the size of I but not I itself. Hence, we can simplify the answer:

$$\mathbb{P}\left(\text{somebody gets their own exam}\right) = \sum_{k=1}^{46}\sum_{\substack{I\subseteq[46]\\ |I|=k}} (-1)^{k+1}\frac{(46-k)!}{46!}$$

$$= \sum_{k=1}^{46}\binom{46}{k}(-1)^{k+1}\frac{(46-k)!}{46!} = \sum_{k=1}^{46}\frac{(-1)^{k+1}}{k!}.$$

⊛

Hence, if we generalize this problem to be calculating

$$\mathbb{P}\left(\pi \text{ has no fixed point}\right),$$

then we can give a definition:

Definition 6.1.2. $\pi \in S_n$ is a derrangement if $\pi(i) \neq i$ for all $i \in [n]$.

Hence, if we define $A_i = \{\pi : \pi(i) = i\}$, then

$$\mathbb{P}\left(\pi \text{ is an derrangement}\right) = \frac{\left|S_n \setminus (\bigcup_{i=1}^n A_i)\right|}{n!} = \frac{1}{n!}\sum_{I\subseteq[n]}(-1)^{|I|}\left|\bigcap_{i=1}^n A_i\right| = \frac{1}{n!}\sum_{k=0}^n\binom{n}{k}(-1)^k(n-k)! = \sum_{k=0}^n\frac{(-1)^k}{k!}.$$

Hence,

$$\lim_{n\to\infty}\mathbb{P}\left(\pi \in S_n \text{ is a derrangement}\right) = \sum_{k=0}^\infty\frac{(-1)^k}{k!} = e^{-1} = \frac{1}{e},$$

which gives

$$\mathbb{P}\left(\pi \text{ is not a derrangement}\right) \to 1 - \frac{1}{e}.$$

There is a fact that the number of derrangements is the integer closest to $\frac{n!}{e}$.

Lecture 20

25 Nov. 15:30

Example 6.1.3. 100 people boarding a plane with 100 seats, and the first person takes a random seat, while everyone else takes their seat if the seat is free, otherwise a randome empty seat is taken. Then, what is the probability that the final passenger is in his own seat?

Proof. If we consider the sequence if random choices, the process ends when

- we choose #1, choosing the loop and guaranteeing everyone else gets their own seat, or
- we choose #100, so every one in between gets their own seat, but the last passenger will not.

Any other choice: we will have a random choice again later. By symmetry, equally likely to choose #1 or #100 each time, so

$$\mathbb{P}\left(\text{last passenger gets own seat}\right) = \frac{1}{2}.$$

⊛

Application

Stack of 30 assignments to grade, and cycle through them, abd give 100% to every k-th problem. To be fair, we need k to be relatively prime to 30.

Question. How many choices for k are there?

Define $S = [30]$ and $A_d = \{x \in S : d \mid x\}$, then we want

$$\left|S \setminus (A_2 \cup A_3 \cup A_5)\right|,$$

which can be derived by Inclusion-Exclusion:

$$|S| - |A_2| - |A_3| - |A_5| + |A_2 \cap A_3| + |A_3 \cap A_5| + |A_2 \cap A_5| - |A_2 \cap A_3 \cap A_5|.$$

Observe that

$$A_d = \left\lfloor\frac{30}{d}\right\rfloor,$$

and if p, q are coprime, then

$$A_p \cap A_q = A_{pq}.$$

Hence, by induction, if $d_1, d_2, \ldots, d_n$ are pairwise coprime, then

$$A_{d_1} \cap A_{d_2} \cap \cdots \cap A_{d_n} = A_{d_1 d_2 \ldots d_n}.$$

Thus,

$$\begin{aligned}&|S| - |A_2| - |A_3| - |A_5| + |A_2 \cap A_3| + |A_3 \cap A_5| + |A_2 \cap A_5| - |A_2 \cap A_3 \cap A_5| \\ &= 30 - \frac{30}{2} - \frac{30}{3} - \frac{30}{5} + \frac{30}{6} + \frac{30}{15} + \frac{30}{10} - \frac{30}{30} = 8.\end{aligned}$$

Definition 6.1.3 (Euler Totient Function). Given $n \in \mathbb{N}$, we define

$$\varphi(n) = |\{r \in [n] : \gcd(r, n) = 1\}|$$

Question. What is $\varphi(n)$?

