Combinatorics

Instructor: Prof. Jie Ma

Combinatorics 2022, Tsinghua

- This class notes will be updating throughout this course.
- The course website can be found at https://ymsc.tsinghua.edu.cn/info/1050/2595.htm

Contents

1 Enumeration

First we give some standard notation that will be used throughout this course.

- Let n be a positive integer. We will use $[n]$ to denote the set $\{1, 2, ..., n\}$.
- Given a set X, let $|X|$ denote the size of X, that is the number of elements contained in X.
- We use "#" to express the word "number".
- The *factorial* of n is the product

$$n! = n \cdot (n-1) \cdots 2 \cdot 1,$$

which can be extended to all non-negative integers by letting $0! = 1$.

1.1 Binomial Coefficients

Let X be a set of size n. Define $2^X = \{A : A \subseteq X\}$ to be the family of all subsets of X. Since the size of 2^X is equal to the number of binary vectors of length $|X|$ or the number of functions from X to $\{0, 1\}$, we have $|2^X| = 2^{|X|} = 2^n$.

Let $\binom{X}{k} = \{A : A \subseteq X, |A| = k\}$, we will use $\binom{n}{k}$ to denote $|\binom{X}{k}|$. For $n < k$, we know that $\binom{n}{k} = 0$ by definition.

Fact 1.1. *For integers $n > 0$ and $0 \leq k \leq n$, we have $|\binom{X}{k}| = \binom{n}{k} = \frac{n!}{k!(n-k)!}$.*

Proof. If $k = 0$, then it is clear that $|\binom{X}{0}| = |\{\emptyset\}| = 1 = \binom{n}{0}$. Now we consider $k > 0$. Let

$$(n)_k := n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}.$$

First we will show that number of ordered k-tuples $(x_1, x_2, \ldots, x_k)$ with distinct $x_i \in X$ is $(n)_k$. There are n choices for the first element x_1. When $x_1, \ldots, x_i$ is chosen, there are exactly $n - i$ choices for the element x_{i+1}. So the number of ordered k-tuples $(x_1, x_2, \ldots, x_k)$ with distinct $x_i \in X$ is $(n)_k$. Since any subset $A \in \binom{X}{k}$ corresponds to $k!$ ordered k-tuples, it follows that $|\binom{X}{k}| = \frac{(n)_k}{k!} = \frac{n!}{k!(n-k)!}$. This finishes the proof. ∎

Next we discuss more properties of binomial coefficients.

Fact 1.2. (1). $\binom{n}{k} = \binom{n}{n-k}$ *for $0 \leq k \leq n$.*
(2). $2^n = \sum_{0 \leq k \leq n} \binom{n}{k}$.
(3). $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$.*(Pascal's identity)*

Proof. (1) is trivial. Since $2^{[n]} = \cup_{0 \leq k \leq n} \binom{[n]}{k}$, we see $2^n = \sum_{0 \leq k \leq n} \binom{n}{k}$, proving (2). Finally, we consider (3). Note that the first term on the right hand side $\binom{n-1}{k-1}$ is the number of k-sets containing a fixed element, while the second term $\binom{n-1}{k}$ is the number of k-sets avoiding this element. So their summation gives the total number of k-sets in $[n]$, which is $\binom{n}{k}$. This finishes the proof. ∎

Pascal's triangle is a triangular array constructed by summing adjacent elements in preceding rows. By Fact 1.2 (3), in the following graph we have that the k-th element in the $n+1$ row is $\binom{n}{k-1}$.

$$
\begin{array}{ccccccccccccccccccccc}
&&&&&&&&&&1&&&&&&&&&& \\
&&&&&&&&&1&&1&&&&&&&&& \\
&&&&&&&&1&&2&&1&&&&&&&& \\
&&&&&&&1&&3&&3&&1&&&&&&& \\
&&&&&&1&&4&&6&&4&&1&&&&&& \\
&&&&&1&&5&&10&&10&&5&&1&&&&& \\
&&&&1&&6&&15&&20&&15&&6&&1&&&& \\
&&&1&&7&&21&&35&&35&&21&&7&&1&&& \\
&&1&&8&&28&&56&&70&&56&&28&&8&&1&& \\
&1&&9&&36&&84&&126&&126&&84&&36&&9&&1& \\
1&&10&&45&&120&&210&&252&&210&&120&&45&&10&&1
\end{array}
$$

Fact 1.3. *The number of integer solutions* $(x_1, \ldots, x_n)$ *to the equation* $x_1 + \cdots + x_n = k$ *with each* $x_i \in \{0, 1\}$ *is* $\binom{n}{k}$.

Fact 1.4. *The number of integer solutions* $(x_1, \ldots, x_n)$ *to the equation* $x_1 + \cdots + x_n = k$ *with each* $x_i > 0$ *is* $\binom{k-1}{n-1}$.

Proof. This question is equivalent to ask: How many ways are there of distributing k sweets to n children such that each child has at least one sweet.

Lay out the sweets in a single row of length k, and cut it into n pieces. Then give the sweets of the i_{th} piece to child i, which means that we need $n-1$ cuts from $k-1$ possibles. ∎

Fact 1.5. *The number of integer solutions* $(x_1, \ldots, x_n)$ *to the equation* $x_1 + \cdots + x_n = k$ *with each* $x_i \geq 0$ *is* $\binom{n+k-1}{n-1}$.

Proof 1. Let $A = \{$integer solutions $(x_1, \ldots, x_n)$ to $x_1 + \cdots + x_n = k$, $x_i \geq 0\}$ and $A = \{$integer solutions $(y_1, \ldots, y_n)$ to $y_1 + \cdots + y_n = n + k$, $y_i > 0\}$. Then $|B| = \binom{n+k-1}{n-1}$ by Fact 1.4.

Define $f : A \to B$, by $f((x_1, \ldots, x_n)) = (x_1 + 1, \ldots, x_n + 1)$. It suffices to check that f is a bijection, which we omit here.

∎

Proof 2. Suppose we have k sweets (of the same sort), which we want to distribute to n children. In how many ways can we do this? Let x_i denote the number of sweets we give to the i-th child, this question is equivalent to that state above.

We lay out the sweets in a single row of length r and let the first child pick them up from left to right (can be 0). After a while we stop him/her and let the second child pick up sweets, etc. The distribution is determined by the specifying the place of where to start a new child. This is equal to select $n-1$ elements from $n+r-1$ elements to be the child, others be the sweets (the first child always starts at the beginning). So the answer is $\binom{n+k-1}{n-1}$. ∎

Exercise 1.6. *Let* $X = [n]$, $A = \{(a_1, a_2, \ldots, a_r) | a_i \in X \ , 1 \le a_1 \le a_2 \le \cdots \le a_r \le n, a_{i+1} - a_i \ge k+1, i \in [r-1]\}$*. Prove that* $|A| = \binom{n-k(r-1)}{r}$*.*

Exercise 1.7. *Give a Combinatorial proof of*

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

Exercise 1.8. *Give a Combinatorial proof of*

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

1.2 Counting Mappings

Define X^Y to be the set of all functions $f : Y \to X$.

Fact 1.9. $|X^Y| = |X|^{|Y|}$.

Proof. Let $|Y| = r$. We can view X^Y as the set of all strings $x_1 x_2 \cdots x_r$ with elements $x_i \in X$, indexed by the r elements of Y. So $|X^Y| = |X|^{|Y|}$. ∎

Fact 1.10. *The number of injective functions* $f : [r] \to [n]$ *is* $(n)_r$*.*

Proof. We can view the injective function f as an ordered k-tuple $(x_1, x_2, \ldots, x_r)$ with distinct $x_i \in X$, so the number of injective functions $f : [r] \to [n]$ is $(n)_r$. ∎

Definition 1.11 (The Stirling number of the second kind). *Let* $S(r, n)$ *be the number of partitions of* $[r]$ *into* n *unordered non-empty parts.*

Exercise 1.12. *Prove that*

$$S(r, 2) = \frac{2^r - 2}{2} = \frac{1}{2} \sum_{i=1}^{r-1} \binom{r}{i}.$$

Fact 1.13. *The number of surjective functions* $f : [r] \to [n]$ *is* $n!S(r, n)$*.*

Proof. Since f is a surjective function if and only if for any $i \in [n], f^{-1}(i) \neq \emptyset$ if and only if $\cup_{i \in [n]} f^{-1}(i) = [r]$, and $S(r, n)$ is the number of partition of $[r]$ into n unordered non-empty parts, we have the number of surjective functions $f : [r] \to [n]$ is $n!S(r, n)$. ∎

We say that any injective $f : X \to X$ is a **permutation** of X (also a bijection). We may view a permutation in two ways: (1) it is a bijective from X to X. (2) a reordering of X.

Cycle notation describes the effect of repeatedly applying the permutation on the elements of the set. It expresses the permutation as a product of cycles; since distinct cycles are disjoint, this is referred to as "decomposition into disjoint cycles".

Definition 1.14 (The Stirling number of the first kind). *Let* $s(r, n)$ *be the number of permutations of* $[r]$ *with exactly* n *cycles multiplied by* $(-1)^{(r-n)}$*.*

The following fact is a direct consequence of Fact 1.10.

Fact 1.15. *The number of permutations of $[n]$ is $n!$.*

Exercise 1.16. *(1) Let $S(r,n) = \left\{ {r \atop n} \right\}$,give a Combinatorial proof of $\left\{ {n \atop k} \right\} = \left\{ {n-1 \atop k-1} \right\} + k \left\{ {n-1 \atop k} \right\}$.*

(2) Let $s(n,k) = (-1)^{n-k} \left[{n \atop k} \right]$, give a Combinatorial proof of $\left[{n \atop k} \right] = \left[{n-1 \atop k-1} \right] + (n-1) \left[{n-1 \atop k} \right]$.

1.3 The Binomial Theorem

Define $[x^k]f$ to be the coefficient of the term x^k in the polynomial $f(x)$.

Fact 1.17. *For $j = 1, 2, \ldots, n$, let $f_j(x) = \sum_{k \in I_j} x^k$ where I_j is a set of non-negative integers, and let $f(x) = \prod_{j=1}^{n} f_j(x)$. Then, $[x^k]f$ equals the number of solutions $(i_1, i_2, \ldots, i_n)$ to $i_1 + i_2 + \cdots + i_n = k$, where $i_j \in I_j$.*

Fact 1.18. *Let $f_1, \ldots, f_n$ be polynomials and $f = f_1 f_2 \cdots f_n$. Then,*

$$[x^k]f = \sum_{i_1+\cdots+i_n=k, i_j \geq 0} \left(\prod_{j=1}^{n} [x^{i_j}] f_j \right).$$

Theorem 1.19 (The Binomial Theorem)**.** *For any real x and any positive integer n, we have*

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

Proof 1. Let $f = (1+x)^n$. By Fact 1.17 we have $[x^k]f$ equals the number of solutions $(i_1, i_2, ..., i_n)$ to $i_1 + i_2 + \cdots + i_n = k$ where $i_j \in \{0, 1\}$, so $[x^k]f = \binom{n}{k}$. ∎

Proof 2. By induction on n. When $n = 1$, it is trivial. If the result holds for $n-1$, then $(1+x)^n = (1+x)(1+x)^{n-1} = (1+x)\sum_{i=0}^{n-1} \binom{n-1}{i} x^i = \sum_{i=1}^{n-1} (\binom{n-1}{i} + \binom{n-1}{i-1}) x^i + 1 + x^n$. Since $\binom{n-1}{i} + \binom{n-1}{i-1} = \binom{n}{i}$ and $\binom{n}{0} = \binom{n}{n} = 1$, we have $(1+x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i$. ∎

Fact 1.20. $\binom{2n}{n} = \sum_{i=0}^{n} \binom{n}{i}^2 = \sum_{i=0}^{n} \binom{n}{i}\binom{n}{n-i}$.

