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 **3**
1.1 Binomial Coefficients 3
1.2 Counting Mappings 5
1.3 The Binomial Theorem 6
1.4 Inclusion and Exclusion Principle (IEP) 8
1.5 Möbius Inversion Formula 10
1.6 Generating Functions 11
1.7 Random Walks 14
1.8 Exponential Generating Functions 16

2 Basics of Graphs **20**

3 Double-counting **22**
3.1 Basics 22
3.2 Sperner's Theorem 22
3.3 Littlewood-Offord Problem 25
3.4 Turán Type Problems 26
3.5 Sperner's Lemma 28

4 The Pigeonhole Principle **31**
4.1 Basics 31
4.2 Erdős-Szekeres Theorem 32
4.3 Ramsey's Theorem 32

5 Trees **37**
5.1 Tree Characterization 37
5.2 The First Proof of Cayley's Formula 38
5.3 The Second Proof of Cayley's Formula 39
5.4 The Third Proof of Cayley's Formula 41

6 System of Distinct Representatives **44**
6.1 Hall's Theorem . . . 44
6.2 Latin Rectangles . . . 45
6.3 Decomposition of Doubly Stochastic Matrices . . . 45
6.4 König's Min-Max Theorem . . . 46
6.5 Matchings in Bipartite Graphs . . . 46

7 Extremal Combinatorics **48**
7.1 Erdős-Ko-Rado Theorem . . . 48
7.2 Turán's Theorem . . . 50

8 Partially Ordered Sets (Poset) **52**

9 The Probabilistic Method **56**
9.1 Basics . . . 56
9.2 Union Bound . . . 56
9.3 The Linearity of Expectation . . . 59
9.4 The Deletion Method . . . 61
9.5 Markov's inequality . . . 62
9.6 Lovász Local Lemma . . . 64

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

Proof. (1) is trivial. Since $2^{[n]} = \cup_{0 \le k \le n} \binom{[n]}{k}$, we see $2^n = \sum_{0 \le k \le n} \binom{n}{k}$, 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 $A = \{\text{integer solutions } (x_1, \ldots, x_n) \text{ to } x_1 + \cdots + x_n = k,\ x_i \geq 0\}$ and $A = \{\text{integer solutions } (y_1, \ldots, y_n) \text{ to } y_1 + \cdots + y_n = n + k,\ y_i > 0\}$. 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* $X = [n]$, $A = \{(a_1, a_2, \dots, a_r) | a_i \in X\ , 1 \le a_1 \le a_2 \le \cdots \le a_r \le n, a_{i+1} - a_i \ge k+1, i \in [r-1]\}$. *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 $(x_1, x_2, \dots, x_r)$ 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).

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

(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*

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

We have two ways to view this power series.

(i). When the power series $\sum_{n\geq 0} a_n x^n$ 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 $\sum_{n\geq 0} a_n x^n$ 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 $a(x) = \sum_{n\geq 0} a_n x^n$ and $b(x) = \sum_{n\geq 0} b_n x^n$.

Addition.

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

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

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

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*

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

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* $n \geq 1$. *Find* a_n.

Solution. Consider the generating function,

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

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

Solution. Let $a_n = |A_n|$. We first observe that $a_1 = 3, a_2 = 8$. For $n \geq 3$, 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 $n \geq 3$, 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 $n \geq 2$. The generating function of $\{a_n\}$ is

$$f(x) = \sum_{n \geq 0} a_n x^n = a_0 + a_1 x + \sum_{n \geq 2} (2a_{n-1} + 2a_{n-2}) x^n = 1 + 3x + 2x(f(x) - 1) + 2x^2 f(x),$$

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 $F = \{(e, v) : e \in E(G), v \in V(G) \text{ such that } v \text{ is incident to } e\}$. 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 $O = \{v \in V(G) : d(v) \text{ is odd}\}$ and $\mathcal{E} = \{v \in V(G) : d(v) \text{ is even}\}$. 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

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

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*

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

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* **symmetric** *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

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

We then define the "partial pairing of parentheses" 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 the same partial pairing, 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 $S = \{(\{u_1, u_2\}, w) : u_1 w u_2$ 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

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

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. ∎

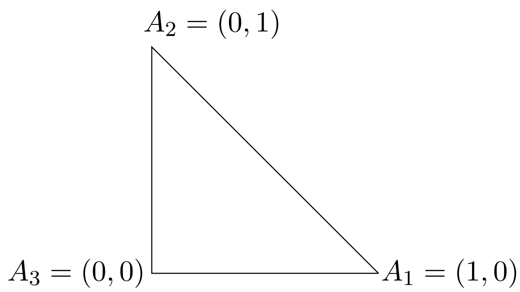
Before we introduce an interesting application of Sperner's lemma, we introduce the following theorem first.

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

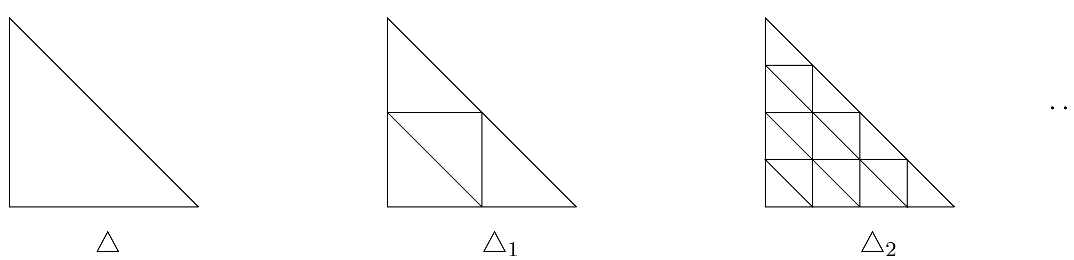
$$\begin{cases} \beta_1(a) &= x, \\ \beta_2(a) &= y, \\ \beta_3(a) &= 1 - x - y. \end{cases}$$

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.

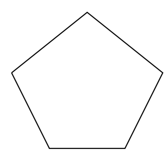
Definition 4.13. *For $k, \ell \geq 2$, the* **Ramsey Number** *$R(k, \ell)$ denotes the smallest integer N such that any 2-edge-coloring of K_N has a blue K_k or a red K_ℓ.*

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 $B_x = \{y : xy \text{ is red}\}$.

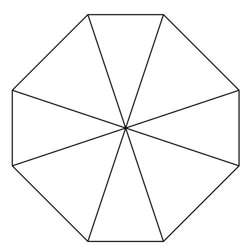
The previous proof tells us that if $|A_x| \ge R(k-1,\ell)$ or $|B_x| \ge R(k,\ell-1)$, then we can find a blue K_k or a red K_ℓ. Thus, we may assume that $|A_x| \le R(k-1,\ell)-1$ and $|B_x| \le R(k,\ell-1)-1$ for any vertex v, which implies that

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

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 $|V| = n$).*

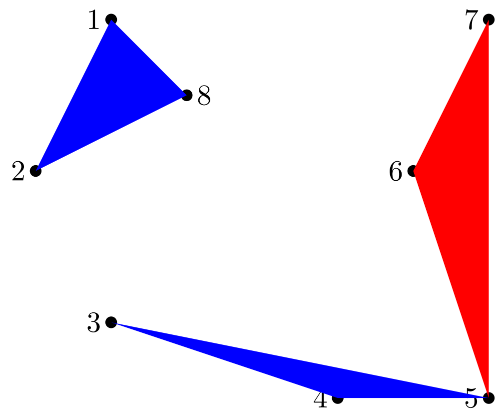
Definition 4.24. *The **hypergraph Ramsey number** $R^{(r)}(s, t)$ is the least integer N such that any 2-edge coloring of $K_n^{(r)}$ has a blue $K_s^{(r)}$ or a red $K_t^{(r)}$.*

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

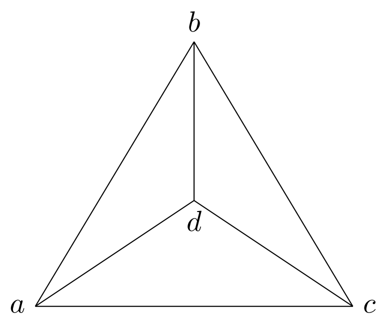
Theorem 4.26 (Erdős-Szekeres Theorem)**.** *For any integer n, there exists an integer $N(n)$ such that any collection of $N \geq N(n)$ points in the plane, no three on a line, has a subset of n points forming a convex n-gon.*

Proof. We observe that n points forms a convex n-gon if and only if every quadrilateral formed by taking 4 points in the n points is convex.

Let $N(n) = R^{(3)}(n, n)$. We will show that it is enough to have $N(n)$ points in the plane. Let S be a set of $N(n)$ such points. Number the points by $[N(n)]$ and color any triangle (every 3-subset) red if the path from the smallest number via the middle one to the largest number is clockwise, color any triangle blue if the path is counter clockwise.



In this way, we get a 2-edge-coloring of $K_N^{(3)}$. By the definition of Ramsey number, we have a monochromatic $K_n^{(3)}$. We may assume it is red. Now we claim that this red $K_n^{(3)}$ gives a convex n-gon. So it suffices to prove that there are no four points forming the following configuration.



If such configuration exists, since Δabc is a red triangle, we may assume $a < b < c$. Now $a < c$, and Δacd is a red triangle, therefore $a < d < c$.

If $b < d$ then $b < d < c$, Δbcd is a blue triangle.