If n is prime, then $\varphi(n) = n - 1$. If $n = p^k$ is a prime power, then

$$\varphi(n) = p^k - p^{k-1}.$$

For the general case, $n = p_1^{k_1} p_2^{k_2} \ldots p_s^{k_s}$ where $p_1, p_2, \ldots, p_s$ are the distinct prime factors of n and $k_i \in \mathbb{N}$ for all i. Let $S = [n]$ and $A_d = \{x \in [n] : d \mid x\}$, then

$$\varphi(n) = \left| S \setminus \bigcup_{i=1}^{s} A_{p_s} \right|,$$

so by Inclusion-Exclusion principle,

$$\begin{aligned}\varphi(n) &= \sum_{I \subseteq [s]} (-1)^{|I|} \left| \bigcap_{i \in I} A_{p_i} \right| = \underbrace{n}_{I = \varnothing} + \sum_{\substack{I \subseteq [s] \\ I \neq \varnothing}} (-1)^{|I|} \left| A_{\prod_{i \in I} p_i} \right| = n + \sum_{\substack{I \subseteq [s] \\ I \neq \varnothing}} (-1)^{|I|} \frac{n}{\prod_{i \in I} p_i} \\ &= n \left(\sum_{I \subseteq [s]} \frac{(-1)^{|I|}}{\prod_{i \in I} p_i} \right) = n \left(\sum_{I \subseteq [s]} \prod_{i \in I} \left(-\frac{1}{p_i} \right) \right) = n \prod_{i=1}^{s} \left(1 - \frac{1}{p_i} \right).\end{aligned}$$

Corollary 6.1.2.

$$\varphi(mn) = \varphi(m)\varphi(n)$$

if and only if m and n are coprime.

6.2 Partially ordered sets (Posets)

Example 6.2.1. Ranking items on a member. If

stinky tofu < anything else < xiaolongbao,

then

beef noodle soup?scallion pancakes.

If items come all at once, we don't want there to be a dish and a better dish, because every one would take from the better dish. If a banquet, where one dish is served at a time, maybe we want each dish to be better than the previous.

Definition 6.2.1. A partially ordered set (poset) is a set P together with a binary relation $\leq \subseteq P \times P$ that is

- reflexive: $x \leq x \ \forall x \in P$.
- anti-symmetric: If $x \leq y$ and $y \leq x$, then $x = y$.
- transitive: If $x \leq y$ and $y \leq z$, then $x \leq z$.

Definition 6.2.2. Two elements $x, y \in P$ are comparable if $x \leq y$ or $y \leq x$, and incomparable otherwise.

We can represent a finite poset via a Hasse diagram: If $x \leq y$, we draw x below y. If $x \leq y$ and $\nexists z$ s.t. $x < z < y$, i.e. $x \leq z \leq y$ but $z \neq x, y$, then we draw an edge between x and y.

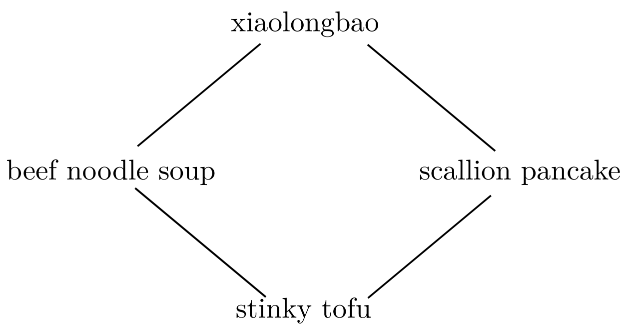


Figure 6.1: A Hasse diagram for the initial example

Definition 6.2.3. A chain is a subset $C \subseteq P$ where every pair is comparable.

Definition 6.2.4. An antichain is a subset $A \subseteq P$ where no two distinct elements are comparable.

Example 6.2.2.

- $\mathbb{N}, \mathbb{Q}, \mathbb{R}$ with their usual ordering.
- $(\mathbb{C}, \leq)$ where $x \leq y$ if $x = y$ or $|x| < |y|$. In particular, circles around the origin are antichains.
- (Boolean poset) $P = 2^{[n]}$ and $a \leq b$ iff $a \subseteq b$.
- (Divisibility poset) $P = [n]$ or $\mathbb{N}$, and $a \leq b$ iff $a \mid b$.