Proof 1. Since $(1+x)^{2n} = (1+x)^n(1+x)^n$, by Fact 1.18, we have $\binom{2n}{n} = [x^n](1+x)^{2n} = \sum_{i=0}^{n} ([x^i](1+x)^n)([x^{n-i}](1+x)^n) = \sum_{i=0}^{n} \binom{n}{i}\binom{n}{n-i} = \sum_{i=0}^{n} \binom{n}{i}^2$. ∎

Proof 2. (It is easy to find a combinatorial proof.) ∎

Exercise 1.21 (Vandermonde's Convolution Formula)**.**

$$\binom{n+m}{k} = \sum_{j=0}^{k} \binom{n}{j}\binom{m}{k-j} = \sum_{i+j=k} \binom{n}{i}\binom{m}{j}.$$

Exercise 1.22.

$$\binom{n+m}{r+m} = \sum_{i-j=r} \binom{n}{i}\binom{m}{j}.$$

Exercise 1.23. *Prove that*

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

by Binomial Theorem.

Fact 1.24. (1).

$$\sum_{\textit{all even } k} \binom{n}{k} = \sum_{\textit{all odd } k} \binom{n}{k} = 2^{n-1}.$$

(2).

$$\sum_{k=0}^{n} k\binom{n}{k} = n2^{n-1}.$$

Proof. (1). We see that $(1+x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i$. Taking $x = 1$ and $x = -1$, we have

$$\sum_{\text{all even } k} \binom{n}{k} = \sum_{\text{all odd } k} \binom{n}{k} = 2^{n-1}.$$

(2). Let $f(x) = (1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$. Then $f'(x) = n(1+x)^{n-1} = \sum_{k=0}^{n} k\binom{n}{k} x^{k-1}$. Let $x = 1$, then we have $\sum_{k=0}^{n} k\binom{n}{k} = n2^{n-1}$. ∎

Definition 1.25. *Let $k_j \geq 0$ be integers satisfying that $k_1 + k_2 + \cdots + k_m = n$. We define*

$$\binom{n}{k_1, k_2, \cdots, k_m} := \frac{n!}{k_1!k_2!\cdots k_m!}.$$

- When $m = 2, \binom{n}{k_1,k_2} = \binom{n}{k_1}$ is the number of binary vectors of length n with k_1 zero and k_2 ones, which is also the number of ordered partitions of $[n]$ into 2 parts such that the i_{th} part has size k_i.
- When $m \geq 3, \binom{n}{k_1,k_2,\cdots,k_m}$ is the number of $m-$ary vectors of length n over $[m]$ such that i occurs k_i times, which is also the number of ordered partitions of $[n]$ into m parts such that the i_{th} part has size k_i.

The following theorem is a generalization of the binomial theorem.

Exercise 1.26 (Multinomial Theorem). *For any reals $x_1, \ldots, x_m$ and any positive integer $n \geq 1$, we have*

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{k_1+k_2+\cdots+k_m=n,\ k_j \geq 0} \binom{n}{k_1, k_2, \cdots, k_m} x_1^{k_1} x_2^{k_2} \cdots x_m^{k_m}.$$

Exercise 1.27. *Suppose $\sum_{i=1}^{m} k_i = n$ with $k_i \geq 1$ for all $i \in [m]$. Then*

$$\binom{n}{k_1, k_2, \cdots, k_m} = \binom{n-1}{k_1 - 1, k_2, \cdots, k_m} + \cdots + \binom{n-1}{k_1, k_2, \cdots, k_m - 1}.$$

1.4 Inclusion and Exclusion Principle (IEP)

This lecture is devoted to Inclusion-Exclusion formula and its applications.

Let Ω be a ground set and let $A_1, A_2, ..., A_n$ be subsets of Ω. Write $A_i^c = \Omega \backslash A_i$. Throughout this lecture, we use the following notation.

Definition 1.28. *Let $A_\emptyset = \Omega$. For any nonempty subset $I \subseteq [n]$, let*

$$A_I = \bigcap_{i \in I} A_i.$$

For any integer $k \geq 0$, let

$$S_k = \sum_{I \in \binom{[n]}{k}} |A_I|.$$

Now we introduce Inclusion-Exclusion formula (in three equivalent forms) and give two proofs as follows.

Theorem 1.29 (Inclusion-Exclusion Formula)**.** *We have*

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{k=1}^{n} (-1)^{k+1} S_k,$$

which is equivalent to

$$\left| \Omega \backslash \bigcup_{i=1}^{n} A_i \right| = |A_1^c \cap A_2^c \cap \cdots \cap A_n^c| = \sum_{k=0}^{n} (-1)^k S_k,$$

and

$$\left| \Omega \backslash \bigcup_{i=1}^{n} A_i \right| = |A_1^c \cap A_2^c \cap \cdots \cap A_n^c| = \sum_{I \subseteq [n]} (-1)^{|I|} |A_I|.$$

Proof (1). For any subset $X \subseteq \Omega$, we define its characterization function $\mathbb{1}_X : \Omega \to \{0, 1\}$ by assigning

$$\mathbb{1}_X(x) = \begin{cases} 1, & x \in X \\ 0, & x \notin X. \end{cases}$$

Then we notice that $\sum_{x \in \Omega} \mathbb{1}_X(x) = |X|$. Let $A = A_1 \cup A_2 \cup \cdots \cup A_n$. Our key observation is that

$$(\mathbb{1}_A - \mathbb{1}_{A_1})(\mathbb{1}_A - \mathbb{1}_{A_2}) \cdots (\mathbb{1}_A - \mathbb{1}_{A_n})(x) \equiv 0,$$

which holds for any $x \in \Omega$. Next we expand this product into a summation of 2^n terms as follows:

$$\mathbb{1}_A + \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|} (\prod_{i \in I} \mathbb{1}_{A_i}) \equiv 0$$

holds for any $x \in \Omega$. Summing over all $x \in \Omega$, this gives that

$$|A| + \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|} |A_I| = 0,$$

which implies that

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = |A| = \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|+1} |A_I| = \sum_{k=1}^{n} (-1)^{k+1} S_k,$$

finishing the proof. ∎

Proof (2). It suffices to prove that

$$\mathbb{1}_{A_1 \cup A_2 \cup \cdots \cup A_n}(x) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{I \in \binom{[n]}{k}} \mathbb{1}_{A_I}(x)$$

holds for all $x \in \Omega$. Denote by LHS (resp. RHS) the left-hand side (resp. right-hand side) of the above equation.

Assume that x is contained in exactly ℓ subsets, say $A_1, A_2, \ldots, A_\ell$. If $\ell = 0$, then clearly $LHS = 0 = RHS$, so we are done. So we may assume that $\ell \geq 1$. In this case, we have $LHS = 1$ and

$$RHS = \ell - \binom{\ell}{2} + \binom{\ell}{3} + \cdots + (-1)^{\ell+1} \binom{\ell}{\ell} = 1.$$

Note that the above equation holds since $\sum_{i=0}^{\ell} (-1)^i \binom{\ell}{i} = (1-1)^\ell = 0$. This finishes the proof. ∎

Next, we will demonstrate the power of Inclusion-Exclusion formula by using it to solve several problems.

Definition 1.30. *Let $\varphi(n)$ be the number of integers $m \in [n]$ which are relatively prime*[1] *to n.*

Theorem 1.31. *If we express $n = p_1^{a_1} p_2^{a_2} \cdots p_t^{a_t}$, where $p_1, \ldots, p_t$ are distinct primes, then*

$$\varphi(n) = n \prod_{i=1}^{t} (1 - \frac{1}{p_i}).$$

Proof. Let the ground set

$$\Omega = [n]$$

and

$$A_i = \{m \in [n] : p_i | m\}$$

for $i \in \{1, 2, \ldots, t\}$. It implies

$$\varphi(n) = \big|\{m \in [n] : m \notin A_i \text{ for all } i \in [t]\}\big| = \big|[n] \backslash (A_1 \cup A_2 \cup \cdots \cup A_t)\big|.$$

By Inclusion-Exclusion formula,

$$\varphi(n) = \sum_{I \subseteq [t]} (-1)^{|I|} |A_I|,$$

[1] Here, "m is relatively prime to n" means that the greatest common divisor of m and n is 1.

where $A_I = \cap_{i\in I} A_i = \{m \in [n] : (\prod_{i\in I} p_i) | m\}$ and thus $|A_I| = \frac{n}{\prod_{i\in I} p_i}$. We can derive that

$$\varphi(n) = \sum_{I\subseteq[t]} (-1)^{|I|} \frac{n}{\prod_{i\in I} p_i} = n(1-\frac{1}{p_1})(1-\frac{1}{p_2})\cdots(1-\frac{1}{p_t}),$$

as desired.

∎

Exercise 1.32. *For any positive integer n,*

$$\sum_{d|n} \varphi(d) = n.$$

1.5 Möbius Inversion Formula

Definition 1.33. *The* **Möbius Function** μ *for a positive integer d is*

$$\mu(d) = \begin{cases} 1, & d \text{ is a product of even number of distinct primes } (d=1 \text{ included}) \\ -1, & d \text{ is a product of odd number of distinct primes} \\ 0, & \text{otherwise} \end{cases}$$

Theorem 1.34. *For any positive integer n,*

$$\sum_{d|n} \mu(d) = \begin{cases} 1, & n = 1 \\ 0, & \text{otherwise} \end{cases}$$

Proof. If $n = 1$, it is trivial. For $n = p_1^{a_1} \dots p_r^{a_r} \geq 2$,

$$\sum_{d|n} \mu(d) = \sum_{i_1 \leq a_1,\dots,i_r \leq a_r} \mu(p_1^{i_1} \dots p_r^{i_r}) = \sum_{i=0}^{r} \binom{r}{i} (-1)^i = 0.$$

∎

Theorem 1.35 (**Möbius Inversion Formula**)**.** *Let $f(n)$ and $g(n)$ be two functions defined for every positive integer n satisfying*

$$f(n) = \sum_{d|n} g(d).$$

Then we have

$$g(n) = \sum_{d|n} \mu(d) f(\frac{n}{d}).$$

Proof.

$$\begin{aligned}\sum_{d|n}\mu(d)f(\frac{n}{d}) &= \sum_{d|n}\mu(\frac{n}{d})f(d)\\ &= \sum_{d|n}\mu(\frac{n}{d})(\sum_{d'|d}g(d'))\\ &= \sum_{d'|n}g(d')\sum_{\frac{n}{d}|\frac{n}{d'}}\mu(\frac{n}{d})\\ &= \sum_{d'|n}g(d')\sum_{m|\frac{n}{d'}}\mu(m)\\ &= \sum_{d'|n,d'\neq n}g(d')\times 0+g(n)\times 1\\ &= g(n)\end{aligned}$$

as desired. ∎

1.6 Generating Functions

Definition 1.36. *The (ordinary)* **generating function (GF)** *for an infinite sequence* $\{a_0, a_1, \dots\}$ *is a power series*

$$f(x)=\sum_{n\geq 0}a_nx^n.$$

We have two ways to view this power series.