If $d < b$ then $a < d < b$, Δabd is a blue triangle.

This contradiction completes the proof. ∎

5 Trees

5.1 Tree Characterization

Definition 5.1. *A graph G is* **connected***, if for any vertices u and v, G contains a path from u to v. Otherwise, G is* **disconnected***.*

Definition 5.2. *A* **component** *of a graph G is a maximal connected subgraph of G.*

Definition 5.3. *A graph T is called a* **tree** *if it is connected but contains no cycle. A vertex in a tree T with degree one is called a* **leaf***.*

Recall Euler's Formula on planar graphs that if a connected plane graph G has exactly n vertices, e edges and f faces, then $n - e + f = 2$. The following fact is the direct consequence of this formula and we will give a separate proof of the fact.

Fact 5.4 (Euler's Formula on trees)**.** *For any tree $T = (V, E)$, we have $|V| = |E| + 1$.*

Proof. First, any tree has at least one leaf. As otherwise, all vertices have degree at least 2, then this gives a cycle, which is a contradiction.

Next we apply induction on n. Consider the base case that $n = 2$, the tree is an edge, then we are done. Now we assume the statement holds for any tree on $n - 1$ vertices. Consider a tree T on n vertices ($n \geq 2$). We know that T contains a leaf, call v. It is easy to see that $T - \{v\}$ is still a tree as it is connected and has no cycles which has $n - 1$ vertices. By induction, we know that $T - \{v\}$ has $n - 2$ edges. So T has $n - 1$ edges. ∎

Fact 5.5. *Any tree T with at least two vertices has at least two leaves.*

Proof. Assume for a contradiction that an n-vertex tree T has exactly one leaf v, then $d(u) \geq 2$ for any $u \in V(T)\backslash\{v\}$. Thus

$$2(n-1) = 2e(T) = \sum_{x \in V(T)} d(x) \geq 2(n-1) + 1 = 2n - 1,$$

which is a contradiction. ∎

Theorem 5.6 (Tree characterization)**.** *Let $T = (V, E)$ be a graph. Then the following statements are equivalent:*

(i). T is a tree. (i.e. connected and no cycle.)

(ii). T is a "minimal" connected graph. (i.e. deleting any edge will result in a disconnected graph.)

(iii). T is a "maximal" graph without a cycle. (i.e. adding any new edge will result in a cycle.)

Proof. (i)$\Rightarrow$(ii): Suppose (ii) fails, then there exists $e = xy \in E(T)$ such that $T - \{e\}$ is still connected. Then $T - \{e\}$ has a path P from x to y. So $P \cup \{e\}$ is a cycle in T, which is a contradiction.

(ii)$\Rightarrow$(i): Suppose (i) fails, then T contains a cycle C. If we delete any edge e from C, $T - \{e\}$ remains connected, which is a contradiction.

(i)⇒(iii): For any new edge $e = xy$. As T is connected, T has a path P from x to y. Thus, $P \cup \{e\}$ gives a cycle.

(iii)⇒(i): Suppose (i) fails, so T is disconnected. Then T has two components (say D_1 and D_2). Pick $x \in D_1$ and $y \in D_2$. If we add the new edge $e = xy$, then it is easy to see that $T + \{e\}$ still has no cycle, which is a contradiction. ∎

Definition 5.7. *Given a graph G, a subgraph H of G is a* **spanning subgraph** *if $V(H) = V(G)$.*

Fact 5.8. *Any graph G is connected if and only if it contains a spanning tree.*

Proof. For the sufficiency, if G has a spanning tree, then it is connected. Suppose G is connected. Deleting edges of G until it satisfies the property (ii) in the Theorem 5.6, then we get a spanning tree. So the necessity follows. ∎

Definition 5.9. *Given a connected graph G with n vertices, say $v_1, ..., v_n$. Let $ST(G)$ be the number of labeled spanning trees in G.*

Theorem 5.10 (Cayley's Formula)**.** *For an integer $n \geq 2$,*

$$ST(K_n) = n^{n-2}.$$

We will give three proofs for this formula.

5.2 The First Proof of Cayley's Formula

Let $V(K_n) = \{v_1, v_2, \cdots, v_n\}$ and given a spanning tree T. Then

$$\sum_{i=1}^{n} d(v_i) = 2e(T) = 2n - 2.$$

Now we introduce a lemma.

Lemma 5.11. *Let $d_1, d_2, ..., d_n$ be positive integers with $\sum_{i=1}^{n} d_i = 2n - 2$. Then the number of spanning trees in K_n on vertex set $\{v_1, ..., v_n\}$ satisfying $d(v_i) = d_i$ is equal to*

$$\frac{(n-2)!}{(d_1-1)!(d_2-1)!\cdots(d_n-1)!} = \binom{n-2}{d_1-1, d_2-1, \cdots, d_n-1}.$$

Proof. We prove by induction on n. Base case is trivial. When $n = 2$, $d_1 = d_2 = 1$, there is only one spanning tree.

Now we assume that this statement holds for any sequence of $n-1$ positive integers. Then consider $d_1, ..., d_n$ with $\sum_{i \in [n]} d_i = 2n-2$. By average, $(\sum d_i)/n < 2$, so there exists some $d_i = 1$, say $d_n = 1$. Let $\mathcal{F}$ be the family of all spanning trees with $d(v_i) = d_i$ for $i \in [n]$. And let $\mathcal{F}_i = \{T - \{v_n\} : T \in \mathcal{F}$, the unique neighbor of v_n in T is $v_i\}$. So $|\mathcal{F}| = \sum_{i=1}^{n-1} |\mathcal{F}_i|$. All trees in $\mathcal{F}_i$ have $n-1$ vertices $\{v_1, v_2, \cdots, v_{n-1}\}$ such that

$$\begin{cases} d(v_j) = d_j, & \text{if} \quad j \neq i \\ d(v_i) = d_i - 1, & \text{otherwise.} \end{cases}$$

By induction, we have

$$|\mathcal{F}_i| = \frac{(n-3)!}{(d_1-1)!\cdots(d_i-2)!\cdots(d_{n-1}-1)!} = \frac{(n-3)!(d_i-1)}{\prod_{j=1}^{n-1}(d_j-1)!}.$$

So

$$|\mathcal{F}| = \sum_{i=1}^{n-1}|\mathcal{F}_i| = \frac{(n-3)!}{\prod_{j=1}^{n-1}(d_j-1)!}\left(\sum_{i=1}^{n-1}(d_i-1)\right) = \frac{(n-2)!}{\prod_{j=1}^{n}(d_j-1)!}.$$

∎

Recall the Multinomial Theorem:

$$(x_1+x_2+\cdots+x_k)^n = \sum_{i_1+\cdots+i_k=n}\frac{n!}{i_1!\cdots i_k!}x_1^{i_1}\cdots x_k^{i_k},$$

which implies

$$k^n = \sum_{i_1+\cdots+i_k=n}\frac{n!}{i_1!\cdots i_k!}.$$

Thus we have

$$ST(K_n) = \sum_{\substack{\sum_{i=1}^n d_i=2n-2\\ d_i\geq 1}}\frac{(n-2)!}{\prod_{j=1}^{n}(d_j-1)!} = n^{n-2}.$$

5.3 The Second Proof of Cayley's Formula

Definition 5.12. *A multigraph is a loopless graph, where we allow multiple edges between vertices.*

For a multigraph G on $[n]$, we define the Laplace matrix $Q=(q_{ij})_{n\times n}$ of G as follows:

$$q_{ij} = \begin{cases} d_G(i), & \text{if } i=j\\ -m, & \text{if } i\neq j \text{ and there are } m \text{ edges between } i \text{ and } j.\end{cases}$$

Note that Q is symmetric, and the sum of each row/column is 0.

For example

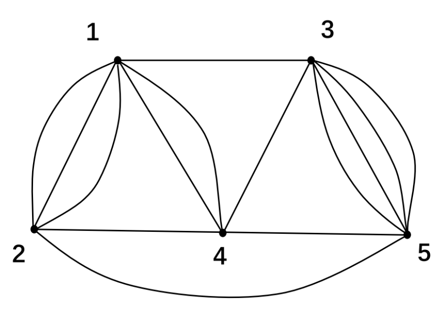


Figure 1: Multigraph G

$$Q = \begin{pmatrix} 6 & -3 & -1 & -2 & 0 \\ -3 & 5 & 0 & -1 & -1 \\ -1 & 0 & 6 & -1 & -4 \\ -2 & -1 & -1 & 5 & -1 \\ 0 & -1 & -4 & -1 & 6 \end{pmatrix}.$$

Note that for an $n \times n$ matrix Q, let Q_{ij} be the $(n-1) \times (n-1)$ matrix obtained from Q by deleting the i^{th} row and j^{th} column. For multigraphs, if two vertices u and v are joined by several edges, then we count each spanning tree in which u and v are adjacent to the corresponding number of times. In other words, we distinguish between spanning trees that use different edges. Then we have the following theorem.

Theorem 5.13. *For any multigraph G, $ST(G) = det(Q_{11})$, where Q_{ij} is the $(n-1) \times (n-1)$ matrix obtained from the Laplace matrix Q of G by deleting the i^{th} row and j^{th} column.*

Proof. We consider a multigraph G in the following two cases.

Case 1. Suppose that the multigraph G has an isolated vertex, say vertex number 1, then we have $ST(G) = 0$. Since the first column of the Laplace matrix consists only of zeros, the sum of the rows of Q_{11} is also zero, which implys zero is the eigenvalue of Q_{11}. Thus, we have $det(Q_{11}) = ST(G)$.