Question. What is the biggest chain in these posets? (maximal of largest cardinality)

Answer.

- $\mathbb{N}, \mathbb{Q}, \mathbb{R}$ respectively are chains.
- Choose one complex number of each modules. (e.g. $\{x \in \mathbb{R}, x \geq 0\}$)
- Cannot have the sets of the same cardinality, so we can pick

$$\varnothing, \{1\}, \{1,2\}, \{1,2,3\}, \ldots, \{1,2,\ldots,n\}.$$

- Powers of 2: $1, 2, 2^2, \ldots, 2^{\lfloor \log_2(n) \rfloor}$. Since we want

$$c_0 \mid c_1 \mid c_2 \mid \cdots \mid c_s$$

and $c_1 \geq 2c_0$ and $c_2 \geq 2c_1$ and so on.

⊛

Question. What are the largest antichain?

Answer.

- Single elements, because every pair is comparable.
- Circles of fixed radius/modules.
- Consider $\binom{[n]}{k}$, then since any two distinct sets of same size are incomparable, so for all $1 \leq k \leq n$ this is true, and the better k is $\lfloor \frac{n}{2} \rfloor$. However, can do better? For example, if n is odd, then

$$A = \binom{[n-1]}{\frac{n+1}{2}} \cup \left\{ S \subseteq [n] : n \in S, |S| = \frac{n-1}{2} \right\}.$$

Then

$$|A| = \binom{n-1}{\frac{n+1}{2}} + \binom{n-1}{\frac{n-3}{2}} < \binom{n-1}{\frac{n+1}{2}} + \binom{n-1}{\frac{n-1}{2}} = \binom{n}{\frac{n+1}{2}} \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

⊛

Theorem 6.2.1 (Sperner 1928). The size of the largest antichain in $(2^{[n]}, \subseteq)$ is

$$\binom{n}{\lfloor \frac{n}{2} \rfloor} \approx \frac{c2^n}{\sqrt{n}}.$$

Lecture 21

28 Nov. 12:20

Proposition 6.2.1 (The LYM inequality). If $A \subseteq 2^{[n]}$ is an antichain, then

$$\sum_{F \in A} \frac{1}{\binom{n}{|F|}} \leq 1.$$

Proof. Let $A \subseteq 2^{[n]}$ be an antichain. We double count pairs (π, F) where

- $\pi \in S_n$ is a permutation of the ground set $[n]$.
- $F \in A$, and
- $F = \{\pi(1), \pi(2), \ldots, \pi(|F|)\}$.

For example, if $n = 7$ and $\pi = 1372465$, then possible F is

$$\{1\}, \{1, 3\}, \{1, 3, 7\}, \{1, 3, 7, 2\}, \ldots \text{e.t.c.}$$

Hence, if we fix π, and we have two pairs (π, F_1) and (π, F_2), then F_1 and F_2 are comparable, but since $F_1, F_2 \in A$, so they should be incomparable. Thus, for each $\pi \in S_n$, we can have at most one set F with (π, F) in our collection. Hence, there are $\leq n!$ pairs. Now fix $F \in A$, how many permutations π give the pair (π, F)? We can order the elements of F arbitrarily at the front, so there are $|F|!$ options. Also, we can order the elements outside F arbitralily at the back, so there are $(n-|F|)!$ options. By the product rule, there are $|F|!(n-|F|)!$ permutations for a given F. By the sum rule, the total number of pairs is

$$\sum_{F\in A} |F|!(n-|F|)! \leq n!.$$

Hence,

$$\sum_{F\in A} \frac{|F|!(n-|F|)!}{n!} \leq 1 \Leftrightarrow \sum_{F\in A} \frac{1}{\binom{n}{|F|}} \leq 1.$$

■

Hence, we can prove Sperner's theorem.

proof of Sperner. Let A be a largest antichain in $2^{[n]}$. Then

$$1 \geq \sum_{F\in A} \frac{1}{\binom{n}{|F|}} \geq \sum_{F\in A} \frac{1}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} = \frac{|A|}{\binom{n}{\lfloor \frac{n}{2} \rfloor}},$$

so

$$|A| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}. \tag{6.1}$$

Suppose $|A| = \binom{n}{\lfloor \frac{n}{2} \rfloor}$, then we have equality in Equation 6.1. Hence, we must have

$$\binom{n}{|F|} = \binom{n}{\lfloor \frac{n}{2} \rfloor} \text{ for every } F \in A.$$