(i). When the power series $\sum_{n\geq 0}a_nx^n$ converges (i.e. there exists a radius $R>0$ of convergence), we view GF as a function of x and we can apply operations of calculus on it (including derivation and integration). For example, we know that

$$a_n=\frac{f^{(n)}(0)}{n!}.$$

Recall the following sufficient condition on the radius of convergence that if $|a_n|\leq K^n$ for some $K>0$, then $\sum_{n\geq 0}a_nx^n$ converges in the interval $(-\frac{1}{K},\frac{1}{K})$.

(ii). When we are not sure of the convergence, we view the generating function as a formal series and take additions and multiplications. Let $a(x)=\sum_{n\geq 0}a_nx^n$ and $b(x)=\sum_{n\geq 0}b_nx^n$.

Addition.

$$a(x)+b(x)=\sum_{n\geq 0}(a_n+b_n)x^n.$$

Multiplication. Let $c_n=\sum_{i=0}^{n}a_ib_{n-i}$. Then

$$a(x)\cdot b(x)=\sum_{n\geq 0}c_nx^n.$$

Example 1.37. *Consider the GF of* $\{1,1,1,\dots\}$. *We note* $\frac{1}{1-x}=\sum_{n=0}^{\infty}x^n$ *holds for all* $-1<x<1$. *From the point view of (i), its first derivative gives*

$$\frac{1}{(1-x)^2}=\sum_{n=1}^{\infty}nx^{n-1}=\sum_{n=0}^{\infty}(n+1)x^n.$$

So we could view $\frac{1}{(1-x)^2}$ as the GF of $\{1,2,3,\dots\}$ for all $-1<x<1$.

Problem 1.38. *Let $a_0=1$ and $a_n=2a_{n-1}$ for $n\ge 1$. Find a_n.*

Solution. Consider the generating function,

$$f(x)=\sum_{n=0}^{\infty}a_nx^n=1+\sum_{n=1}^{\infty}a_nx^n=1+2x\sum_{n=1}^{\infty}a_{n-1}x^{n-1}=1+2xf(x).$$

So $f(x)=\frac{1}{1-2x}$, which implies that $f(x)=\sum_{n=0}^{+\infty}2^nx^n$ and $a_n=2^n$. ∎

From this problem, we see one of the basic ideas for using generating function: in order to find the general expression of a_n, we work on its generating function $f(x)$; once we find the formula of $f(x)$, then we can expand $f(x)$ into a power series and get a_n by choosing the coefficient of the right term.

Problem 1.39. *Let A_n be the set of strings of length n with entries from the set $\{a,b,c\}$ and with no "aa" occuring (in the consecutive positions). Find $|A_n|$ for $n\ge 1$.*

Solution. Let $a_n=|A_n|$. We first observe that $a_1=3, a_2=8$. For $n\ge 3$, we will find a_n by recursion as follows. If the first string is 'a', the second string has two choices, 'b' or 'c'. Then the last $n-2$ strings have a_{n-2} choices. If the first string is 'b' or 'c', the last $n-1$ strings have a_{n-1} choices. They are all different. Totally, for $n\ge 3$, we have

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

Set $a_0=1$, then $a_n=2a_{n-1}+2a_{n-2}$ holds for $n\ge 2$. The generating function of $\{a_n\}$ is

$$f(x)=\sum_{n\ge 0}a_nx^n=a_0+a_1x+\sum_{n\ge 2}(2a_{n-1}+2a_{n-2})x^n=1+3x+2x(f(x)-1)+2x^2f(x),$$

which implies that

$$f(x)=\frac{1+x}{1-2x-2x^2}.$$

By Partial Fraction Decomposition, we calculate that

$$f(x)=\frac{1-\sqrt{3}}{2\sqrt{3}}\frac{1}{\sqrt{3}+1+2x}+\frac{1+\sqrt{3}}{2\sqrt{3}}\frac{1}{\sqrt{3}-1-2x},$$

which implies that

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

∎

Remark 1.40. *Note that a_n must be an integer but its expression is a combination of irrational terms! Observe that $\left|\frac{-2}{\sqrt{3}+1}\right|<1$, so $\left(\frac{-2}{\sqrt{3}+1}\right)^n\to 0$ as $n\to\infty$. Thus, when n is sufficiently large, this integer a_n is about the value of the second term $\frac{1+\sqrt{3}}{2\sqrt{3}}\frac{1}{\sqrt{3}-1}\left(\frac{2}{\sqrt{3}-1}\right)^n$. Equivalently a_n will be the nearest integer to that.*

Exercise 1.41. *Define Fibonacci number F_n as follows: $F_1 = 0, F_1 = 1, F_{n+2} = F_{n+1} + F_n$ for all $n \geq 0$. Find F_n.*

Definition 1.42. *For any real r and an integer $k \geq 0$, let*

$$\binom{r}{k} = \frac{r(r-1)\cdots(r-k+1)}{k!}.$$

Exercise 1.43. *Prove that $\binom{\frac{1}{2}}{k} = \frac{(-1)^{k-1}\cdot 2}{4^k}\frac{(2k-2)!}{k!(k-1)!}$.*

Theorem 1.44 (Newton's Binomial Theorem)**.** *For any real number r and $x \in (-1, 1)$,*

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

Proof. By Taylor series, it is obvious. ∎

Corollary 1.45. *Let $r = -n$ for some integer $n \geq 0$. Then*

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

Therefore

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

which is equivalent to

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

Noting that

$$\binom{n+k-1}{k} = \#\text{ integer solutions to } x_1 + x_2 + \cdots + x_n = k \text{ where } x_i \geq 0, 1 \leq i \leq n,$$

we can explain Equation (3.21) from another point of view as follows.

Recall the following facts.

Fact 1.46. *For $j \in [n]$, let $f_j(x) := \sum_{i \in I_j} x^i$, where $I_j \subset \mathbb{N}$. Let b_k be the number of solutions to $i_1 + i_2 + \cdots + i_n = k$ for $i_j \in I_j$. Then*

$$\prod_{j=1}^{n} f_j(x) = \sum_{k=0}^{\infty} b_k x^k.$$

Fact 1.47. *If $f(x) = \prod_{i=1}^{k} f_i(x)$ for polynomials $f_1, ..., f_k$, then*

$$[x^n]f = \sum_{i_1+i_2+\cdots+i_k=n} \prod_{j=1}^{k} \left([x^{i_j}]f_j\right),$$

where $[x^n]f$ is the coefficient of x^n in f.

Let $f_j = (1-x)^{-1} = \sum_{i\geq 0} x^i$, $\forall j \in [n]$. By Fact 1.46, we can get Equation 3.21 by considering as $(1-x)^{-n} = \prod_{j=1}^{n} f_j$ easily.

Exercise 1.48. *Show* $(1-x)^{-n} = \sum_{k=0}^{\infty} \binom{n+k-1}{k} x^k$ *by taking the* n^{th} *derivative of* $(1-x)^{-1}$.

Problem 1.49. *Let* a_n *be the number of ways to pay* n *Yuan using 1-Yuan bills, 2-Yuan bills and 5-Yuan bills. What is the generating function of this sequence* $\{a_n\}$*?*

Solution. Observe that a_n is the number of integer solutions (i_1, i_2, i_3) to $i_1 + i_2 + i_3 = n$, where $i_1 \in I_1 := \{0, 1, 2, ...\}$, $i_2 \in I_2 := \{0, 2, 4, ...\}$ and $i_3 \in I_3 := \{0, 5, 10, ...\}$. Let $f_j(x) := \sum_{m \in I_j} x^m$ for $j = 1, 2, 3$. By Fact 1.46, we have

$$\sum_{n=0}^{+\infty} a_n x^n = f_1(x) f_2(x) f_3(x) = \frac{1}{1-x} \cdot \frac{1}{1-x^2} \cdot \frac{1}{1-x^5}.$$

∎

1.7 Random Walks

Consider a real axis with integer points $(0, \pm 1, \pm 2, \pm 3, \ldots)$ marked. A frog leaps among the integer points according to the following rules:

(1). At beginning, it sits at 1.

(2). In each coming step, the frog leaps either by distance 2 to the right (from i to $i+2$), or by distance 1 to the left (from i to $i-1$), each of which is randomly chosen with probability $\frac{1}{2}$ independently of each other.

Problem 1.50. *What is the probability that the frog can reach "0"?*

Solution. In each step, we use "+" or "−" to indicate the choice of the frog that is either to leap right or leap left. Then the probability space Ω can be viewed as the set of infinite vectors, where each entry is in $\{+, -\}$.

Let A be the event that the frog reaches "0". Let A_i be the event that the frog reaches "0" at the i^{th} step for the first time. So $A = \cup_{i=1}^{+\infty} A_i$ is a disjoint union. So $P(A) = \sum_{i=1}^{+\infty} P(A_i)$.

To compute $P(A_i)$, we can define a_i to be the number of trajectories (or vectors) of the first i steps such that the frog starts at "1" and reaches "0" at the i^{th} step for the first time. So

$$P(A_i) = \frac{a_i}{2^i}.$$

Then,

$$P(A) = \sum_{i=1}^{+\infty} \frac{a_i}{2^i}.$$

Let $f(x) = \sum_{i=0}^{+\infty} a_i x^i$ be the generating function of $\{a_i\}_{i\geq 0}$, where $a_0 := 0$. Thus,

$$P(A) = \sum_{i=1}^{+\infty} \frac{a_i}{2^i} = f\left(\frac{1}{2}\right).$$

We then turn to find the expression of $f(x)$.

Let b_i be the number of trajectories of the first i steps such that the frog starts at "2" and reaches "0" at the i^{th} step for the first time.

Let c_i be the number of trajectories of the first i steps such that the frog starts at "3" and reaches "0" at the i^{th} step for the first time.

First we express b_i in terms of $\{a_j\}_{j\geq 1}$. Since the frog only can leap to left by distance 1, if the frog can successfully jump from "i" to "0" in i steps, then this frog must reach "1" first. Let j be the number of steps by which the frog reaches "1" for the first time. So there are a_j trajectories from "2" to "1" at the j^{th} step for the first time. In the remaining $i-j$ steps the frog must jump from "1" to "0" and reach "0" at the coming $(i-j)^{th}$ step for the first time, so there are a_{i-j} trajectories that the frog can finish in exactly $i-j$ steps. In total,

$$b_i = \sum_{j=1}^{i-1} a_j a_{i-j}.$$

As $a_0 = 0$,

$$b_i = \sum_{j=0}^{i} a_j a_{i-j}.$$

We can get

$$\sum_{i\geq 0} b_i x^i = (\sum_{i\geq 0} a_i x^i)^2 = f^2(x).$$

Similarly, if we count the number c_i of trajectories from 3 to 0, we can obtain that

$$c_i = \sum_{j=0}^{i} a_j b_{i-j},$$

which implies that

$$\sum_{i\geq 0} c_i x^i = \left(\sum_{i\geq 0} b_i x^i\right)\left(\sum_{i\geq 0} a_i x^i\right) = f^3(x).$$

Let us consider a_i from another point of view. After the first step, either the frog reaches "0" directly (if it leaps to left, so $a_1 = 1$), or it leaps to "3". In the latter case, the frog needs to jump from "3" to "0" using $i-1$ steps. Thus for $i \geq 2$, $a_i = c_{i-1}$.