Case 2. Suppose that the multigraph G has no isolated vertices, we prove this by induction on the number of edges in G. Base case, suppose that $e(G) = 1$. Then it holds trivially.

Now we assume this holds for any multigraph with less than $e(G)$ edges. Take any edge e in G. Define two multigraph as follows.

1. $G - e$ is the multigraph obtained from G by deleting the edge e.
2. G/e is the multigraph obtained from G by contracting the two endpoints x, y of e into a new vertex z and adding new edges in $\{zu : xu \in E(G)\} \cup \{zv : yv \in E(G)\}$.

Let Q' and Q'' be the Laplace matrices of $G - e$ and G/e respectively. Since G has no isolated vertices, the vertex number 1 is incident to at least one edge. More precisely, assume that the edge e has endpoints 1 and 2. Take the multigraph G in Figure 1 as an example, we have

$$Q' = \begin{pmatrix} 5 & -2 & -1 & -2 & 0 \\ -2 & 4 & 0 & -1 & -1 \\ -1 & 0 & 6 & -1 & -4 \\ -2 & -1 & -1 & 5 & -1 \\ 0 & -1 & -4 & -1 & 6 \end{pmatrix}, \quad Q'' = \begin{pmatrix} 5 & -1 & -3 & -1 \\ -1 & 6 & -1 & -4 \\ -3 & -1 & 5 & -1 \\ -1 & -4 & -1 & 6 \end{pmatrix}.$$

Let $Q_{11,22}$ be the matrix obtained from Q by deleting the first two rows and the first two columns. Then we have

$$det(Q_{11}) = det((Q')_{11}) + det(Q_{11,22}). \tag{5.2}$$

We also see that

$$Q_{11,22} = (Q'')_{11}. \tag{5.3}$$

By (5.2) and (5.3) we have

$$det(Q_{11}) = det((Q')_{11}) + det((Q'')_{11}). \tag{5.4}$$

Claim. For any edge e in G, we have

$$ST(G) = ST(G - e) + ST(G/e). \tag{5.5}$$

Proof. We divide the spanning trees of G into two classes:

-the 1^{st} class contains those spanning trees of G NOT containing e, which are exactly $ST(G - e)$.

-the 2^{nd} class contains those spanning trees of G containing e. We can easily see that the trees in the 2^{nd} class are one-to-one corresponding to the spanning trees of G/e.

This proves (5.5). ∎

By induction, we have $ST(G - e) = det(Q'_{11})$, $ST(G/e) = det((Q'')_{11})$. By (5.4), we have $ST(G) = det(Q_{11})$. ∎

For K_n, we have

$$Q = \begin{pmatrix} n-1 & -1 & \cdots & -1 \\ -1 & n-1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & n-1 \end{pmatrix}_{n \times n},$$

which implies that $ST(G) = det(Q_{11}) = n^{n-2}$.

5.4 The Third Proof of Cayley's Formula

Definition 5.14. *A digraph $D = (V, A)$ consists of a vertex set V and an arc set $A \subseteq \{(i, j) : i, j \in V\}$*

Let $\mathscr{D}$ be the family of digraphs $D = ([n], A)$ such that each vertex in D has exactly one arc going out from this vertex (i.e. each vertex has out degree one).

Fact 5.15.

$$|\mathscr{D}| = n^n.$$

Proof. Consider the set $\mathscr{F} = \{\text{all mapping } f : [n] \to [n]\}$. It is easy to see there exists a bijection between $\mathscr{D}$ and $\mathscr{F}$. So $|\mathscr{D}| = |\mathscr{F}| = n^n$. ∎

Definition 5.16. *Given a spanning tree of K_n, we choose 2 special vertices (one marked by a circle and the other marked by a square; these two vertices can be the same vertex). We call such a subject (the spanning tree with 2 special vertices) as a vertebrate.*

Let $\mathscr{V}$ be a family of all vertebrates on $[n]$. Clearly, we have $|\mathscr{V}| = ST(K_n)n^2$. So it suffices to show $|\mathscr{V}| = n^n$ to get the Cayley's formula.

Lemma 5.17. *There exists a bijection between $\mathscr{V}$ and $\mathscr{D}$.*

Proof. Consider a $W \in \mathscr{V}$ (see Figure 2). Let P be the unique path in W between the two special vertices (marked by a circle and a square), and view P as a directed path from the circle to the square.

We then define a digraph D_1 on $V(P)$ by assign the following arcs (Figure 3): that is, we place two rows, where the 1^{st} row is from P and the 2^{nd} row is the increasing sequence of $V(P)$, then we orient the arcs of D_1 from the vertices of the 2^{nd} row to the one above it. Thus each vertex in D_1 has exactly one arc going out and one arc going in.

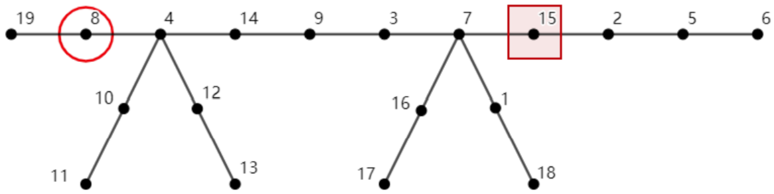


Figure 2: A vertebrate

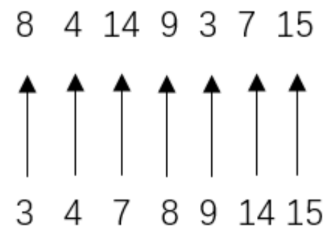


Figure 3: D_1

Exercise 5.18. *D_1 consists of vertex-disjoint directed cycle (possibly loops and 2-cycles).*

Next, we extend D_1 to a digraph D on $[n]$, by the following:

(1) We remove all edges of P from W.

(2) Then $W - E(P)$ consists of subtrees, each having one vertex from $V(P)$. We direct the edges of these subtrees such that they point to the unique vertex of the component contained in $V(P)$.

(3) These arcs added in (2) together with the arcs of D_1, define a new graph D_W on $[n]$. This should be easy to see that $D_W \in \mathscr{D}$.

So we just define a mapping $\varphi : \mathscr{V} \to \mathscr{D}$, by assigning $\varphi(W) = D_W,\ W \in \mathscr{V}$. Next, We show φ is a bijection.

Step 1. We can define $\varphi^{-1} : \mathscr{D} \to \mathscr{V}$ such that $\varphi^{-1} \cdot \varphi = Id$.

Remark: In any D_W, $V(D_1)$ consists of all vertices in D_W contained in a directed cycle.

Take any $D \in \mathscr{D}$, there exists some vertex of D contained in a directed cycle. Let X be the set of all such vertices of D. Since $D[X]$ consists of vertex-disjoint directed cycles, there is a nature way to define a path as follows (see Figure 4):

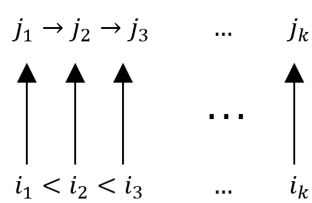


Figure 4: Define a path

First, list the vertices of X in the increasing order. Second, list the out-neighbor vertices of X in another row, respectively. Then the second row defines a path P which is the special path in the vertebrate. Then it is easy to define the rest part of the vertebrate say W. So we have $D \in \mathscr{D} \xrightarrow{\varphi^{-1}} W \in \mathscr{V}$. We can check that $\varphi^{-1} \cdot \varphi = Id$.

Step 2. φ is a surjective.

We have proved in Step 1 that for any $D \in \mathscr{D}$, there exists $W \in \mathscr{V}$ satisfying $\varphi(W) = D$.

Therefore indeed φ is a bijection. ∎

Combining Fact 5.15 with Lemma 5.17, we get $ST(K_n) = n^{n-2}$.

6 System of Distinct Representatives

6.1 Hall's Theorem

Definition 6.1. *A* **system of distinct representatives** *for a sequence of (not necessarily distinct) sets* $S_1, S_2, \ldots, S_m$ *is a sequence of distinct elements* $x_1, \ldots, x_m$ *such that* $x_i \in S_i$ *for* $i = 1, 2, \ldots, m$.

Fact 6.2. *If* $S_1, S_2, \ldots, S_m$ *have a system of distinct representatives then the union of any* k *sets has at least* k *elements. i.e., the following* **Hall's condition** *is fulfilled:*

$$\left|\bigcup_{i\in I} S_i\right| \geq |I| \text{ for all } I \subseteq [m].$$

Example 6.3. $\{1,2\}, \{1\}, \{2\}$ *have no systems of distinct representatives.*

Theorem 6.4 (**Hall's Theorem**)**.** *The sets* $S_1, S_2, \ldots, S_m$ *have a system of distinct representatives if and only if* $S_1, S_2, \ldots, S_m$ *satisfy Hall's condition.*

Proof. It is clear that the necessity of Hall's condition holds. We prove the sufficiency of Hall's condition by induction on m. The case $m = 1$ is clear. Assume the claim holds for any collection with less than m sets.