■

Hence, if n is even, then the only possibility is

$$|F| = \frac{n}{2} \text{ for all } F \in A,$$

and thus

$$\binom{[n]}{\frac{n}{2}}$$

is the unique maximum. However, if n is odd, then

$$\binom{n}{|F|} = \binom{n}{\lfloor \frac{n}{2} \rfloor} \text{ is only possible if } |F| \in \left\{\frac{n-1}{2}, \frac{n+1}{2}\right\}.$$

Thus, any maximum antichain satisfies

$$A \subseteq \binom{[n]}{\frac{n-1}{2}} \cup \binom{[n]}{\frac{n+1}{2}},$$

and A is not unique.

Summary

For $\left(2^{[n]}, \subseteq\right)$,

- Largest chain: $n+1$. (Any permutation add elements one at a time)
- Largest antichain: $\binom{n}{\lfloor \frac{n}{2} \rfloor}$. (all sets of size $\binom{n}{\lfloor \frac{n}{2} \rfloor}$)

Definition 6.2.5. Let $(P, \leq)$ be a finite poset. A chain partition/decomposition of P is a partition of P into disjoint chains

$$P = C_1 \uplus C_2 \uplus \ldots \uplus C_m,$$

where an antichain partition is a partition into disjoint antichains

$$P = A_1 \uplus A_2 \uplus \ldots \uplus A_t.$$

Corollary 6.2.1.

$$|\text{Largest chain}| \leq |\text{Smallest antichain partition}|.$$
$$|\text{Largest antichain}| \leq |\text{Smallest chain partition}|.$$

Proof. Observe that in any poset, if $C \subseteq P$ is a chain and $A \subseteq P$ is an antichain, then

$$|C \cap A| \leq 1.$$

Let C be a chain, and

$$A_1 \uplus A_2 \uplus \ldots \uplus A_t \text{ is an antichain partition,}$$

then

$$|C| = |C \cap P| = \left| C \cap \left(\bigcup_{i=1}^{t} A_i \right) \right| = \left| \bigcup_{i=1}^{t} (C \cap A_i) \right| = \sum_{i=1}^{t} |C \cap A_i| \leq \sum_{i=1}^{t} 1 = t.$$

Now let A be an antichain, and

$$C_1 \uplus C_2 \uplus \ldots \uplus C_\ell$$

be a chain decomposition, then

$$|A| = |A \cap P| = \left| A \cap \left(\bigcup_{i=1}^{\ell} C_i \right) \right| = \left| \bigcup_{i=1}^{\ell} (A \cap C_i) \right| = \sum_{i=1}^{\ell} |A \cap C_i| \leq \sum_{i=1}^{\ell} 1 = \ell.$$

■

Lecture 22

Definition 6.2.6. Let $(P, \leq)$ be a finite poset. The height of P, $h(P)$ is the size of the largest chain in P. The width of P, $w(P)$, is the size of the largest antichain in P. 2 Dec. 15:30

Theorem 6.2.2. Let $(P, \leq)$ be a finite poset. Then,

$$\begin{aligned} &\text{(Mirsky)} \quad h(P) = \text{size of smallest antichain partition } P = A_1 \uplus A_2 \uplus \ldots \uplus A_h \\ &\text{(Dilworth)} \quad w(P) = \text{size of smallest chain partition } P = C_1 \uplus C_2 \uplus \ldots \uplus C_w. \end{aligned}$$

proof of Mirsky. We do induction on $h(P)$.

- Base case: $h(P) = 1$, then P must be an antichain since any two elements of P are incomparable, so $h = 1$, and we're done.
- Induction step: $(h(P) \geq 2)$ Let M be the set of maximal elements in P, i.e.

$$M = \{x \in P : \nexists y \in P \setminus \{x\} \text{ s.t. } x \leq y\}.$$

Claim 6.2.1.

(i) M is an antichain.

(ii) Every maximum chain in P intersects M.

Proof.

(i) If $x_1, x_2 \in M$ and $x_1 \leq x_2$, then $x_1 \notin M$ by definition.

(ii) If C is a chain, then we can write

$$C = \{c_1, c_2, \ldots, c_t\} \text{ with } c_1 \leq c_2 \leq \cdots \leq c_t.$$

If $c_t \notin M$, then by definition, there exists $y \neq c_t$ s.t. $y \geq c_t$, and thus

$$c_1 \leq c_2 \leq \cdots \leq c_t \leq y$$

is a larger chain. Hence, every maximum chain in P intersects M.