Combining the above facts, we have

$$f(x) = \sum_{i=0}^{+\infty} a_i x^i = x + \sum_{i\geq 2} a_i x^i = x + \sum_{i\geq 2} c_{i-1} x^i = x + x\left(\sum_{j=0}^{+\infty} c_j x^j\right) = x + x \cdot f^3(x).$$

Let $a := P(A) = f(1/2)$. Then we have $a = \frac{1}{2} + \frac{a^3}{2}$, i.e., $(a-1)(a^2+a-1) = 0$, implying that

$$a = 1,\ \frac{\sqrt{5}-1}{2} \text{ or } \frac{-\sqrt{5}-1}{1}.$$

Since $P(A) \in [0,1]$, we see $P(A) = 1$ or $\frac{\sqrt{5}-1}{2}$.

Note that $f(x) = x + xf^3(x)$. Consider the inverse function of $f(x)$, that is, $g(x) := \frac{x}{1+x^3}$. Consider the figure of $g(x)$. We find that $g(x)$ is increasing around $\frac{\sqrt{5}-1}{2}$ but decreasing around 1. Since $f(x) = \sum a_i x^i$ is increasing, $g(x)$ also increases. Thus it doesn't make sense for $g(x)$ being around $x = 1$. This explains that $P(A) = \frac{\sqrt{5}-1}{2}$, which is the golden section! ∎

1.8 Exponential Generating Functions

Let $\mathbb{N}, \mathbb{N}_e$ and $\mathbb{N}_o$ be the sets of non-negative integers, non-negative even integers and non-negative odd integers, respectively.

Given n sets I_j of non-negative integers for $j \in [n]$, let $f_j(x) = \sum_{i \in I_j} x^i$. Let a_k be the number of integer solutions to $i_1 + i_2 + \cdots + i_n = k$, where $i_j \in I_j$. Then $\prod_{j=1}^n f_j(x)$ is the ordinary generating function of $\{a_k\}_{k \geq 0}$.

Problem 1.51. *Let S_n be the number of selections of n letters chosen from an unlimited supply of a's, b's and c's such that both of the numbers of a's and b's are even.*

Solution. We can write S_n as

$$S_n = \sum_{e_1+e_2+e_3=n,\ e_1,e_2 \in \mathbb{N}_e,\ e_3 \in \mathbb{N}} 1.$$

Using the previous fact, we see that $S_n = [x^n]f$, where

$$f(x) = \left(\sum_{i \in \mathbb{N}_e} x^i\right)^2 \left(\sum_{j \in \mathbb{N}} x^j\right) = \left(\frac{1}{1-x^2}\right)^2 \cdot \frac{1}{1-x}.$$

∎

Problem 1.52. *Let T_n be the number of arrangements (or words) of n letters chosen from an unlimited supply of a's, b's and c's such that both of the numbers of a's and b's are even. What is the value of T_n?*

Solution. To solve this, we define a new kind of generating functions.

Definition 1.53. *The exponential generating function for the sequence $\{a_n\}_{n \geq 0}$ is the power series*

$$f(x) = \sum_{n=0}^{\infty} a_n \cdot \frac{x^n}{n!}.$$

Then we have the following fact.

Fact 1.54. *If we have n letters including x a's, y b's and z c's (i.e. $x + y + z = n$), then we can form $\frac{n!}{x!y!z!}$ distinct words using them.*

Therefore, a selection (say x a's, y b's and z c's) can contribute $\frac{n!}{x!y!z!}$ arrangements to T_n. This implies that

$$T_n = \sum_{e_1+e_2+e_3=n,\ e_1,e_2 \in \mathbb{N}_e,\ e_3 \in \mathbb{N}} \frac{n!}{e_1!e_2!e_3!}.$$

Similar to defining the above $f(x)$ for S_n, we define the following for T_n. Let

$$g(x) := \left(\sum_{i\in\mathbb{N}_e} \frac{x^i}{i!}\right)^2 \left(\sum_{j\in\mathbb{N}} \frac{x^j}{j!}\right).$$

Claim. We have

$$[x^n]g = \frac{T_n}{n!}.$$

Proof. To see this, we expand $g(x)$. Then the term x^n in $g(x)$ becomes

$$\sum_{\substack{e_1+e_2+e_3=n,\\ e_1,e_2\in\mathbb{N}_e,\ e_3\in\mathbb{N}}} \frac{x^{e_1}}{e_1!}\cdot\frac{x^{e_2}}{e_2!}\cdot\frac{x^{e_3}}{e_3!} = \left(\sum_{\substack{e_1+e_2+e_3=n,\\ e_1,e_2\in\mathbb{N}_e,\ e_3\in\mathbb{N}}} \frac{n!}{e_1!e_2!e_3!}\right)\frac{x^n}{n!} = T_n\cdot\frac{x^n}{n!}.$$

So $[x^n]g = \frac{T_n}{n!}$, i.e., $g(x)$ is the exponential generating function of $\{T_n\}$. This finishes the proof of Claim. ∎

Using Taylor series: $e^x = \sum_{j\geq 0}\frac{x^j}{j!}$ and $e^{-x} = \sum_{j\geq 0}(-1)^j\frac{x^j}{j!}$, we have

$$\frac{e^x+e^{-x}}{2} = \sum_{j\in\mathbb{N}_e}\frac{x^j}{j!} \quad\text{and}\quad \frac{e^x-e^{-x}}{2} = \sum_{j\in\mathbb{N}_o}\frac{x^j}{j!}.$$

By the previous fact, we get

$$g(x) = \left(\frac{e^x+e^{-x}}{2}\right)^2\cdot e^x = \frac{e^{3x}+2e^x+e^{-x}}{4} = \sum_{n\geq 0}\left(\frac{3^n+2+(-1)^n}{4}\right)\cdot\frac{x^n}{n!}.$$

Therefore, we get that

$$T_n = \frac{3^n+2+(-1)^n}{4}.$$

∎

Recall that the *exponential generating function* for the sequence $\{a_n\}_{n\geq 0}$ is the power series

$$f(x) = \sum_{n=0}^{+\infty} a_n\cdot\frac{x^n}{n!}.$$

As we shall see, ordinary generation functions can be used to find the number of selections; while exponential generation functions can be used to find the number of arrangements or some combinatorial objects **involving ordering**. We summarize this as the following facts.

Fact 1.55. *Given $I_j\subseteq\mathbb{N}$ for $j\in[n]$, let $f_j(x) = \sum_{i\in I_j} x^i$. And let $a_k = \sum_{\substack{i_1+\cdots+i_n=k,\\ i_j\in I_j}} 1$. Then*

$$\prod_{j=1}^{n} f_j(x) = \sum_{k=0}^{+\infty} a_k x^k.$$

Fact 1.56. *Given $I_j \subseteq \mathbb{N}$ for $j \in [n]$, let $g_j(x) = \sum_{i \in I_j} \frac{x^i}{i!}$. And let $b_k = \sum_{\substack{i_1+\cdots+i_n=k,\\ i_j \in I_j}} \frac{k!}{i_1!i_2!\cdots i_n!}$. Then*

$$\prod_{j=1}^{n} g_j(x) = \sum_{k=0}^{+\infty} \frac{b_k}{k!} x^k.$$

Fact 1.57. *Let $f(x) = \prod_{j=1}^{n} f_j(x)$. Then*

$$[x^k]f = \sum_{\substack{i_1+\cdots+i_n=k,\\ i_j \geq 0}} \prod_{j=1}^{n} [x^{i_j}] f_j.$$

Fact 1.58. *Let $f(x) = \prod_{j=1}^{n} f_j(x)$ and let $f_j(x) = \sum_{k=0}^{+\infty} \frac{a_k^{(j)}}{k!} x^k$. Then*

$$f(x) = \sum_{k=0}^{+\infty} \frac{A_k}{k!} x^k,$$

if and only if

$$A_k = \sum_{\substack{i_1+\ldots+i_n=k,\\ i_j \geq 0}} \frac{k!}{i_1!i_2!\cdots i_n!} \Big(\prod_{j=1}^{n} a_{i_j}^{(j)} \Big).$$

Exercise 1.59. *Find the number a_n of ways to send n students to four different classes (say R_1, R_2, R_3, R_4) such that each class has at least one student.*

Solution.

$$a_n = \sum_{\substack{i_1+i_2+i_3+i_4=n,\\ i_j \geq 1}} \frac{n!}{i_1!i_2!i_3!i_4!}.$$

Let $I_j \subseteq \mathbb{N}$ for $j \in [4]$ and $g_j(x) = \sum_{i \geq 1} \frac{x^i}{i!} = e^x - 1$. By Fact 1.56, we have that

$$\sum_{n=0}^{+\infty} \frac{a_n}{n!} x^n = g_1 g_2 g_3 g_4 = (\sum_{i \geq 1} \frac{x^i}{i!})^4 = (e^x - 1)^4 = e^{4x} - 4e^{3x} + 6e^{2x} - 4e^x + 1 = \sum_{n=0}^{+\infty} (4^n - 4 \cdot 3^n + 6 \cdot 2^n - 4) \frac{x_n^n}{n!} + 1.$$

Thus $a_n = 4^n - 4 \cdot 3^n + 6 \cdot 2^n - 4$ for $n \geq 4$. ∎

Exercise 1.60. *Let a_n be the number of arrangements of type A for a group of n people, and let b_n be the number of arrangements of type B for a group of n people.*

Define a new arrangement of n people called type C as follows:

- *Divide the n people into 2 groups (say 1^{st} and 2^{nd}).*
- *Then arrange the 1^{st} group by an arrangement of type A, and arrange the 2^{nd} group by an arrangement of type B.*

Let c_n be the number of arrangements of type C of n people. Let $A(x), B(x), C(x)$ be the exponential generation function for $\{a_n\}, \{b_n\}, \{c_n\}$ respectively. Prove that $C(x) = A(x)B(x)$.

Proof. We can easily see that

$$c_n = \sum_{\substack{i+j=n, \\ i,j \geq 0}} \frac{n!}{i!j!} a_i b_j.$$

Then by Fact 1.58, $C(x) = A(x)B(x)$. ∎

Exercise 1.61. *Recall that $S(n,k) \cdot k!$ is equal to the number of surjections from $[n]$ to $[k]$. For fixed k, compute the exponential generating function of $S(n,k) \cdot k!$. Then find the value of $S(n,k) \cdot k!$.*

Theorem 1.62 (Lagrange Inversion Formula). *Let $f(x)$ be analytic (convergent power series) in a neighborhood of $z = 0$ and $f(0) \neq 0$. If $w = \frac{z}{f(z)}$, then z can be expressed as a power series*

$$z = \sum_{k=1}^{\infty} c_k w^k$$

with a positive radius of convergence, where

$$c_k = \frac{1}{k!} \left\{ \left(\frac{\mathrm{d}}{\mathrm{d}z} \right)^{k-1} (f(z))^k \right\}_{z=0}.$$

2 Basics of Graphs

In this second part of our course, we will introduce some basic definitions about graphs.

Definition 2.1. *A graph $G = (V, E)$ consists of a vertex set V and an edge set E, where the elements of V are called* **vertices** *and the elements of $E \subseteq \binom{V}{2} = \{\{x, y\} : x, y \in V\}$ are called* **edges**.

This provides the definition of a simple undirected graph. The word "undirected" means that the edge set E contains unordered pairs. Otherwise, G is called a directed graph. A graph is *simple* if it has no loops or multiple edges. A *loop* is an edge whose endpoints are equal. *Multiple edges* are edges having the same pair of endpoints.