Case 1: Suppose that for any $I \subsetneq [m]$, we have $|\cup_{i\in I} S_i| > |I|$. Take any $x_1 \in S_1$, and let $S'_i = S_i \backslash \{x_1\}, i = 2, \ldots, m$. Then $|\cup_{i\in I} S'_i| \geq |I|$ for all $I \subseteq [2, m]$, which satisfies Hall's condition. By induction hypothesis, $S'_2, \ldots, S'_m$ have a system of distinct representatives $x_2, \ldots, x_m$. Consequently, $x_1, \ldots, x_m$ is a system of distinct representatives of $S_1, \ldots, S_m$.

Case 2: Suppose that there exists some $I \subsetneq [m]$ satisfying that $|\cup_{i\in I} S_i| = |I| = k < m$. By the induction hypothesis, these k sets have a system of distinct representatives. Without loss of generality, let $I = \{1, 2, \ldots, k\}$, and $x_1, \ldots, x_k$ be a system of distinct representatives of the sets $S_1, S_2, \ldots, S_k$. For each i, $k+1 \leq i \leq m$, let $S'_i = S_i \backslash \{x_1, \ldots, x_k\}$. We claim that the sets $S'_{k+1}, S'_{k+2}, \ldots, S'_m$ satisfy Hall's condition. Otherwise, there exists some $J \subseteq [k+1, m]$ satisfying that $|\cup_{i\in J} S'_i| < |J|$. Then $|\cup_{i\in I\cup J} S_i| = |(\cup_{i\in J} S_{i'}) \cup \{x_1, x_2, \ldots, x_k\}| < |J| + k = |I \cup J|$, which is a contradiction. By induction, $S'_{k+1}, S'_{k+2}, \ldots, S'_m$ have a system of distinct represetatives $x_{k+1}, x_{k+2} \ldots, x_m$. Consequently, $x_1, x_2, \ldots, x_m$ is a system of distinct representatives of the sets $S_1, S_2, \ldots, S_m$. ∎

Definition 6.5. *Let* $m_0 \leq m_1 \leq \cdots \leq m_{n-1}$. *Define*

$$F_n(m_0, m_1, \ldots, m_{n-1}) = \prod_{i=0}^{n-1} (m_i - i)_*$$

where $\alpha_* = \max(1, \alpha)$.

Theorem 6.6. *Let* $S_0, S_1, \ldots, S_{n-1}$ *be a sequence of set satisfying the Hall's condition. Let* $m_i = |S_i|$ *for* $i \in \{0, 1, \ldots, n-1\}$ *such that* $m_0 \leq m_1 \leq \cdots \leq m_{n-1}$. *Then the number of SDR for* $S_0, S_1, \ldots, S_{n-1}$ *is at least* $F_n(m_0, m_1, \ldots, m_{n-1})$.

Exercise 6.7. *Prove Theorem 6.6.*

6.2 Latin Rectangles

Definition 6.8. *A* **Latin rectangle** *is an $r \times n$ matrix with entries in $[n]$ such that each row and each column has no repeated elements. A* **Latin square** *is an $n \times n$ Latin rectangle.*

Example 6.9. *Here is a 2×3 Latin rectangle and a 3×3 Latin square.*

1	2	3
3	1	2

1	2	3
3	1	2
2	3	1

Exercise 6.10. *Construct an $n \times n$ Latin square.*

Theorem 6.11 (**Evans' conjecture(1960)**). *If fewer than n entries in an $n \times n$ matrix are filled, then one can always complete it to obtain a Latin square.*

Proof. The conjecture was proved by Smetaniuk (1981) using a quite subtle induction argument. We omit it here. ∎

Remark 6.12. *Evans' condition is best possible, i.e. it is possible to fill n entries so that the resulting partial matrix cannot be completed. Here is an example when $n = 4$.*

1	2	3	?
			4

Theorem 6.13. *If $r < n$, then any given $r \times n$ Latin rectangle can be extended to an $(r+1) \times n$ Latin rectangle.*

Proof. Let R be an $r \times n$ Latin rectangle. For $j \in [n]$, define S_j to be the set of integers in $[n]$ which do not appear in the j-th column. Then it suffices to prove that the sets $S_1, \ldots, S_n$ have a system of distinct representatives. Note that $|S_j| = n - r$, and each i, $i \in [n]$, occurs in exactly $n - r$ sets S_j, so $S_1, \ldots, S_n$ satisfy Hall's condition. By Theorem 6.4, $S_1, \ldots, S_n$ have a system of distinct representatives. ∎

6.3 Decomposition of Doubly Stochastic Matrices

Definition 6.14. *An $n \times n$ matrix $A = (a_{ij})_{n \times n}$ with $a_{ij} \geq 0$ is called* **doubly stochastic matrix** *if the sum of each row and each column equals 1. If $a_{ij} \in \{0, 1\}$, then A is called a* **permutation matrix**.

Theorem 6.15 (**Birkhoff-Von Neumann Theorem**). *Every doubly stochastic matrix A is a convex combination of permutation matrices. That is, there exist permutation matrices $P_1, P_2, \ldots, P_s$ and non-negative reals $\lambda_1, \lambda_2, \ldots, \lambda_s$ such that*

$$A = \sum_{i=1}^{s} \lambda_i P_i \text{ and } \sum_{i=1}^{s} \lambda_i = 1.$$

Proof. Let m be the number of nonzero entries in A then $m \geq n$. To prove the theorem, we apply induction on m. The case $m = n$ is clear since A is a permutation matrix. The following exercise is a hint to complete the proof. ∎

Exercise 6.16. *Define $S_i = \{j : a_{ij} > 0\}$, $i \in [n]$. Show that the sets $S_1, S_2, \ldots, S_n$ have a system of distinct representatives.*

6.4 König's Min-Max Theorem

Definition 6.17. *Let A be a 0-1 matrix. Two 1's are* **dependent** *if they are in the same row or in the same column; otherwise, they are* **independent**. *The size of the largest set of independent 1's is also known as the term* **rank** *of A.*

Theorem 6.18 (König). *Let A be an $m \times n$ 0-1 matrix. The maximum number r of independent 1's is equal to the minimum number R of rows and columns required to cover all 1's in A.*

Proof. Clearly, $r \leq R$, since we can find r independent 1's such that each row or column covers at most one of them.

Next we show that $r \geq R$. Suppose that some a rows and b columns cover all 1's and $a+b = R$. Without loss of generality, we may assume that the first a rows and the first b columns cover all the 1's. Write A in the form

$$A = \begin{pmatrix} B_{a\times b} & C_{a\times(n-b)} \\ D_{(m-a)\times b} & O_{(m-a)\times(n-b)} \end{pmatrix}.$$

We will show that there are a independent 1's in C. Define $S_i = \{j : \ c_{ij} = 1\} \subseteq [n-b]$. We claim that $S_1, \ldots, S_a$ have a system of distinct representatives and thus we can choose a 1's from C, no two in the same column. Otherwise, by Hall's theorem, there are some k $(1 \leq k \leq a)$ sets, say $S_{i_1}, \ldots, S_{i_k}$, such that $\left|\bigcup_{j=1}^{k} S_{i_j}\right| < k$, i.e. the 1's in these k rows in C can be covered by less than k columns. Together with the first b columns and the other $a-k$ rows from the first a rows of A, we obtain a covering of all the 1's in A with at most $(k-1)+b+(a-k) = a+b-1$ rows and columns, a contrdiction.

Similarly, there are b independent 1's in D. Since altogether these $a+b$ 1's are independent, this shows that $r \geq a+b = R$, as desired. ∎

Remark 6.19. *By König's theorem, the maximum number of independent 1's in A is r if and only if the maximum value of $c+d$ is $m+n-r$ for any 0-submatrix of size $c \times d$ in A.*

6.5 Matchings in Bipartite Graphs

Definition 6.20. *In a graph G, two edges are* **disjoint** *if they have no vertex in common. A* **matching** *is a set of pairwise disjoint edges. A* **perfect matching** *is a matching that covers all vertices. For a bipartite graph $G = (A \sqcup B, E)$, if a matching of G covers all vertices of A, then we say it is a* **matching of A into B**. *Clearly, we have $|B| \geq |A|$.*

Fact 6.21. *For a bipartite graph $G = (A \sqcup B, E)$, define $S_x = N_G(x) \subseteq B$ as the set of all neighbours of x for all $x \in A$. The following are equivalent:*

(1) There exists a matching of A into B;

(2) The sets S_x with $x \in A$ have a system of distinct representatives;

(3) Any k vertices from A have at least k neighbours.