⊛

Now let $P' = P \setminus M$, then $h(P') \leq h(P) - 1$. By the induction hypothesis, we have an antichain partition

$$P' = A_1 \cup A_2 \cup \cdots \cup A_{h(P)-1},$$

then $P = A_1 \cup A_2 \cup \cdots \cup A_{h(P)-1} \cup M$ is an antichain partition of P.

Remark 6.2.1. If $h(P') < h(P) - 1$, then by induction hypothesis there exists an antichain decomposition:

$$A_1 \uplus A_2 \uplus \ldots \uplus A_{h(P')},$$

and we can seperate some antichain into 2 parts or more parts to make an antichain decomposition of size $h(P) - 1$.

■

proof of Dilworth. We prove by induction on $|P|$.

- Base case: $|P| = 1$. In this case, P is itself a chain, so $P = P$ is the smallest chain partition, and we're done.
- Induction step: ($|P| \geq 2$) If there exists a chain C that intersecting every maximum antichain, then we can proceed as before:
  - Let $P' = P \setminus C$.
  - $w(P') \leq w(P) - 1$.
  - By induction hypothesis there exists a chain partition $P' = C_1 \cup C_2 \cup \cdots \cup C_{w(P)-1}$ and thus

    $$P = C_1 \cup C_2 \cup \cdots \cup C_{w(P)-1} \cup C$$

    finish the proof.
- Otherwise, let $C \subseteq P$ be a maximum chain. Then, there must be a maximum antichain

  $$A = \{a_1, a_2, \ldots, a_w\} \subseteq P$$

  that is disjoint from C. Define

  $$P^+ = \{x \in P : x \geq a_i \text{ for some } a_i\}$$
  $$P^- = \{x \in P : x \leq a_i \text{ for some } a_i\}.$$

Claim 6.2.2.

(i) $P^+ \cap P^- = A$

(ii) $P^+ \cup P^- = P$

(iii) $P^-, P^+ \subsetneq P$.

Proof. Since for all i, we know $a_i \leq a_i$, so $a_i \in P^+$ and $a_i \in P^-$, and thus $A \subseteq P^+ \cap P^-$. Now if $z \in P^+ \cap P^-$ but $z \notin A$, then there exists i, j s.t. $z \geq a_i$ and $z \leq a_j$. By transitivity, $a_i \leq z \leq a_j$ gives $a_i \leq a_j$. However, A is an antichain, so $a_i = a_j$, but the anti-symmetry gives $z = a_i \in A$, which is a impossible. Hence, $z \in A$. This proves (i).

Now we show (ii). If $P^+ \cup P^- \neq P$, then there exists $z \in P$ s.t. $z \notin P^+$ and $z \notin P^-$, then z is incomparable to everything in A, so $A \cup \{z\}$ is a larger antichain, which is a contradiction.

Now we show (iii). We argue that $C \cap (P^+ \setminus A)$ and $C \cap (P^- \setminus A)$ are both non-empty. Let $C = \{c_1 \leq c_2 \leq \cdots \leq c_t\}$. If $c_t \in P^-$, then there exists $a_i \in A$ s.t. $c_t \leq a_i$, but $c_t \neq a_i$ because $C \cap A = \varnothing$. Thus,

$$c_1 \leq c_2 \leq \cdots \leq c_t \leq a_i$$

is a larger chain, which is a contradiction. Hence, $c_t \in P^+ \setminus A$ since $P^+ \cup P^- = P$, and this shows $P^- \neq P$. Similarly, $c_1 \in P^- \setminus A$ and thus $P^+ \neq P$. ⊛

By induction, we fixed a chain decompositions:

$$\begin{aligned} P^+ &= C_1^+ \cup C_2^+ \cup \cdots \cup C_w^+ \\ P^- &= C_1^- \cup C_2^- \cup \cdots \cup C_w^-. \end{aligned}$$

Relabeling the chains if needed, we may assume C_i^- ends at a_i and C_i^+ starts at a_i. Hence,

$$C_i = C_i^- \cup C_i^+$$

is a chain, and then

$$P = C_1 \cup C_2 \cup \cdots \cup C_w$$

is a chain partition.