- We say vertices x and y are *adjacent* if $\{x, y\} \in E$, write $x \sim_G y$ or $x \sim y$ or $xy \in E$.
- We say the edge xy is *incident* to the endpoints x and y.
- Let $e(G)$ be the number of edges in G, i.e., $e(G) = |E(G)|$.
- The *degree* of a vertex v in G, denoted by $d_G(v)$, is the number of edges in G incident to v.
- The *neighborhood* of a vertex v is the set of vertices that are adjacent to v, i.e., $N_G(v) = \{u \in V(G) : u \sim v\}$. Thus we have $d_G(v) = |N_G(v)|$.
- A graph $G' = (V', E')$ is a *subgraph* of $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E \cap \binom{V'}{2}$, i.e., $G' \subseteq G$.
- A subgraph $G' = (V', E')$ of $G = (V, E)$ is *induced*, if $E' = E \cap \binom{V'}{2}$, write $G' = G[V']$.

Definition 2.2. *Two graphs $G = (V, E)$ and $G' = (V', E')$ are* **isomorphic** *if there exists a bijection $f : V \to V'$ such that $i \sim_G j$ if and only if $f(i) \sim_{G'} f(j)$.*

- A graph on n vertices is a *complete graph* (or a *clique*), denoted by K_n, if all pairs of vertices are adjacent. So we have $e(K_n) = \binom{n}{2}$.
- A graph on n vertices is called an *independent set*, denoted by I_n, if it contains no edge at all.
- Given a graph $G = (V, E)$, its *complement* is a graph $\overline{G} = (V, E^c)$ with $E^c = \binom{V}{2} \backslash E$.
- The *degree sequence* of a graph $G = (V, E)$ is a sequence of degrees of all vertices listed in a non-decreasing order.
- The *path* P_k of length $k - 1$ is a graph $v_1 v_2 ... v_k$ where $v_i \sim v_{i+1}$ for $i \in [k - 1]$ and $v_j \neq v_l$ for any $j \neq l \in [k]$. Note that the length of a path P (denoted by $|P|$) is the number of edges in P.
- A *cycle* C_k of length k is a graph $v_1 v_2 ... v_k v_1$ where $v_i \sim v_{i+1}$ for $i \in [k]$, $v_{k+1} = v_1$, and $v_j \neq v_l$ for any $j \neq l \in [k]$.
- Let G be a simple graph with vertex set $V(G) = \{v_1, \ldots, v_n\}$ and edge set $E(G) = \{e_1, \ldots, e_m\}$. The *adjacency matrix* of G, denoted by $A(G)$, is the n-by-n matrix in which entry $a_{i,j}$ is the number of edges in G with endpoints $\{v_i, v_j\}$. The *incidence matrix* $M(G)$ is the n-by-m matrix in which entry $m_{i,j}$ is 1 if v_j is an endpoint of e_j and 0 otherwise.
- A graph G is *planar*, if we can draw G on the plane such that its edges intersect only at their endpoints.

Theorem 2.3 (Euler's Formula)**.** *Let $G = (V, E)$ be a connected planar graph with v vertices and e edges, and let r be the number of regions in which some given embedding of G divides the plane. Then $v - e + r = 2$.*

Exercise 2.4. *Show that K_4 is planar but K_5 is not.*

Exercise 2.5. *Show that $K_{3,3}$ is not planar.*

The following Handshaking Lemma is the most basic lemma in graph theory.

Lemma 2.6 (Handshaking Lemma)**.** *In any graph $G = (V, E)$,*

$$\sum_{v \in V} d_G(v) = 2e(G).$$

Proof. Let $F = \{(e, v) : e \in E(G), v \in V(G)$ such that v is incident to $e\}$. Then

$$\sum_{e \in E(G)} 2 = |F| = \sum_{v \in V} d_G(v).$$

∎

Corollary 2.7. *In any graph G, the number of vertices with odd degree is even.*

Proof. Let $O = \{v \in V(G) : d(v)$ is odd$\}$ and $\mathcal{E} = \{v \in V(G) : d(v)$ is even$\}$. Then by Lemma 2.6,

$$2e(G) = \sum_{v \in O} d_G(v) + \sum_{v \in \mathcal{E}} d_G(v).$$

Thus we have $\sum_{v \in O} d_G(v)$ is even, moreover we have $|O|$ is even. ∎

Corollary 2.8. *In any graph G, if there exists a vertex with odd degree, then there are at least two vertices with odd degree.*

3 Double-counting

3.1 Basics

The basic setting of the double counting technique is as follows. Suppose that we are given two finite sets A and B, and a subset $S \subseteq A \times B$. If $(a, b) \in S$, then we say that a and b are incident. Let N_a be the number of elements $b \in B$ such that $(a, b) \in S$, and N_b be the number of elements $a \in A$ such that $(a, b) \in S$. Then we have

$$\sum_{a \in A} N_a = |S| = \sum_{b \in B} N_b.$$

Theorem 3.1. *Let $T(j)$ be the number of divisions of a positive integer j. Let $\overline{T(n)} = \frac{1}{n}\sum_{j=1}^{n} T(j)$. Then we have $|\overline{T(n)} - H(n)| < 1$, where $H(n) = \sum_{i=1}^{n} \frac{1}{i}$ is the n^{th} Harmonic number.*

Proof. Define a table $X = (x_{ij})$ where

$$x_{ij} = \begin{cases} 1, & if\ i|j \\ 0, & otherwise. \end{cases}$$

Then

$$\sum_{j=1}^{n} T(j) = \sum_{1 \le i \le j \le n} x_{ij} = \sum_{i=1}^{n} \lfloor \frac{n}{i} \rfloor,$$

which implies that

$$\overline{T(n)} = \frac{1}{n} \sum_{i=1}^{n} \lfloor \frac{n}{i} \rfloor.$$

Then we have

$$|\overline{T(n)} - H(n)| < 1.$$

∎

Exercise 3.2. *Prove that*

$$\left| \frac{1}{n} \sum_{i=1}^{n} \lfloor \frac{n}{i} \rfloor - \sum_{i=1}^{n} \frac{1}{i} \right| < 1.$$

3.2 Sperner's Theorem

Definition 3.3. *Let $\mathcal{F} \subseteq 2^{[n]}$ be a family of subsets of [n]. We say $\mathcal{F}$ is* **independent** *(or $\mathcal{F}$ is an* **independent system***), if for any two $A, B \in \mathcal{F}$, we have $A \not\subset B$ and $B \not\subset A$. In other words, $\mathcal{F}$ is independent if and only if there is no "containment" relationship between any two subsets of $\mathcal{F}$.*

Fact 3.4. *For a fixed $k \in [n]$, $\binom{[n]}{k}$ is an independent system.*

Theorem 3.5 (Sperner's Theorem)**.** *For any independent system $\mathcal{F}$ of [n], we have*

$$|\mathcal{F}| \le \binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

First we define a chain.

Definition 3.6. *A chain of subsets of* $[n]$ *is a sequence of distinct subsets such that*

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots \subseteq A_k.$$

First proof of Sperner's Theorem (Double-Counting). A *maximal chain* is a chain with the property that no other subsets of $[n]$ can be inserted into it to find a longer chain. We have the following observations.

(1). Any maximal chain looks like:

$$\phi \subseteq \{x_1\} \subseteq \{x_1, x_2\} \subseteq \cdots \subseteq \{x_1, ..., x_k\} \subseteq \cdots \subseteq \{x_1, ..., x_n\}.$$

(2). There are exactly $n!$ maximal chains.

This is because any such a maximal chain, say $\mathcal{C} : \phi \subseteq \{x_1\} \subseteq \{x_1, x_2\} \subseteq \cdots \subseteq \{x_1, x_2, ..., x_n\}$, defines a unique permutation:

$$\pi : [n] \to [n], \pi(i) = x_i, \forall i \in [n].$$

Now we count the number of pairs $(\mathcal{C}, A)$ satisfying that:

- $\mathcal{C}$ is a maximal chain of $[n]$.
- $A \in \mathcal{C} \cap \mathcal{F}$.

Recall the rule of double counting given at the beginning that

$$\sum_{\mathcal{C}} N_{\mathcal{C}} = \text{the number of pairs } (\mathcal{C}, A) = \sum_{A} N_A,$$

where $N_{\mathcal{C}}$ is the number of subsets $A \in \mathcal{C} \cap \mathcal{F}$ and N_A is the number of maximal chains $\mathcal{C}$ containing A. It is key to observe that

- $N_{\mathcal{C}} \leq 1$,
- $N_A = |A|!(n - |A|)!$

So we have

$$\begin{aligned} n! = \sum_{\mathcal{C}} 1 \geq \sum_{\mathcal{C}} N_{\mathcal{C}} &= \sum_{A \in \mathcal{F}} N_A = \sum_{A \in \mathcal{F}} |A|!(n - |A|)! \\ &= \sum_{A \in \mathcal{F}} \frac{n!}{\binom{n}{|A|}} \geq \sum_{A \in \mathcal{F}} \frac{n!}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} = \frac{n!}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} |\mathcal{F}|, \end{aligned}$$

which implies that

$$|\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

This finishes the proof. ∎

Now we give another proof of Sperner's Theorem.

Definition 3.7. *A chain is* <u>*symmetric*</u> *if it consists of subsets of sizes* $k, k+1, ..., \lfloor \frac{n}{2} \rfloor, ..., n-k-1, n-k$ *for some* $k \geq 0$.

For example, when $n = 3$, $\{\{2\}, \{2,3\}, \{1,2,3\}\}$ is not symmetric. And when $n = 4$, $\{\phi, \{1,2,3\}\}$ is not symmetric.

Theorem 3.8. *The family* $2^{[n]}$ *can be partitioned into a disjoint union of symmetric chains.*

First proof of Theorem 3.8. We prove by induction on n.

The base case is $n = 1$. The family $2^{[n]} = 2^{[1]} = \{\emptyset, \{1\}\}$, which itself is a symmetric chain. Thus this theorem is true for $n = 1$.

Now we may assume that $2^{[n]}$ can be partitioned into a disjoint union of symmetric chains $e_1, e_2, \ldots, e_t$. Consider $2^{[n+1]}$, For any

$$e_i = \{P_k \subseteq P_{k+1} \subseteq \cdots \subseteq P_{n-k}\},$$

define two new symmetric chains for $2^{[n+1]}$:

$$e_i' = \{P_{k+1} \subseteq P_{k+2} \subseteq \cdots \subseteq P_{n-k}\},$$

and

$$e_i'' = \{P_k \subseteq (P_k \cup \{n+1\}) \subseteq (P_{k+1} \cup \{n+1\}) \subseteq \cdots \subseteq (P_{n-k} \cup \{n+1\})\}.$$

We assert that $\cup_i \{e_i', e_i''\}$ is a disjoint union of symmetric chain for $2^{[n+1]}$. ∎

Exercise 3.9. *Prove that* $\cup_i \{e_i', e_i''\}$ *is a disjoint union of symmetric chain for* $2^{[n+1]}$.

Second proof of Theorem 3.8. For each $A \in 2^{[n]}$, we define a sequence "$a_1 a_2 ... a_n$" consisting of left and right parentheses by defining

$$a_i = \begin{cases} \text{"(", if } & i \in A \\ \text{")", otherwise.} & \end{cases}$$

We then define the <u>"partial pairing of parentheses"</u> as follows:

(1). First, we pair up all pairs "()" of adjoint parentheses.

(2). Then, we delete these already paired parentheses.

(3). Repeat the above process until nothing can be done.