Definition 6.22. *A* **vertex cover** *in $G = (V, E)$ is a set of vertices $S \subseteq V$ such that each edge is incident to at least one vertex in S.*

Theorem 6.23. *For a bipartite graph, the maximum size of a matching is equal to the minimum size of a vertex cover.*

Proof. Let $G = (A \sqcup B, E)$ be a bipartite graph. Consider a 0-1 matrix $M = (m_{a,b})$ with $|A|$ rows labeled by vertices $a \in A$ and with $|B|$ columns labeled by vertices $b \in B$ such that

$$m_{a,b} = \begin{cases} 1, & \text{if } \ a \sim b, \\ 0, & \text{otherwise.} \end{cases}$$

Then there is a natural one-to-one correspondence between a matching of G and a family of independent 1's in M. Also, there is a natural one-to-one correspondence between a vertex cover in G and a family of rows and columns that covers all 1's in M. By Theorem 6.18, the maximum size of a matching is equal to the minimum size of a vertex cover. ∎

7 Extremal Combinatorics

7.1 Erdős-Ko-Rado Theorem

Definition 7.1. *A family $\mathcal{F} \subseteq 2^{[n]}$ is intersecting if for any $A, B \in \mathcal{F}$, we have $A \cap B \neq \emptyset$.*

Example 7.2. *Here are two intersecting families of size $|\mathcal{F}| = 2^{n-1}$.*

(1) $\mathcal{F} = \{A \subseteq [n] : 1 \in A\}$;

(2) $\mathcal{F} = \{A \subseteq [n] : |A| > \frac{n}{2}\}$ *when n is odd.*

Fact 7.3. *For any intersecting family $\mathcal{F} \subseteq 2^{[n]}$, we have $|\mathcal{F}| \leq 2^{n-1}$.*

Proof. Consider pairs $\{A, A^c\}$ for all $A \subseteq [n]$. Note that there are exactly 2^{n-1} such pairs, and $\mathcal{F}$ contains at most one subset from each pair. Therefore we have $|\mathcal{F}| \leq 2^{n-1}$ and Example 7.2 implies that the upper bound is tight. ∎

Example 7.4. *Here are two intersecting families of uniform size.*

(1) $\mathcal{F} = \{A \in \binom{[n]}{k} : 1 \in A\}$ *when $k \leq \frac{n}{2}$. We have $|\mathcal{F}| = \binom{n-1}{k-1}$.*

(2) $\mathcal{F} \subseteq \binom{[n]}{k}$ *when $k > \frac{n}{2}$. We have $|\mathcal{F}| = \binom{n}{k}$.*

Let $\mathcal{F}$ be an intersecting family of k-element subsets of $[n]$. The basic question is: how large can such a family be? To avoid trivialities, we assume $n \geq 2k$ since otherwise any two k-element sets intersect, and there is nothing to prove. The following theorem answers the question.

Theorem 7.5 (Erdős-Ko-Rado, 1961). *For $n \geq 2k$, the largest intersecting family $\mathcal{F} \subseteq \binom{[n]}{k}$ has size $\binom{n-1}{k-1}$. Moreover, if $n > 2k$, the intersecting family $\mathcal{F} \subseteq \binom{[n]}{k}$ with $|\mathcal{F}| = \binom{n-1}{k-1}$ must be a star. That is, $\mathcal{F} = \{A \in \binom{[n]}{k} : t \in A\}$ for some fixed $t \in [n]$.*

Proof. (Due to Katona 1972) Take a cyclic permutation $\pi = (a_1\ a_2\ \ldots\ a_n)$ of the elements of $[n]$. We say a set A is contained in π if all elements of A appear consecutively in π. Let $\mathcal{F}_\pi = \{A \in \mathcal{F}$: A is contained in π.$\}$.
Claim. $|\mathcal{F}_\pi| \leq k$ for all cyclic permutation π.

Proof of claim. Pick $A \in \mathcal{F}_\pi$, then there exist $2k-2$ sets B contained in π such that $B \cap A \neq \emptyset$ and $B \neq A$. But these $2k-2$ sets can be partitioned into $k-1$ pairs of disjoint subsets. Since $\mathcal{F}_\pi$ is intersecting, $\mathcal{F}_\pi$ contains at most one subset from each pair. So $|\mathcal{F}_\pi| \leq k$. This proves the claim.

We now do a double counting of

$$N = |\{(\pi, A) : \pi \text{ is a cyclic permutation of } [n], \text{ and } A \in \mathcal{F}_\pi\}|.$$

By Claim, when we fix π,

$$N = \sum_\pi |\mathcal{F}_\pi| \leq \sum_\pi k = k(n-1)!.$$

Then if we fix A, the number of cyclic permutations containing A is $k!(n-k)!$. So we have

$$k(n-1)! \geq N = \sum_{A \in \mathcal{F}} k!(n-k)! = |\mathcal{F}|k!(n-k)!,$$

which implies that

$$|\mathcal{F}| \le \frac{k \cdot (n-1)!}{k!(n-k)!} = \binom{n-1}{k-1}.$$

Next we show the extremal case: when $n > 2k$, the intersecting family $\mathcal{F} \subseteq \binom{[n]}{k}$ with $|\mathcal{F}| = \binom{n-1}{k-1}$ must be a star. That is, there exists some $t \in [n]$ such that $\mathcal{F} = \{A \in \binom{[n]}{k} : t \in A\}$.

Proof of the extremal case. From the proof above, we see that if $|\mathcal{F}| = \binom{n-1}{k-1}$, then $|\mathcal{F}_\pi| = k$ for any cyclic permutation π.

Claim 1. Fix any cyclic permutation $\pi = (a_1\ a_2\ \dots\ a_n)$. If $\mathcal{F}_\pi = \{A_1, A_2, \dots, A_k\}$, then $\mathcal{F}_\pi = \{\{a_{j+r}, a_{j+r+1}, \dots, a_{j+r+k-1}\} : 1 \le j \le k\}$ for some fixed $r \in [n]$. Without loss of generality, we may say $A_j = \{a_j, a_{j+1}, \dots, a_{j+k-1}\}$ for $1 \le j \le k$ and $A_1 \cap A_2 \cap \cdots \cap A_k = \{a_k\}$ (where the indices are taken under the additive group $\mathbb{Z}_n$). That is, if $A_1, A_k \in \mathcal{F}_\pi$ and $A_1 \cap A_k = \{a_k\}$, then all $B \subseteq \binom{A_1 \cup A_k}{k}$ whose elements appear consecutively in π are in $\mathcal{F}_\pi$.

Proof of Claim 1. Let B_1, B_2 be sets in $\mathcal{F}_\pi$. When $|B_1 \cap B_2|$ is minimized, it can be easily checked that $|B_1 \cap B_2| = 1$. Then for any k-subset $B \subseteq B_1 \cup B_2$ whose elements appear consecutively in π, we have $B \in \mathcal{F}_\pi$. This proves Claim 1.

If for any $A \in \mathcal{F}$ we have $a_k \in A$, then we are done. So we assume that there exists some $A_0 \in \mathcal{F}$ such that $a_k \notin A_0$. We will show that $\mathcal{F} = \binom{A_1 \cup A_k}{k}$. Then $|\mathcal{F}| = \binom{2k-1}{k} = \binom{2k-1}{k-1} < \binom{n-1}{k-1}$, which is a contradiction.

Claim 2. For any $B \in \binom{A_1 \cup A_k \setminus \{a_k\}}{k-1}$, we have $B \cup \{a_k\} \in \mathcal{F}$.

Proof of Claim 2. Let $B = B_1 \cup B_2$ with $B_1 \subseteq A_1, B_2 \subseteq A_k$. Consider another cyclic permutation π' whose elements appear in the following sequence: $A_1 \setminus (B_1 \cup \{a_k\}), B_1, a_k, B_2, A_2 \setminus (B_2 \cup \{a_k\})$. Then $A_1, A_k \in \mathcal{F}_{\pi'}$. By Claim 1, it is easy to check that $B \cup \{a_k\} \in \binom{A_1 \cup A_k}{k} \subseteq \mathcal{F}$. This proves Claim 2.

Claim 3. The subset $A_0 \in \mathcal{F}$ (with $a_k \notin A_0$) satisfies $A_0 \subseteq A_1 \cup A_k \setminus \{a_k\}$.

Proof of Claim 3. Otherwise, we have $A_0 \not\subseteq A_1 \cup A_k$ and $|A_0 \cap (A_1 \cup A_k)| \le k-1$. Then there exists some $B \in \binom{(A_1 \cap A_k) \setminus A_0}{k}$ such that $a_k \in B$. By Claim 2, we have $B \in \mathcal{F}$. But $A_0 \cap B = \emptyset$, which contradicts the fact that $\mathcal{F}$ is intersecting. This proves Claim 3.

Claim 4. We have $\binom{A_1 \cup A_k}{k} \subseteq \mathcal{F}$.

Proof of Claim 4. Consider any $i \in A_0$, let $B_i = ((A_1 \cup A_k) \setminus A_0) \cup \{i\}$. Since $a_k \in B_i$, by Claim 2, we have $B_i \in \mathcal{F}$. Repeating the proof of Claim 2, we can obtain that any k-subset of $A_1 \cup A_k$ containing i belongs to $\mathcal{F}$. In other words, any k-subset B of $A_1 \cup A_k$ must intersect A_0, and thus B belongs to $\mathcal{F}$. Then we have $\binom{A_1 \cup A_k}{k} \subseteq \mathcal{F}$. This proves Claim 4.

Claim 5. We have $\mathcal{F} \subseteq \binom{A_1 \cup A_k}{k}$.

Proof of Claim 5. Suppose there exists some k-subset $C \in \mathcal{F}$ such that $C \not\subseteq A_1 \cup A_k$, then $|(A_1 \cup A_k) \setminus C| \ge k$. So there exists some $B \in \binom{A_1 \cup A_k}{k}$ such that $B \cap C = \emptyset$. By Claim 4, we have $B \in \mathcal{F}$ and thus $B \cap C \neq \emptyset$, which is a contradiction. This proves Claim 5.

Under the assumption that there exists some $A_0 \in \mathcal{F}$ such that $a_k \notin A_0$, we have shown that $\mathcal{F} = \binom{A_1 \cup A_k}{k}$ and $|\mathcal{F}| = \binom{2k-1}{k} = \binom{2k-1}{k-1} < \binom{n-1}{k-1}$, which is a contradiction. This completes the proof of the extremal case. ∎

Theorem 7.6 (A generalization of Erdős-Ko-Rado Theorem)**.** *Let k be a fixed integer. Let $A_1, A_2, \ldots, A_m$ be m subsets of $[n]$ such that:*

- *For any $i \in [m]$, $|A_i| \leq k \leq \frac{n}{2}$.*
- *For any distinct $i, j \in [m]$, $A_i \not\subseteq A_j$ and $A_i \cap A_j \neq \emptyset$.*

Then $m \leq \binom{n-1}{k-1}$.