Remark 6.2.2. Each C_i^+ can contain only one element in it, and C_i^- has same property, so there exists a chain partition of P^+ and P^-, of size w by induction.

■

Remark 6.2.3. It is not true in general that every maximum chain intersects every maximal antichain. Note that the difference between maximum and maximal:

- maximum means largest in size.
- maximal means is not contained in a large ...

Example 6.2.3. $(2^{[3]}, \subseteq)$, and $C = \{\varnothing, \{1\}, \{1,2\}, \{1,2,3\}\}$ does not intersected $A = \{\{2\}, \{1,3\}\}$.

Application

Definition 6.2.7. A monotone subsequence that is either increasing or decreasing.

Question. Given a sequence of m real numbers where $m \geq 1$, then how large a

(i) increasing subsequence

(ii) decreasing subsequence

(iii) monotone subsequence

are guaranteed to find?

For (i), any number is an increasing subsequence, and if original sequence is decreasing, then cannot find larger. As for 2, cannot find a decreasing subsequence of length larger than 1 if the original sequence is increasing. As for (iii), the below theorem tells us the answer is $\sqrt{m}$.

Theorem 6.2.3 (Erdos-Szekeres, 1935). If $m \geq (r-1)(s-1)+1$, then any sequence of m real numbers contains an increasing subsequence of length r or a decreasing subsequence of length s.

Corollary of Dilworth/Mirsky. Let $x_1, x_2, \ldots, x_m$ be our sequence, then we define

$$P = \{(i, x_i) : i \in [m]\}$$

where $(i, x_i) \leq (j, x_j)$ iff $i \leq j$ and $x_i \leq x_j$. We can observe that the chain in this poset is equivalent to an increasing subsequence, and the antichain is equivalent to an (strictly) decreasing subsequence. By Mirsky, we have an antichain partition

$$P = A_1 \cup A_2 \cup \cdots \cup A_{h(P)}.$$

If $h(P) \geq r$, we have a chain of length $\geq r$, and then we have an increasing subsequence of length $\geq r$. Otherwise $h(P) \leq r-1$, so by averaging, some A_i has

$$|A_i| \geq \frac{m}{r-1} > s-1,$$

so $|A_i| \geq s$, which means there is a decreasing subsequence of length s. ■

Remark 6.2.4. $(r-1)(s-1)+1$ is best possible. For $m = (r-1)(s-1)$, then consider

$$\begin{array}{cccc}
(r-1)(s-2)+1, & (r-1)(s-2)+2, & \ldots & (r-1)(s-1), \\
(r-1)(s-3)+1, & (r-1)(s-3)+2, & \ldots & (r-1)(s-2), \\
\vdots & & & \\
(r-1)+1, & (r-1)+2, & \ldots & (r-1)\cdot 2, \\
1, & 2, & \cdots & (r-1),
\end{array}$$

and if we read this from top to down then from left to right, we find this sequence's largest decreasing subsequence is of size $s-1$ and the largest increasing subsequence is of size $r-1$.

6.3 Mobius Inversion

Example 6.3.1. Let $a_1, a_2, \ldots$ be a sequence of real numbers. Let $s_1, s_2, \ldots$ be the sequence of cummulative sums:

$$s_n = \sum_{i \leq n} a_i.$$

Then, given $s_1, s_2, \ldots$, can we recover $a_1, a_2, \ldots$?