Note that when this process stops, the remaining unpaired parentheses must look like this:

$$))))((((($$

We say two subsets $A, B \in 2^{[n]}$ have <u>the same partial pairing</u>, if the paired parentheses are the same (even in the same positions).

We can define an equivalence "$\sim$" on $2^{[n]}$ by letting $A \sim B$ if and only if A, B have the same partial pairing.

Exercise 3.10. *Each equivalence class indeed forms a symmetric chain.*

Using this fact, now we see that $2^{[n]}$ can be partitioned into disjoint equivalence classes, which are disjoint symmetric chains. This finishes the proof. ∎

Theorem 3.8 can rapidly imply Sperner's Theorem.

Second proof of Sperner's Theorem. Note that by definition, any symmetric chain contains exactly one subset of size $\lfloor \frac{n}{2} \rfloor$. Since there are $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ many subsets of size $\lfloor \frac{n}{2} \rfloor$, by Theorem 3.8, we see that any partition of $2^{[n]}$ into symmetric chains has to consist of exactly $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ symmetric chains. Each symmetric chain can contain at most one subset from $|\mathcal{F}|$ and thus we see $|\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$. ∎

3.3 Littlewood-Offord Problem

Theorem 3.11. *Fix a vector $\vec{a} = (a_1, a_2, ..., a_n)$ with each $|a_i| \geq 1$. Let $S = \{\vec{\epsilon} = (\epsilon_1, \epsilon_2, ..., \epsilon_n) : \epsilon_i \in \{1, -1\}$ and $\vec{\epsilon} \cdot \vec{a} \in (-1, 1)\}$, then $|S| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$.*

Remark: Note that this is tight for many vectors $\vec{a}$.

Proof. For any $\vec{\epsilon} \in S$, define $A_{\vec{\epsilon}} = \{i \in [n] : a_i \epsilon_i > 0\}$. Let $\mathcal{F} = \{A_{\vec{\epsilon}} : \vec{\epsilon} \in S\}$. Then we have

$$|S| = |\mathcal{F}|.$$

Now we claim that $\mathcal{F}$ is an independent system. Suppose for a contradiction that there exist $A_{\vec{\epsilon}_1}, A_{\vec{\epsilon}_2} \in \mathcal{F}$ with $A_{\vec{\epsilon}_1} \subseteq A_{\vec{\epsilon}_2}$. That also says,

$$\begin{cases} \vec{\epsilon}_1 \cdot \vec{a} \in (-1, 1), \\ \vec{\epsilon}_2 \cdot \vec{a} \in (-1, 1), \end{cases}$$

which imply that

$$|\epsilon_1 \cdot \vec{a} - \epsilon_2 \cdot \vec{a}| < 2.$$

By definition, we have

$$\vec{\epsilon}_1 \cdot \vec{a} = \sum_{i \in A_{\vec{\epsilon}_1}} |a_i| - \sum_{i \notin A_{\vec{\epsilon}_1}} |a_i| = 2 \sum_{i \in A_{\vec{\epsilon}_1}} |a_i| - \sum_{i=1}^{n} |a_i|.$$

Since $A_{\vec{\epsilon}_1} \subseteq A_{\vec{\epsilon}_2}$, we also have that

$$\vec{\epsilon}_2 \cdot \vec{a} - \vec{\epsilon}_1 \cdot \vec{a} = 2(\sum_{i \in A_{\vec{\epsilon}_2}} |a_i| - \sum_{j \in A_{\vec{\epsilon}_1}} |a_j|) \geq 2|a_k| \geq 2, \text{ for some } k \in A_{\vec{\epsilon}_2} \setminus A_{\vec{\epsilon}_1}.$$

This is a contradiction. By Sperner's Theorem, we have $|S| = |\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$. This finishes the proof. ∎

3.4 Turán Type Problems

Definition 3.12. *A graph G is* **bipartite** *if its vertex set can be partitioned into two parts (say A and B) such that each edge joints one vertex in A and another in B.*

This is equivalent to say that $V(G)$ can be partitioned into two independent subsets. And we say (A, B) is a bipartition of G. For example, all even cycles C_{2k} are bipartite, while all odd cycles C_{2k+1} are not.

Definition 3.13. *Let $K_{a,b}$ be the* **complete bipartite** *graph with two parts of sizes a and b. This is a bipartite graph with edge set $\{\{i,j\} : i \in A,\ j \in B\}$ where $|A| = a$ and $|B| = b$.*

Definition 3.14. *Given a graph H, we say a graph G is* ***H*-free** *if G dose not contain a copy of H as its subgraph.*

For example, $K_{a,b}$ is K_3-free.

Definition 3.15. *For fixed graph H, let the* **Turán number of *H***, *denoted by* $\mathrm{ex}(n, H)$, *be the maximum number of edges in an n-vertex H-free graph G.*

Theorem 3.16. $\mathrm{ex}(n, C_4) \leqslant \dfrac{n}{4}(1 + \sqrt{4n-3})$.

Proof. Let G be a C_4-free graph with n vertices. We need to show that $e(G) \leqslant \dfrac{n}{4}(1 + \sqrt{4n-3})$. Consider $S = \{(\{u_1, u_2\}, w) : u_1wu_2$ is a path of length 2 in $G\}$. Since G is C_4-free, for fixed $\{u_1, u_2\}$, there is at most one vertex w such that $(\{u_1, u_2\}, w) \in S$. So we have

$$|S| = \sum_{\{u_1,u_2\}} \text{the number of } (\{u_1, u_2\}, w) \in S \leqslant \sum_{\{u_1,u_2\}} 1 = \binom{n}{2}.$$

On the other hand, fixed a vertex w, the number of $\{u_1, u_2\}$ such that $(\{u_1, u_2\}, w) \in S$ exactly equals $\binom{d(w)}{2}$, which implies that

$$|S| = \sum_{w \in V(G)} \binom{d(w)}{2} = \frac{1}{2} \sum_{w \in V(G)} d^2(w) - e(G).$$

Putting the above together, we have

$$\binom{n}{2} \geq |S| = \frac{1}{2} \sum_{w \in V(G)} d^2(w) - e(G).$$

Using Cauchy-Schwarz inequality, we have

$$\frac{n^2 - n}{2} \geq \frac{n}{2} \sum_{w \in V(G)} \frac{d^2(w)}{n} - e(G) \geq \frac{n}{2} \sum_{w \in V(G)} \left(\frac{d(w)}{n}\right)^2 - e(G),$$

which implies that

$$\frac{2e^2(G)}{n} - e(G) \leq \frac{n^2 - n}{2}.$$

Solving it, we can derive easily that $e(G) \leq \frac{n}{4}(1 + \sqrt{4n-3})$. ∎

Exercise 3.17. *Prove that for all positive integer $n \geq 4$, $\mathrm{ex}(n, C_4) < \frac{n}{4}(1+\sqrt{4n-3})$.*

Hint: Look up the Friendship Theorem.

Corollary 3.18. *We have $\mathrm{ex}(n, C_4) \leqslant (\frac{1}{2} + o(n))n^{\frac{3}{2}}$, where $o(n) \to 0$ as $n \to \infty$.*

The upper bound in Corollary 3.18 is asymptotically tight because there is a construction as follows.

Let p be a prime. Let

$$V = (\mathbb{Z}_p \setminus \{0\}) \times \mathbb{Z}_p$$

and

$$E = \{\{(a,b),(c,d)\} : a, c \in \mathbb{Z}_p \setminus \{0\}, b, d \in \mathbb{Z}_p \text{ and } ac = b + d\}.$$

We have $|V| = (p-1)p$ and $d((a,b)) = p-1$, for any $(a,b) \in V$. Thus we have $|E| = \frac{(p-1)^2 p}{2} \sim \frac{|V|^{\frac{3}{2}}}{2}$. Finally we explain that $G = (V, E)$ is C_4-free. For any $(a_1, b_1), (a_2, b_2) \in V$, if there exist a vertex (say (c,d)) which is their common neighbour, (c,d) satisfies the following condition:

$$\begin{cases} a_1 c = b_1 + d \\ a_2 c = b_2 + d. \end{cases}$$

There is no mutiple solution of this equation.

Theorem 3.19 (Kövári-Sós-Turán Theorem)**.**

$$\mathrm{ex}(n, K_{s,t}) \leq \frac{1}{2}(t-1)^{\frac{1}{s}} n^{2-\frac{1}{s}} + \frac{1}{2}(s-1)n$$

for all $t, s \geq 2$.

Proof. Let G be an n-vertices $K_{s,t}$ free graph with $e(G) \geq \frac{1}{2}sn$ (otherwise we are done). We aim to show $e(G) \leq \frac{1}{2}(t-1)^{\frac{1}{s}} n^{2-\frac{1}{s}} + \frac{1}{2}(s-1)n$. We count the number T of s-stars $K_{1,s}$ as follows. On one hand, $T = \sum_{w \in V(G)} \binom{d(w)}{s}$. On the other hand, $T \leq (t-1)\binom{n}{s}$.

We define

$$f(x) = \begin{cases} 0 \quad , & \text{if } x < s, \\ \dbinom{x}{s}, & \text{if } x \geq s. \end{cases}$$

When $x \geq 0$, $f(x)$ is a convex function. Let $d = \frac{2e(G)}{n}$, by Jensen's inequality,

$$\frac{(t-1)\binom{n}{s}}{n} \geq \frac{T}{n} = \frac{1}{n}\sum_w f(d(w)) \geq f(\frac{\sum_w d(w)}{n}) = f(\frac{2e(G)}{n}) \geq \frac{(d-s+1)^s}{s!}.$$

Thus

$$d \leq ((t-1)(n-1)(n-2)\dots(n-s+1))^{\frac{1}{s}} + (s-1) \leq (t-1)^{\frac{1}{s}} n^{1-\frac{1}{s}} + (s-1).$$

Then we have

$$e(G) = \frac{nd}{2} \leq \frac{1}{2}(t-1)^{\frac{1}{s}} n^{2-\frac{1}{s}} + \frac{1}{2}(s-1)n,$$

finishing the proof. ∎

3.5 Sperner's Lemma

Let us consider the following application of Corollary 2.8. First we draw a triangle in the plane, with 3 vertices $A_1A_2A_3$. Then we divide this triangle $\triangle = A_1A_2A_3$ into small triangles such that no triangle can have a vertex inside an edge of any other small triangle. Then we assign 3 colors (say 1,2,3) to all vertices of these triangles, under the following rules.

(1) The vertex A_i is assigned by color i for $i \in [3]$.

(2) All vertices lying on the edge A_iA_j of the large triangle are assigned by the color i or j.

(3) All interior vertices are assigned by any color 1,2,3.

Lemma 3.20 (Sperner's Lemma (a planar version))**.** *For any assignment of colors described as above, there always exists a small triangle whose three vertices are assigned by three colors* $1, 2, 3$.

Proof. Define an auxiliary graph G as follows.

- Its vertices are the faces of small triangles and the outer face. Let z be the vertex representing the outer face.
- Two vertices of G are adjacent, if the two corresponding faces are neighboring faces and the two endpoints of their common edge are colored by 1 and 2.

We consider the degree of any vertex $v \in V(G)\backslash\{z\}$.

(1) If the face of v has NO two endpoints with color 1 and 2, then $d_G(v) = 0$.

(2) If the face of v has 2 endpoints with color 1 and 2, then let k be the color of the third endpoint of this face. If $k \in \{1, 2\}$, then $d_G(v) = 2$. Otherwise $k = 3$, then $d_G(v) = 1$ and the vertices of this triangle are assigned by three different colors 1,2,3.