Proof. By Theorem 3.8, the family $2^{[n]}$ can be partitioned into a disjoint union of symmetric chains. We fixed this partition. Since for any distinct $i, j \in [m]$, $A_i \not\subseteq A_j$, each A_i must be in different chains. Notice that $\forall i \in [m]$, $|A_i| \leq k \leq \frac{n}{2}$, thus there exists $B_i \subseteq [n]$ in the same chain with A_i such that $|B_i| = k$ and $A_i \subseteq B_i$. Since B_i is in different chains, B_i are distinct. Now B_i is m distinct sets of size k, for any distinct $i, j \in [m]$, $B_i \cap B_j \supseteq A_i \cap A_j \neq \emptyset$. By Theorem 7.5 (Erdős-Ko-Rado Theorem), $m \leq \binom{n-1}{k-1}$. ∎

7.2 Turán's Theorem

Definition 7.7. *A* **Turán graph $T_r(n)$** *is an n-vertex complete r-partite graph, where each part has almost equal size (differs by at most 1). Let $n = tr + s$ where $1 \leq s \leq r$. Then in $T_r(n)$, we have s parts of size $t+1$ and $r - s$ parts of size t. The number of edges of Turán graph $T_r(n)$ is $e(T_r(n)) = \sum_{1 \leq i < j \leq r} |V_i||V_j| = \frac{r-1}{2r}n^2 - \frac{s(r-s)}{2r}$.*

Theorem 7.8 (Turán's Theorem)**.**

$$ex(n, K_{r+1}) = e(T_r(n)).$$

Moreover, the unique n-vertex K_{r+1}-free graph G with $e(G) = e(T_r(n))$ is $G = T_r(n)$.

Theorem 7.9 (Turán's Theorem approximate form)**.** *If G has n vertices and is K_{r+1}-free, then*

$$e(G) \leq \frac{r-1}{2r}n^2.$$

Proof of Theorem 7.9. We are given an n-vertex K_{r+1}-free graph G, where $V(G) = [n]$. Consider a function $p : [n] \to [0, 1]$ such that

$$\sum_{i \in [n]} p_i = 1.$$

We want to find the maximum of $f(p) = \sum_{ij \in E(G)} p_i p_j$ over all such functions $p : [n] \to [0, 1]$. Suppose p is the function obtaining the maximum $f(p)$, and subject to this, the number of vertices i with $p(i) \neq 0$ is minimized.
Claim. $\{i : p(i) > 0\}$ is a clique in G.

proof of claim. Suppose NOT, say $p(i), p(j) > 0$ and $ij \notin E(G)$. Let $S_i = \sum_{k\in N_G(i)} p_k$ and $S_j = \sum_{k\in N_G(j)} p_k$. Let $S_i \geqslant S_j$. Then we can assign a new function $p^* : [n] \to [0,1]$ such that

$$p^*(i) = p(i) + p(j),\ p^*(j) = 0 \text{ and } p^*(k) = p(k) \text{ for } k \in [n] \setminus \{i,j\}.$$

Now we have

$$f(p^*) = f(p) - (p_iS_i + p_jS_j) + (p_i + p_j)S_i = f(p) + (S_i - S_j)p_j \geqslant f(p).$$

By the choice of p, we see $f(p^*) = f(p)$, but p^* has fewer vertices i with positive weight than p, a contradiction. This proves the claim.

Let $S = \{1, 2, ..., s\} \subseteq V(G)$ be the set of vertices with positive weight. Then by the claim, we see $G[S] = K_s$, where $s \le r$ as G is K_{r+1}-free. Then

$$\begin{aligned}\max_p f(p) =& \frac{1}{2}\left[\Big(\sum_{1\le i\le s} p(i)\Big)^2 - \sum_{1\le i\le s} p^2(i)\right] = \frac{1}{2}[1 - \sum_{1\le i\le s} p^2(i)] \le \frac{1}{2}\left[1 - s\left(\frac{\sum_{1\le i\le s} p(i)}{s}\right)^2\right] \\ =& \frac{1}{2}(1 - \frac{1}{s}) \le \frac{1}{2}(1 - \frac{1}{r}).\end{aligned}$$

On the other hand,

$$\max_p f(p) \ge \frac{e(G)}{n^2}.$$

Combining, we have

$$e(G) \le \frac{r-1}{2r} \cdot n^2.$$

∎

Proof of Turán's theorem. We prove for a fixed $r \ge 2$. Recall that $n = tr + s$, $1 \le s \le r$. We will prove by induction on t that any K_{r+1}-free n-vertex graph G has $e(G) \le e(T_r(n))$.

The base case $t = 0$ is trivial. Now we may assume this statement holds for those G' with $t' < t$. Let G be a maximal K_{r+1}-free n-vertex graph, where $n = tr + s$, $1 \le s \le r$. Then G contains a K_r, say H. Then $\forall x \in G - V(H)$, x has at most $r-1$ neighbors in H. Consider $G' = G - V(H)$, it is a K_{r+1}-free graph with $n' = (t-1)r + s$ vertices. By induction, $e(G') \le e(T_r(n-r))$. Therefore $e(G) \le \binom{r}{2} + (n-r)(r-1) + e(T_r(n-r)) = e(T_r(n))$.

The following exercise completes the proof. ∎

Exercise 7.10. *Prove that the unique n-vertex K_{r+1}-free graph G with $e(G) = e(T_r(n))$ is $G = T_r(n)$.*

8 Partially Ordered Sets (Poset)

Definition 8.1. *Let X be a finite set. R is a (binary)* **relation** *on X, if $R \subseteq X \times X = \{(x_1, x_2) : \forall x_1, x_2 \in X\}$. If $(x, y) \in R$, then we often write xRy.*

Definition 8.2. *A* **partially ordered set** (*poset for short*) *is an ordered pair (X, R), where X is a finite set and R is a relation on X such that the following holds:*

(1) R is **reflective***: xRx for any $x \in X$,*

(2) R is **antisymmetric***: if xRy and yRx, then $x = y$,*

(3) R is **transitive***: if xRy and yRz, then xRz.*

Example 8.3. *Consider the poset $(2^{[n]}, \subseteq)$, where "$\subseteq$" denotes the inclusion relationship.*

Example 8.4. *Consider the poset $([n], |)$, where "$|$" denotes the divisibility relationship.*

We often use "$\preccurlyeq$" to replace "R". So poset $(X, R) = (X, \preccurlyeq)$ and $xRy = x \preccurlyeq y$. If $x \preccurlyeq y$ but $x \neq y$, then $x \prec y$, and we say x is a predecessor/child of y.

Definition 8.5. *Let $(X_1, \preccurlyeq_1)$ and $(X_2, \preccurlyeq_2)$ be two posets. A mapping $f : X_1 \to X_2$ is called an* **embedding** *of $(X_1, \preccurlyeq_1)$ in $(X_2, \preccurlyeq_2)$ if*

(1) f is injective,

(2) $f(x) \preccurlyeq_2 f(y)$ if and only if $x \preccurlyeq_1 y$.

Theorem 8.6. *For every poset $(X, \preccurlyeq)$ there exists an embedding of $(X, \preccurlyeq)$ in poset $(2^X, \subseteq)$.*

Proof. Consider the mapping $f : X \to 2^X$ by letting $f(x) = \{y \in X : \; y \preccurlyeq x\}$ for any $x \in X$. It suffices to verify that f is an embedding of $(X, \preccurlyeq)$ in $(2^X, \subseteq)$.

First, it's easy to check that f is injective. If $f(x) = f(y)$ for $x, y \in X$, then $x \in f(x) = f(y)$ and $x \preccurlyeq y$. Similarly we have $y \preccurlyeq x$ which implies that $x = y$.

Second, if $x \preccurlyeq y$, then clearly $f(x) \subseteq f(y)$ by transitive property. Now we suppose that $f(x) \subseteq f(y)$. Since $x \in f(x)$ and $y \in f(y)$, we have $x \preccurlyeq y$. Thus we have that f indeed is an embedding. ∎

Definition 8.7. *Let $(X, \preccurlyeq)$ be a poset. We say an element x is an* **immediate predecessor** *of y or y covers x, if*

(1) $x \prec y$,

(2) there is no element $t \in X$ such that $x \prec t \prec y$.

In this case, we write $x \lhd y$.

Fact 8.8. *For $x, y \in (X, \preccurlyeq)$, $x \prec y$ if and only if there exist $z_1, z_2, ..., z_k \in X$ such that $x \lhd z_1 \lhd z_2 \lhd \cdots \lhd z_k \lhd y$. (Note that here k can be 0, i.e., $x \lhd y$.)*

Proof. (⇐) This direction is trivial by transitive property.

(⇒) Let $x \prec y$. Let $M_{xy} = \{t \in X : x \prec t \prec y\}$. We prove the statement by induction on $|M_{xy}|$.