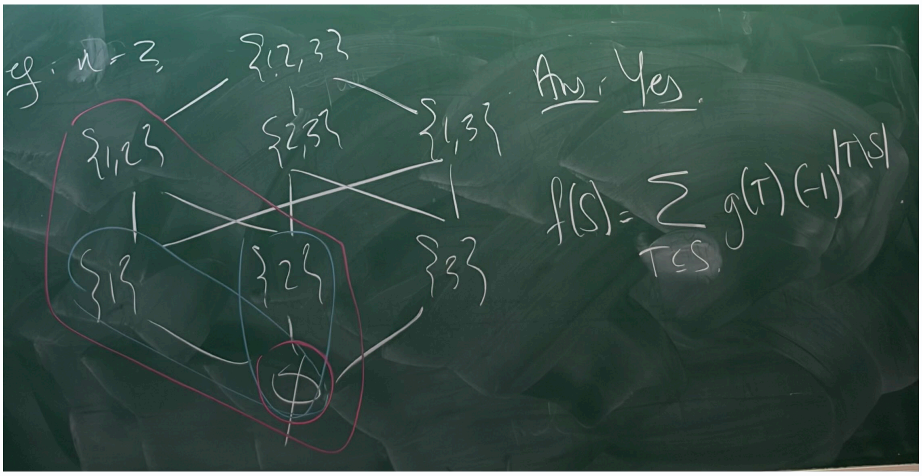
Proof. Yes. Since

$$s_n - s_{n-1} = \sum_{i \leq n} a_i - \sum_{i \leq n-1} a_i = a_n.$$

⊛

Example 6.3.2. Suppose we have a function $f : 2^{[n]} \to \mathbb{R}$. Define $g(S) = \sum_{T \subseteq S} f(T)$, then given g can we recover f?

Proof. Yes. $f(S) = \sum_{T \subseteq S} g(T)(-1)^{|S \setminus T|}$.



Since

$$\begin{aligned}\sum_{T \subseteq S} g(T)(-1)^{|S\setminus T|} &= \sum_{T\subseteq S}\left(\sum_{S' \subseteq T} f(S')\right)(-1)^{|S\setminus T|}\\ &= \sum_{S'\subseteq S} f(S') \sum_{S' \subseteq T \subseteq S} (-1)^{|S\setminus T|},\end{aligned}$$

and

$$\begin{aligned}\sum_{S'\subseteq T\subseteq S} (-1)^{|S\setminus T|} = \sum_{T' \subseteq S\setminus S'} (-1)^{|(S\setminus S')\setminus T'|} &= \prod_{x \in S\setminus S'} (1 + (-1))\\ &= \begin{cases} 0, & \text{if } S \setminus S' \neq \varnothing; \\ 1, & \text{if } S' = S. \end{cases}\end{aligned}$$

Hence,

$$\sum_{T\subseteq S} g(T)(-1)^{|S\setminus T|} = f(S).$$

⊛

Lecture 23

5 Dec. 12:20

Definition 6.3.1. A poset $(P, \leq)$ is locally finite if for every $x, y \in P$, the interval

$$[x, y] = \{z \in P : x \leq z \leq y\}$$

is finite.

Example 6.3.3. For $(\mathbb{N}, \leq), (\mathbb{Z}, \leq), \left(2^{[n]}, \subseteq\right), (\mathbb{N}, \cdot \mid \cdot)$, then they are all locally finite.

Definition 6.3.2. Given a locally finite poset $(P, \leq)$, the incidence algebra is the set of functions

$$I(P) = \left\{ f : P^2 \to \mathbb{R} \text{ s.t. } f(x, y) = 0 \quad \forall x \not\leq y \right\}.$$

Remark 6.3.1. We can think of the incidence algebra as the space of functions on the non-empty intervals $[x, y]$.

Theorem 6.3.1. The incidence algebra is equipped with operations:

- (pointwise) sums: $\forall f, g \in I(P)$,
$$(f + g)(x, y) = f(x, y) + g(x, y).$$
- scalar multiplication: $\forall f \in I(P), \lambda \in \mathbb{R}$,
$$(\lambda f)(x, y) = \lambda f(x, y).$$
- multiplication (convolution): $\forall f, g \in I(P)$,
$$(f * g)(x, y) = \sum_{z \in [x,y]} f(x, z)g(z, y).$$

Remark 6.3.2. Let $<$ be a linear extension of $(P, \leq)$ is a total ordering of P, i.e. $\forall x \neq y$, $x < y$ or $y < x$, and if $x \leq y$ but $x \neq y$, then $x < y$.

Proposition 6.3.1. We can write $f \in I(P)$ as a matrix whose rows/columns are indexed by P, and the corresponding matrix is upper triangular.