Thus we have that $d_G(v)$ is odd if and only if $d_G(v) = 1$, and then the face of v has colors 1,2,3. Now we consider $d_G(z)$ and claim that it must be odd. Indeed, the edge of G incident to z obviously have to go across A_1A_2. Consider the sequence of the colors of the endpoints on A_1A_2, from A_1 to A_2. Then $d_G(z)$ equals the number of alternations between 1 and 2 in this sequence. It is easy to check that $d_G(z)$ must be odd. By Corollary 2.8, since the graph G has a vertex z with odd degree, there must be another vertex $v \in V(G)\backslash\{z\}$ with odd degree. Then $d(v) = 1$ and the face of v has colors 1,2,3. ∎

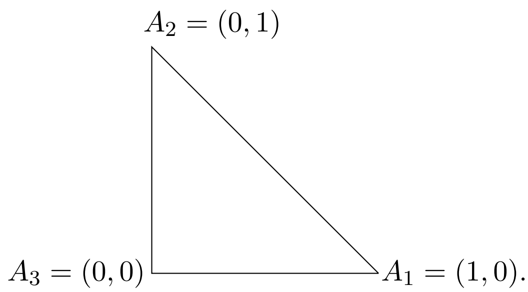
Before we introduce an interesting application of Sperner's lemma, we introduce the following theorem first.

Theorem 3.21 (One-dimensional fixed point theorem)**.** *For any continuous function* $f : [0, 1] \to [0, 1]$*, there exists a point* $x \in [0, 1]$ *such that* $f(x) = x$.

Such an x is called a fixed point of the function f. The theorem can be proved by considering the function $g(x) = f(x) - x$. This is a continuous function with $g(0) \geq 0$ and $g(1) \leq 0$. Intuitively it is quite clear that the graph of such a continuous function can not jump across the x-axis and therefore it has to intersect it, and hence g is 0 at some point of $[0, 1]$. Prove the existence of such a point rigorously requires quite some work. In analysis, this result appears under heading "Darboux theorem".

If we replace the 1-dimensional interval from the Theorem 3.21 by a triangle in the plane, or by a tetrahedron in the 3-dimensional space, or by their analogs in higher dimensions, we will have Brouwer's fixed point theorem. Here we prove only the 2-dimensional version by Spener's lemma.

Let $\triangle$ denote a triangle in the plane. For simplicity, let us take the triangle with vertices $A_1 = (1, 0), A_2 = (0, 1)$, and $A_3 = (0, 0)$:



Theorem 3.22 (Brouwer's Fixed Point theorem in 2-dimension)**.** *Every continuous function $f : \triangle \to \triangle$ has a fixed point x, that is, $f(x) = x$.*

Proof. Define three auxiliary functions $\beta_i : \triangle \to R$ for $i \in \{1, 2, 3\}$ as follows:

For any $a = (x, y) \in \triangle$,

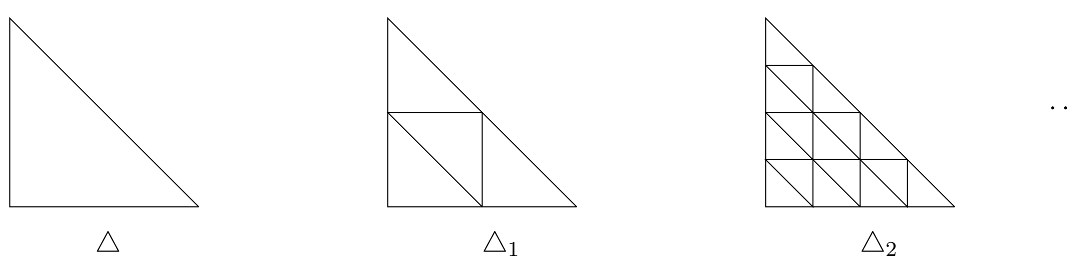
$$\begin{cases} \beta_1(a) &= x, \\ \beta_2(a) &= y, \\ \beta_3(a) &= 1 - x - y. \end{cases}$$

For any continuous $f : \triangle \to \triangle$, define $M_i = \{a \in \triangle : \beta_i(a) \geqslant \beta_i(f(a))\}$ for $i \in \{1, 2, 3\}$. Then we have the following facts.

(1) Any point $a \in \triangle$ belongs to at least one M_i.

(2) If $a \in M_1 \cap M_2 \cap M_3$, then a is a fixed point.

Consider a sequence of refinements $\{\triangle_1, \triangle_2, ...\}$ of $\triangle$ such that the maximum diameter of small triangles in $\triangle_n$ is going to 0 as $n \to +\infty$. For example, we can consider the refining triangulations of the triangle $\triangle$ as follows:



We want to define a coloring $\phi : \triangle \to \{1, 2, 3\}$ such that

(a) Any $a \in \triangle$ with $\phi(a) = i$ belongs to M_i.

(b) The coloring ϕ satisfies the conditions of Sperner's Lemma for any subdivision $\triangle_n$ of $\triangle$.

Next we show such ϕ exists. This is because

- For the point A_i (say $i = 1$), we have that $A_1 = (1, 0) \in M_1$, so we can let $\phi(A_i) = i$;
- Consider a vertex $a = (x, y) \in A_1A_2$, i.e., $x+y = 1$. Since $\beta_1(f(a))+\beta_2(f(a)) \leq 1 = x+y = \beta_1(a) + \beta_2(a)$, so we must have at least one of $\beta_1(f(a)) \leq \beta_1(a)$ and $\beta_2(f(a)) \leq \beta_2(a)$ holds, which means that $a \in M_1 \cup M_2$.

Applying Sperner's Lemma to each $\triangle_n$ and the coloring ϕ, we get that there exists a small triangle $A_1^{(n)}A_2^{(n)}A_3^{(n)}$ in $\triangle_n$ which has three different colors 1,2,3.

Consider the sequence $\{A_1^{(n)}\}_{n\geq 1}$. Since everything is bounded, there is a subsequence $\{A_1^{(n_k)}\}_{k\geq 1}$ such that $\lim_{k\to+\infty} A_1^{(n_k)} = p \in \triangle$ exists. Since the diameter of $A_1^{(n)}A_2^{(n)}A_3^{(n)}$ is going to be 0 as $n \to +\infty$, we see that $\lim_{k\to+\infty} A_2^{(n_k)} = \lim_{k\to+\infty} A_3^{(n_k)} = p$. Since $\beta_i(A_i^{(n_k)}) \geqslant \beta_i(f(A_i^{(n_k)}))$ for $i \in [3]$ and f is continuous, we get $\beta_i(p) = \lim_{k\to+\infty} \beta_i(A_i^{(n_k)}) \geq \lim_{k\to+\infty} \beta_i(f(A_i^{(n_k)})) = \beta_i(f(p))$ for $i \in [3]$. This implies that $p \in M_1 \cap M_2 \cap M_3$, so p is a fixed point of f, that is, $f(p) = p$. ∎

4 The Pigeonhole Principle

4.1 Basics

Theorem 4.1 (The Pigeonhole Principle). *Let X be a set with at least $1+\sum_{i=1}^{k}(n_i-1)$ elements and let $X_1, X_2, ..., X_k$ be disjoint sets forming a partition of X. Then, there exists some i, such that $|X_i| \geq n_i$.*

Now we introduce some applications of the Pigeonhole Principle.

4.1.1 Two equal degrees

Theorem 4.2. *Any graph has two vertices of the same degree.*

Proof. Let G be a graph with n vertices. Suppose that G does not have two vertices of same degree. So the only exceptional case will be that there is exactly one vertex of degree i for all $i \in \{0, 1, ..., n-1\}$. But this is impossible to have a vertex with degree 0 and a vertex with degree $n-1$ at the same time. ∎

Exercise 4.3. *For any n, find an n-vertex graph G, which has exactly two vertices with the same degree.*

4.1.2 Chromatic number

Definition 4.4. *A* **vertex-coloring** *of a graph $G = (V, E)$ is a mapping $f : V \to C$, where C is the set of colors. A coloring is* **proper** *if no two adjacent vertices have the same color. The* **chromatic number** *$\chi(G)$ is the minimum size of the proper colorings.*

Theorem 4.5. *For any graph G on n vertices, we have $\alpha(G)\chi(G) \geq n$, where $\alpha(G)$ is the maximum size of an independent set in G.*

Proof. Given a proper coloring of G with $\chi(G)$ colors. Then we can partition V into $\chi(G)$ parts such that the vertices in each part have the same color. Since each part is an independent set, there exists an independent set with size equal to or larger than $\frac{n}{\chi(G)}$, which implies that $\alpha(G) \geq \frac{n}{\chi(G)}$. Thus we have $\alpha(G)\chi(G) \geq n$. ∎

4.1.3 Subsets without divisors

Question 4.6. *How large can a subset $S \subset [2n]$ be such that for any $i, j \in S$, we have $i \nmid j$ and $j \nmid i$?*

Obviously, we can take $S = \{n+1, n+2, ..., 2n\}$ with $|S| = n$.

Theorem 4.7. *For any $S \subset [2n]$ with $|S| \geq n+1$, there exist $i, j \in S$ such that $i|j$.*

Proof. For any odd integer $2k-1$, $k \in [n]$, let's define $S_{2k-1} = \{2^i \cdot (2k-1) \in S : i \geq 0\}$. Clearly, $S = \bigcup_{k=1}^{n} S_{2k-1}$. Since $|S| \geq n+1$, there exists some $|S_{2k-1}| \geq 2$. We say $x, y \in S_{2k-1}$. It is easy to see that we have $x|y$ or $y|x$. ∎

4.1.4 Rational approximation

Theorem 4.8. *Given $n \in \mathbb{Z}^+$, for any $x \in \mathbb{R}^+$, there is a rational number $\frac{p}{q}$ such that $1 \leq q \leq n$ and $|x - \frac{p}{q}| < \frac{1}{nq}$.*

Proof. For any $x \in \mathbb{R}^+$, let $\{x\} = x - \lfloor x \rfloor$ be the fractional part of x. Consider $\{ix\} \in [0,1)$, for any $i = 1, 2, ..., n+1$. Partition $[0,1)$ into n subintervals $[0, \frac{1}{n}), [\frac{1}{n}, \frac{2}{n}), ..., [\frac{n-1}{n}, 1)$. By Pigeonhole Principle, there exists a subinterval $[\frac{k}{n}, \frac{k+1}{n})$ containing two reals say $\{ix\}$ and $\{jx\}$ for $1 \leq i < j \leq n+1$. It's easy to check that $|\{jx\} - \{ix\}| < \frac{1}{n}$, so $|(j-i)x - (\lfloor jx \rfloor - \lfloor ix \rfloor)| < \frac{1}{n}$. Let $q = j - i$ $1 \leq q \leq n$ and $p = \lfloor jx \rfloor - \lfloor ix \rfloor \geq 0$. We know that q, p are integers. Thus we have $|qx - p| < \frac{1}{n}$, which implies $|x - \frac{p}{q}| < \frac{1}{nq}$. ∎

4.2 Erdős-Szekeres Theorem

Theorem 4.9 (Erdős-Szekeres Theorem)**.** *For any sequence of $mn+1$ real numbers $\{a_0, a_1, ..., a_{mn}\}$, there is an increasing subsequence of length $m+1$ or a decreasing subsequence of length $n+1$.*

Proof. Consider any sequence $\{a_0, a_1, ..., a_{mn}\}$. For any $i \in \{0, 1, ..., mn\}$, let f_i be the maximum size of an increasing subsequence starting at a_i. We may assume $f_i \in \{1, 2, ..., m\}$ for any $i \in \{0, 1, ..., mn\}$. By Pigeonhole Principle, there exists an $s \in \{1, 2, ..., m\}$ such that there are at least $n+1$ elements $i \in \{0, 1, ..., m\}$ satisfying $f_i = s$. Let these elements be $i_1 < i_2 < ... < i_{n+1}$.