Base case is clear, since if $|M_{xy}| = 0$, then $x \lhd y$. Now we may assume that $M_{xy} \neq \emptyset$ and the statement holds for any $u \prec v$ with $|M_{uv}| < n$. Suppose $x \prec y$ with $|M_{xy}| = n \geqslant 1$. Pick any $t \in M_{xy}$ and consider M_{xt} and M_{ty}. Clearly $M_{xt} \subsetneq M_{xy}$ and $M_{ty} \subsetneq M_{xy}$ because of transitive property. By inductional assumption, there exist $x_1, x_2, ..., x_m \in X$ and $y_1, y_2, ..., y_l \in X$ such that $x \lhd x_1 \lhd x_2 \lhd \cdots \lhd x_m \lhd t$ and $t \lhd y_1 \lhd y_2 \lhd \cdots \lhd y_l \lhd y$. Thus, we have $x \lhd x_1 \lhd x_2 \lhd \cdots \lhd x_m \lhd t \lhd y_1 \lhd \cdots \lhd y_l \lhd y$ and we are done. ∎

Now we can express a poset in a diagram.

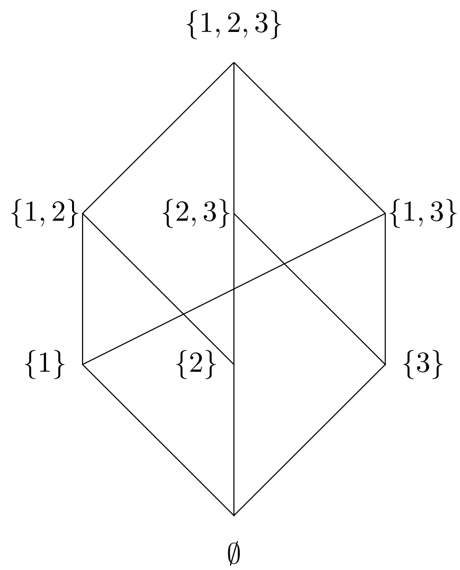
Definition 8.9. *The* **Hasse diagram** *of a poset* $(X, \preccurlyeq)$ *is a drawing in the plane such that*

(1) each element of X *is drawn as a nod in the plane,*

(2) each pair $x \lhd y$ *is connected by a line segment,*

(3) if $x \lhd y$*, then the nod* x *must appear lower in the plane than the nod* y*.*

Example 8.10. *The Hasse diagram of poset* $(2^{[3]}, \subseteq)$ *is as follows.*



The fact that $x \prec y$ if and only if $x \lhd x_1 \lhd x_2 \lhd \cdots \lhd x_k \lhd y$ now can be restated as follows: $x \prec y$ if and only if we can find a path in the Hasse diagram from nod x to nod y, strictly from bottom to top.

Definition 8.11. *Let* $P = (X, \preccurlyeq)$ *be a poset.*

(1) For distinct $x, y \in X$*, if* $x \prec y$ *or* $y \prec x$*, then we say that* x, y *are* **comparable***; otherwise,* x, y *are* **incomparable**.

(2) The set $A \subseteq X$ *is an* **antichain** *of* P*, if any two elements in* A *are incomparable. Let* $\boldsymbol{\alpha(P)}$ *be the maximum size of an antichain of* P*.*

(3) *The set $B \subseteq X$ is a* **chain** *of P, if any two elements of B are comparable. Let* $\boldsymbol{\omega(P)}$ *be the maximum size of a chain of P.*

Consider the Hasse diagram, $\omega(P)$ means the maximum number of vertices in a path (from bottom to top) in this diagram. So $\omega(P)$ is also called the height of P and $\alpha(P)$ is called the width of P similarly.

Definition 8.12. *An element $x \in X$ is* **minimal** *in $P = (X, \preccurlyeq)$, if x has no predecessor in P.*

Fact 8.13. *The set of all minimal elements of $P = (X, \preccurlyeq)$ forms an antichain of P.*

Theorem 8.14. *For any poset $P = (X, \preccurlyeq)$, $\alpha(P) \cdot \omega(P) \geq |X|$.*

Proof. We inductively define a sequence of posets $P_i = (X_i, \preccurlyeq)$ and a sequence of sets $M_i \subset P_i$ such that each M_i is the set of minimal elements of P_i, and $X_i = X - \bigcup_{j=0}^{i-1} M_j$.

First, set $P_1 = P = (X, \preccurlyeq), X_1 = X$ and $M_0 = \emptyset$. Assume posets $P_i = (X_i, \preccurlyeq)$ and M_{i-1} are defined for all $1 \leqslant i \leqslant k$, where k is big enough. Let $M_i = \{$ all minimal elements of $P_i\}$ and let $X_{i+1} = X - M_1 \bigcup \cdots \bigcup M_i$. Then let P_{i+1} be the subposet of P restricted on X_{i+1}. We keep doing this until $X_{\ell+1} = \emptyset$. By Fact 8.13, each M_i is an antichain of P_i. Since P_i is the restricted subposet of P on X_i, M_i is also an antichain of P. So we have

$$|M_i| \leq \alpha(P).$$

It suffices to find a chain $x_1 \prec x_2 \prec \cdots \prec x_\ell$ in P, such that $x_i \in P_i = (X_i, \preccurlyeq)$ for $i \in [\ell]$. Indeed, if this holds, then

$$X = M_1 \bigcup M_2 \bigcup \cdots \bigcup M_\ell \quad \text{and} \quad |X| = \sum_{i=1}^{\ell} |M_i| \leq \alpha(P) \cdot \ell \leq \alpha(P) \cdot \omega(P).$$

In fact, by the definition of M_i, we can claim something stronger holds: For any $x \in M_i$ $(2 \leq i < \ell)$, there exists $y \in M_{i-1}$, such that $y \prec x$. This completes the proof. ∎

Corollary 8.15. *Consider a poset $(X, \prec)$. If $|X| = rs + 1$ where r, s are positive integers, there exists a chain of szie $s + 1$ or an antichain of size $r + 1$.*

Definition 8.16. *Consider a sequence $X = (x_1, x_2, ..., x_n)$ of n real numbers. A subsequence $(x_{i_1}, x_{i_2}, ..., x_{i_m})$ of X, where $i_1 < i_2 < \cdots < i_m$, is* **monotone**, *if either $x_{i_1} \leq x_{i_2} \leq \cdots \leq x_{i_m}$ or $x_{i_1} \geq x_{i_2} \geq \cdots \geq x_{i_m}$.*

For example, $(10, 9, 7, 5, 1)$ is a monotone subsequence of $(10, 9, 7, 4, 5, 1, 2, 3)$.

Theorem 8.17 (Erdős-Szekeres Theorem)**.** *For any sequence $(x_1, x_2, ..., x_{n^2+1})$ of length $n^2 + 1$, there exists a monotone subsequence of length $n + 1$.*

Proof. Let $X = [n^2 + 1]$. We define a poset $P = (X, \preccurlyeq)$ as follows: $i \preccurlyeq j$ if and only if $i \leq j$ and $x_i \leq x_j$.

It is easy to check that $P = (X, \preccurlyeq)$ indeed defines a poset. By the previous result that $\alpha(P) \cdot w(P) \geq |X| = n^2 + 1$, we have either $w(P) \geq n + 1$ or $\alpha(P) \geq n + 1$.

Case 1. $w(P) \geq n + 1$.

There exists a chain of size $n+1$, which we say $\{x_{i_1}, x_{i_2}, .., x_{i_{n+1}}\}$ such that $i_1 \leq i_2 \leq \cdots \leq i_{n+1}$. By definition, we have $x_{i_1} \leq x_{i_2} \leq \cdots \leq x_{i_{n+1}}$ which is an increasing subsequence of length $n+1$.

Case 2. $\alpha(P) \geq n+1$.

There exists an antichain of size $n+1$, which we say $\{x_{i_1}, x_{i_2}, .., x_{i_{n+1}}\}$ such that $i_1 \leq i_2 \leq \cdots \leq i_{n+1}$. By definition, we have $x_{i_1} > x_{i_2} > \cdots > x_{i_{n+1}}$ which is a decreasing subsequence of length $n+1$. ∎

Remark 8.18. *What we proved is a bit stronger: there is either an increasing subsequence of length $n+1$ or a strictly decreasing subsequence of length $n+1$.*

Exercise 8.19. *Find examples to show that Erdős-Szekeres Theorem is optimal: there exists a sequence of n^2 reals such that there is no monotone subsequence of length $n+1$.*

Theorem 8.20 (Dilworth's Theorem)**.** *Let $P = (X, \preccurlyeq)$ be a poset. Then the minimum number m of disjoint chains which together contain all elements of X is equal to $\alpha(P)$. (that is the minimal m such that $X = \cup_{i=1}^{m} c_i$, c_i are pairwise disjoint chains in P).*

Proof. In a disjoint chain decomposition $X = \cup_{i=1}^{m} c_i$ and a max antichain A, we have $|c_i \cap A| \leq 1$. Thus $m \geq |A| = \alpha(P)$.

Let $\alpha = \alpha(P)$. We use induction on $|X|$ to show that $m \leq \alpha$. The base case is $|X| = 0$, it is obvious that $m \leq \alpha$.

Let C be a fixed maximum chain in $(X, \preccurlyeq)$. If every antichain in $(X - C, \preccurlyeq)$ contains at most $\alpha' \leq \alpha - 1$ elements, then by induction there exists $X - C = \cup_{i=1}^{\alpha'} c_i$ where c_i's are pairwise disjoint chains in $(X - C, \preccurlyeq)$. Thus $X = C \cup [\cup_{i=1}^{\alpha'} c_i]$ has at most $\alpha' + 1 \leq \alpha$ chains.