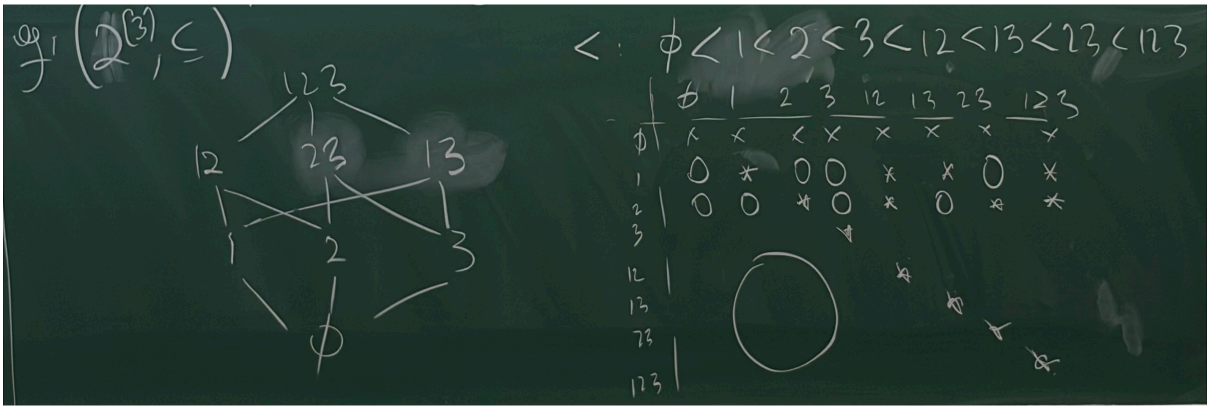


Figure 6.2: An example of $\left(2^{[3]}, \subseteq\right)$

With this correspondence,
$$M_{f*g} = M_f M_g.$$
Thus, the multiplication (convolution) is not necessarily commutative, i.e.
$$f * g \neq g * f,$$
but it is associative, i.e.
$$(f * g) * h = f * (g * h).$$

Identity

There is an identity: the delta function
$$\delta(x, y) = \begin{cases} 1, & \text{if } x = y; \\ 0, & \text{if } x \neq y. \end{cases},$$
then $M_\delta = I_{|P|}$.

Inverses

For every $f \in I(P)$, there is a unique inverse $f^{-1} \in I(P)$ s.t.

$$f * f^{-1} = \delta = f^{-1} * f \Leftrightarrow f(x,x) \neq 0 \quad \forall x \in P.$$

We can define it recursively:

$$f^{-1}(x,x) = \frac{1}{f(x,x)}.$$

Then,

$$1 = \delta(x,x) = \left(f * f^{-1}\right)(x) = f(x,x) f^{-1}(x,x).$$

What about $f^{-1}(x,y)$ for $x \leq y$ but $x \neq y$? Suppose we have already define $f^{-1}(z,y)$ for all $x < z \leq y$, then

$$\begin{aligned} 0 = \delta(x,y) = \left(f * f^{-1}\right)(x,y) &= \sum_{z \in [x,y]} f(x,z) f^{-1}(z,y) \\ &= f(x,x) f^{-1}(x,y) + \sum_{x < z \leq y} f(x,z) f^{-1}(z,y), \end{aligned}$$

so we can define

$$f^{-1}(x,y) = \frac{-\sum_{x<z\leq y} f(x,z) f^{-1}(z,y)}{f(x,x)}.$$

Remark 6.3.3. f^{-1} is unique since

$$I = M_\delta = M_{f * f^{-1}} = M_f M_{f^{-1}},$$

and if the inverse of M_f exists, then this inverse, which is $M_{f^{-1}}$, is unique, so f^{-1} is unqiue. Also, since M_f is upper triangular, so $\det M_f$ is the product of all entries on the diagonal, which is

$$\prod_{x \in P} f(x,x),$$

so we need $f(x,x) \neq 0$ for all $x \in P$ to ensure the existence of the inverse.

Definition 6.3.3. The zeta function ζ_P of a locally finite poset P is

$$\zeta_P(x,y) = \begin{cases} 1, & \text{if } x \leq y; \\ 0, & \text{if } x \not\leq y. \end{cases},$$

i.e. $\zeta_P \equiv 1$ on the non-empty intervals.

Definition 6.3.4. Given a locally finite poset $(P, \leq)$, the mobius function μ_P is the inverse of the zeta function:

$$\mu_P = \zeta_P^{-1}.$$

Observe that since $\zeta_P(x,x) = 1 \neq 0$ for all $x \in P$, so $\mu_P = \zeta_P^{-1}$ exists:

$$\mu_P(x,x) = \frac{1}{\zeta_P(x,x)} = 1 \quad \forall x \in P$$

and for all $x < y$ we have

$$\mu_P(x,y) = -\sum_{x<z\leq y} \mu_P(z,y) = -\sum_{x \leq z < y} \mu_P(x,z).$$

Appendix