We claim that $a_{i_1} \geq a_{i_2} \geq \cdots \geq a_{i_{n+1}}$. Indeed, If $a_{i_j} < a_{i_{j+1}}$ for some $j \in [n]$, then we would extend the longest increasing subsequence of length s starting at $a_{i_{j+1}}$ by adding a_{i_j} to obtain an increasing subsequence starting at a_{i_j} of length s+1, which is a contradiction to $f_{i_j} = s$.

∎

Remark: We may require the increasing or decreasing subsequence to be strictly increasing or strictly decreasing given that all a_i are distinct.

4.3 Ramsey's Theorem

Fact 4.10 (A party of six)**.** *Suppose a party has six participants. Participants may know each other or not. Then there must be three participants who know each other or do not know each other, i.e. any 6-vertex graph G has a K_3 or an I_3.*

Proof. We consider a graph G on six vertices, say $V(G) = [6]$. Each vertex i represents one participant: i and j are adjacent if and only if they know each other. Then we need to show that there are three vertices in G which form a triangle K_3 or an independent set I_3.

Consider vertex 1. There are five other persons. So 1 is adjacent to three vertices or not adjacent to three vertices. By symmetry, we may assume that 1 is adjacent to three vertices, say $2, 3, 4$. If one of pairs $\{2,3\}, \{2,4\}, \{3,4\}$ is adjacent, then we have a K_3. Otherwise, $\{2,3,4\}$ forms an independent set of size three. This finishes the proof. ∎

Definition 4.11. *An **r-edge-coloring** of K_n is a mapping $f : E(K_n) \longrightarrow \{1, 2, \ldots, r\}$ which assigns one of the colors $1, 2, \ldots, r$ to each edge of K_n.*

Definition 4.12. *Given an r-edge-coloring of K_n, a clique in K_n is called* **monochromatic***, if all its edges are colored by the same color.*

Then the example of a party of six says that any 2-edge-coloring of K_6 has a monochromatic K_3.

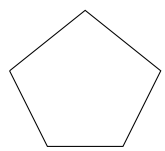
Definition 4.13. *For $k, \ell \geq 2$, the* **Ramsey Number** *$R(k, \ell)$ denotes the smallest integer N such that any 2-edge-coloring of K_N has a blue K_k or a red K_ℓ.*

Let us try to understand this definition a bit more:

- $R(k, \ell) \leq L$ if and only if any 2-edge-coloring of K_L has a blue K_k or a red K_ℓ.
- $R(k, \ell) > M$ if and only if there exists a 2-edge-coloring of K_M which has no blue K_k nor red K_ℓ.

Fact 4.14. *(1) $R(k, \ell) = R(\ell, k)$.*
(2) $R(2, \ell) = \ell$ and $R(k, 2) = k$.
(3) $R(3, 3) = 6$.

Proof. It is easy to know that (1) and (2) is right. We have $R(3, 3) \leq 6$ from the fact on a party of six. On the other hand, we have $R(3, 3) > 5$ from the following graph (if u, v are adjacent, we color edge uv blue, otherwise we color edge uv red).



∎

Proof. ∎

Theorem 4.15 (Ramsey's Theorem)**.** *The Ramsey number is finite. In fact we have that $R(k, \ell) \leq R(k-1, \ell) + R(k, \ell-1)$. Thus in particular $R(k, \ell) \leq \binom{k+\ell-2}{k-1}$ for $\ell, k \geq 2$.*

Proof. first we prove that $R(k, \ell) \leq R(k-1, \ell) + R(k, \ell-1)$. Let $n = R(k-1, \ell) + R(k, \ell-1)$. Consider any 2-edge-coloring of $G = K_n$. For any vertex x, define $A = \{y \in V(G) \setminus \{x\} : xy$ is blue$\}$ and $B = \{y \in V(G) \setminus \{x\} : xy$ is red$\}$. Then

$$|A| + |B| = n - 1 = R(k-1, \ell) + R(k, \ell-1) - 1.$$

By The Pigeonhole Principle, we have either $|A| \geq R(k-1, \ell)$ or $|B| \geq R(k, \ell-1)$.
Case 1. $|A| \geq R(k-1, \ell)$.
The induced subgraph $G[A]$ contains a blue K_{k-1} or a red K_ℓ. If $G[A]$ contains a red K_ℓ, G contains a red K_ℓ. In the former case, by adding the vertex x to that blue K_{k-1}, we can obtain a blue K_k in the G.
Case 2. $|B| \geq R(k, \ell-1)$.
This case is similar.

Next we will prove $R(k,\ell) \le \binom{k+\ell-2}{k-1}$ by induction on $k+\ell$. Base case is trivial, since $R(2,\ell) = R(\ell,2) = \ell$. Assume the claim holds for all $R(s,t)$ with $s+t < k+\ell$. Then

$$R(k,\ell) \le R(k-1,\ell) + R(k,\ell-1) \le \binom{k-1+\ell-2}{k-2} + \binom{k+\ell-1-2}{k-1} = \binom{k+\ell-2}{k-1}.$$

∎

Theorem 4.16. *If for some (k,ℓ), the numbers $R(k-1,\ell)$ and $R(k,\ell-1)$ are both even, then*

$$R(k,\ell) \le R(k-1,\ell) + R(k,\ell-1) - 1.$$

Proof. Let $n = R(k-1,\ell) + R(k,\ell-1) - 1$. So n is odd. Consider any 2-edge-coloring of K_n. For any vertex x, define the following as before $A_x = \{y : xy \text{ is blue}\}$ and $B_x = \{y : xy \text{ is red}\}$.

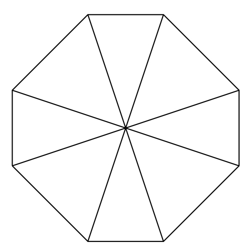
The previous proof tells us that if $|A_x| \ge R(k-1,\ell)$ or $|B_x| \ge R(k,\ell-1)$, then we can find a blue K_k or a red K_ℓ. Thus, we may assume that $|A_x| \le R(k-1,\ell)-1$ and $|B_x| \le R(k,\ell-1)-1$ for any vertex v, which implies that

$$n \le A_x + B_x + 1 \le R(k-1,\ell) + R(k,\ell-1) - 1.$$

This shows that for each x, $|A_x| = R(k-1,\ell)-1$ and $|B_x| = R(k,\ell-1)-1$. Now we consider the graph G consisting of all blue edges. Note that G has an odd number of vertices and any vertex has odd degree. But this contradicts to the Handshaking Lemma. ∎

Corollary 4.17. $R(3,4) = 9$.

Proof. By the previous theorem, we have $R(3,4) \le R(2,4) + R(3,3) - 1 = 4 + 6 - 1 = 9$. On the other hand, we have $R(3,4) > 8$ from the following 8-vertex graph (if u, v are adjacent, we color edge uv blue, otherwise we color edge uv red).



∎

Definition 4.18. *For any $k \ge 2$ and any integers $s_1, s_2, \ldots, s_k \ge 2$, the* **multi-color Ramsey number** *$R_k(s_1, s_2, \ldots, s_k)$ is the least integer N such that any k-edge-coloring of K_N has a clique K_{s_i} in color i, for some $i \in [k]$.*

Exercise 4.19. $R_k(s_1, s_2, \ldots, s_k) < +\infty$.

Theorem 4.20 (Schur's Theorem)**.** *For $k \ge 2$, there exists some integer $N = N(k)$ such that for any coloring $\varphi : [N] \to [k]$, there exist three integers $x, y, z \in [N]$ satisfying that $\varphi(x) = \varphi(y) = \varphi(z)$ and $x + y = z$.*

Proof. Let $N = R_k(3, 3, ..., 3)$. Define a k-edge-coloring of K_N from the coloring φ as follows: for any $i, j \in [N]$, define the color of ij to be $\varphi(|i - j|)$. By the definition of $R_k(3, 3, ..., 3)$, we can find a monochromatic triangle, say $ij\ell$. Suppose $i < j < \ell$, we have $\varphi(\ell - j) = \varphi(\ell - i) = \varphi(j - i)$. Let $x = \ell - j, y = j - i, z = \ell - i \in [N]$. We have $\varphi(x) = \varphi(y) = \varphi(z)$ and $x + y = z$. This finishes the proof. ∎

Exercise 4.21. *Prove that Schur's Theorem is also true while x, y, z are required to be distinct.*

Using this theorem, Schur proved the restricted version of Fermat's last problem in $\mathbb{Z}_p$ for sufficiently large prime p.

Theorem 4.22 (Schur). *For any integer $m \geq 1$, there is an integer $p(m)$ such that for any prime $p \geq p(m)$, $x^m + y^m = z^m$ (mod p) has a nontrivial solution in $\mathbb{Z}_p$.*

Proof. For prime p, consider the multiplicative group $\mathbb{Z}_p^* = \{1, 2, ..., p - 1\}$. Let g be a generator of $\mathbb{Z}_p^*$. Then for $x \in \mathbb{Z}_p^*$, there exists exactly one pair of integers (i, j) such that $x = g^{im+j} \pmod p$ for some $0 \leq j \leq m-1$ and $0 \leq im+j \leq p-2$. Then we define a coloring $\varphi : \mathbb{Z}_p^* \to \{0, 1, ..., m-1\}$ by letting $\varphi(x) = j$.

By Schur's Theorem, choose $p(m) = N(m)$, and for any $p \geq p(m)$, the coloring φ gives $x, y, z \in \mathbb{Z}_p^*$ satisfying $\varphi(x) = \varphi(y) = \varphi(z)$ and $x + y = z$. Let $x = g^{i_1 m+j},\ y = g^{i_2 m+j},\ z = g^{i_3 m+j}$ (mod p). Then $x + y = z$ implies that

$$g^{i_1 m+j} + g^{i_2 m+j} = g^{i_3 m+j} \pmod p, \tag{4.1}$$

thus

$$g^{i_1 m} + g^{i_2 m} = g^{i_3 m} \pmod p.$$

Let $\alpha = g^{i_1},\ \beta = g^{i_2},\ \gamma = g^{i_3}$. We have

$$\alpha^m + \beta^m = \gamma^m \pmod p.$$

∎

Remark: Schur's theorem holds in $\mathbb{Z}$, but we need to restrict the calculation in a multiplication cyclic group when deducing equation (4.1).

Definition 4.23. *Let $r \geq 3$. An* ***r*-uniform hypergraph** *(or an r-graph) is a pair (V, E) such that $E \subset \binom{V}{r}$. Let $K_n^{(r)}$ be the* **complete *r*-uniform hypergraph** *on n vertices ($K_n^{(r)} = (V, \binom{V}{r})$ with $|V| = n$).*

Definition 4.24. *The* **hypergraph Ramsey number** *$R^{(r)}(s, t)$ is the least integer N such that any 2-edge coloring of $K_n^{(r)}$ has a blue $K_s^{(r)}$ or a red $K_t^{(r)}$.*

Exercise 4.25. *Prove that for any integer $s, t > r$, $R^{(r)}(s, t) < +\infty$.*