Hence, we may assume that $\{a_1, a_2, \ldots, a_\alpha\}$ is an antichain in $(X - C, \preccurlyeq)$. Now we define $S^- = \{x \in X : \text{if } \exists a_i \text{ such that } x \preccurlyeq a_i\}$, $S^+ = \{y \in X : \text{if } \exists a_j \text{ such that } a_j \preccurlyeq y\}$. We claim that $S^- \cup S^+ = X$. If it is not the case then there exists $x \in X$ such that $x \notin S^-$, $x \notin S^+$, $\{a_1, a_2, \ldots, a_\alpha, x\}$ is an antichain of size $\alpha+1$, which leads to contradiction. Since $\{a_1, a_2, \ldots, a_\alpha\}$ is an antichain, $\{a_1, a_2, \ldots, a_\alpha\} \subseteq S^- \cap S^+$. Therefore $\alpha(S^-, \preccurlyeq) = \alpha = \alpha(S^+, \preccurlyeq)$. Since C is a maximal chain, the largest element of C is not in S^-. Therefore $|S^-| < |X|$.

Since $\alpha(S^-, \preccurlyeq) = \alpha$, by induction, S^- is the union of $\alpha = \alpha(S^-, \preccurlyeq)$ disjoint chains $s_1^-, s_2^-, \ldots, s_\alpha^-$, where $a_i \in s_i^-$.

Claim: a_i is the maximal element of the chain s_i^-.

Proof of Claim. Otherwise $\exists x \in s_i^-$ with $a_i \preccurlyeq x$. By definition of S^-, $\exists a_j$ such that $x \preccurlyeq a_j$, thus $a_i \preccurlyeq x \preccurlyeq a_j$, a contradiction.

Now we can do the same for S^+, S^+ is the union of $\alpha = \alpha(S^-, \preccurlyeq)$ disjoint chains $s_1^+, s_2^+, \ldots, s_\alpha^+$, where $a_i \in s_i^+$ and a_i is the minimal element in s_i^+.

We can combine s_i^- and s_i^+ to get a chain $s_i^- \cup s_i^+$ in $(X, \preccurlyeq)$. As $S^- \cup S^+ = X$, X is the union of α disjoint chains $s_i^- \cup s_i^+$ $(1 \leq i \leq \alpha)$, thus $m \leq \alpha = \alpha(P)$. ∎

Dilworth's Theorem has a dual version:

Theorem 8.21 (Mirsky's Theorem)**.** *Let $P = (X, \preccurlyeq)$ be a poset. If $P = (X, \preccurlyeq)$ contains no chain of $m+1$ elements, then X is the union of m disjoint antichains.*

Exercise 8.22. *Prove Mirsky's Theorem by a similar way as the proof in Theorem 8.14.*

9 The Probabilistic Method

9.1 Basics

Definition 9.1. *A* **probability space** *is a pair* (Ω, P)*, where* Ω *is a finite set and* $P : 2^{\Omega} \to [0,1]$ *is a function assigning a number in the interval* $[0,1]$ *to every subset of* Ω *such that*

(i) $P(\emptyset) = 0$*,*

(ii) $P(\Omega) = 1$*, and*

(iii) $P(A \cup B) = P(A) + P(B)$ *for disjoint sets* $A, B \subset \Omega$*.*

Definition 9.2. *Any subset* A *of* Ω *is called an* **event***, and* $P(A) = \sum_{\omega \in \Omega} P(\{\omega\})$*.*

Definition 9.3. *A* **random variable** *is a function* $X : \Omega \to R$*.*

Definition 9.4. *The* **expectation** *of a random variable* X *is:*

$$E[X] := \sum_{\omega \in \Omega} P(\{\omega\}) \cdot X(\omega).$$

Fact 9.5 (Union bound)**.** $P(A \cup B) \leq P(A) + P(B)$ *for any* $A, B \subset \Omega$*.*

Fact 9.6 (The linearity of expectations)**.** *for any two random variables* X *and* Y *on* Ω*, we have*

$$E[X + Y] = E[X] + E[Y].$$

9.2 Union Bound

Now we discuss the following basic form of the probabilistic methods in Combinatorics:

(i) Imagine we need to find some combinatorial object satisfying certain property, call it a "good" property. We consider a big family for candidates and randomly pick one from this family and call it a random object. If the probability that the random object has "good" property is positive, then there must exist "good" objects.

(ii) To compute the probability of being "good", we often compute the probability of being "bad" and aim to show that this probability of being "bad" is strictly less than 1.

Theorem 9.7. *Let* n, s *satisfy* $\binom{n}{s} \cdot 2^{1-\binom{s}{2}} < 1$*. Then* $R(s,s) > n$*.*

Proof. We need to find a 2-edge-coloring of K_n such that it has no monochromatic clique K_s.

Let Φ be the family of all 2-edge-colorings of K_n. Let $c \in \Phi$ be chosen uniformly at random. Then c is a random 2-edge-coloring of K_n, where each edge of K_n is colored by red and blue, each with probability $\frac{1}{2}$, independent of each other edge.

Let B be the event that this random 2-edge-coloring has no monochromatic K_s. We want to prove $P(B) > 0$. Consider its complement event $A = \Phi \backslash B$ and its probability $P(A)$, where A is the event that c has a monochromatic K_s. Next, we compute $P(A)$ for any $S \in \binom{[n]}{s}$. Let A_S be the event that S forms a monochromatic K_s for c. Then we have $A = \cup_{S \in \binom{[n]}{s}} A_S$, and $P(A_S) = 2^{1-\binom{s}{2}}$.

So by the union bound,

$$P(A) = P\left(\cup_{S\in\binom{[n]}{s}} A_S\right) \le \sum_{S\in\binom{[n]}{s}} P(A_S) = \binom{n}{s} 2^{1-\binom{s}{2}} < 1.$$

This shows that $P(B) > 0$. ∎

Corollary 9.8. $R(s,s) \ge \frac{1}{e\sqrt{2}} s 2^{\frac{s}{2}}$.

Proof. Let $n = \frac{1}{e\sqrt{2}} s 2^{\frac{s}{2}} \left(\frac{e}{2}\right)^{1/s}$. Recall that $\binom{n}{s} < \frac{n^s}{s!}$ and $s! \ge e\left(\frac{s}{e}\right)^s$, thus we have that

$$\binom{n}{s} 2^{1-\binom{s}{2}} < \frac{n^s}{e\left(\frac{s}{e}\right)^s} 2^{1-\binom{s}{2}} = 1.$$

So by the above theorem, we get

$$R(s,s) > n = \frac{1}{e\sqrt{2}} s 2^{\frac{s}{2}} \left(\frac{e}{2}\right)^{1/s} \ge \frac{1}{e\sqrt{2}} s 2^{\frac{s}{2}}.$$

∎

Definition 9.9. *The* **random graph $G(n,p)$** *for some real $p \in (0,1)$ is a graph with vertex set $\{1,2,...,n\}$, where each of potential $\binom{n}{2}$ edges appears with probability p, independent of other edges.*

In the proof of the previous theorem, in fact we consider $G(n,1/2)$.

Let A be the property we are interested in. Let

$$\begin{aligned} P(A) &= P(G(n,\frac{1}{2}) \text{ satisfies the property } A) \\ &= \frac{\text{the number of graphs with vertex set } [n] \text{ satisfying the property } A}{2^{\binom{n}{2}}}. \end{aligned}$$

Then $P(A)$ is a function of n, taking value in $[0,1]$.

Definition 9.10. *We say the random graph $G(n,\frac{1}{2})$* **almost surely** *satisfies property A, if*

$$\lim_{n\to+\infty} P(A) = 1.$$

If $\lim_{n\to+\infty} P(A) = 0$, then $G(n,\frac{1}{2})$ almost surely does not satisfy the property A.

Theorem 9.11. *Random graph $G(n,\frac{1}{2})$ almost surely is not bipartite.*

Proof. Let A be the event that $G(n,\frac{1}{2})$ is bipartite. For any $U \subseteq [n]$, let A_U be the event that all edges of G are between U and $[n]\backslash U$. Then we know $A = \bigcup_{U\subseteq[n]} A_U$. We have

$$P(A_U) = \frac{\text{the number of graphs satisfying } A_U}{2^{\binom{n}{2}}} = \frac{2^{|U|(n-|U|)}}{2^{\binom{n}{2}}} \le \frac{2^{\frac{n^2}{4}}}{2^{\frac{n(n-1)}{2}}} = 2^{-\frac{n^2}{4}+\frac{n}{2}}.$$

So by the union bound,

$$0 \le P(A) = P(\bigcup_{U\subseteq[n]} A_U) \le \sum_{U\subseteq[n]} P(A_U) \le 2^n \cdot 2^{-\frac{n^2}{4}+\frac{n}{2}} = 2^{-\frac{n^2}{4}+\frac{3n}{2}}.$$

Thus we have $\lim_{n\to+\infty} P(A) = 0$. ∎

Definition 9.12. *Given a probability space (Ω, P), we say events $A_1, A_2, ..., A_k$ are* **independent** *if for any $I \subset [n]$, we have $P(\bigcap_{i\in I} A_i) = \prod_{i\in I} P(A_i)$.*

Definition 9.13. *An* **arc** *is an edge in a directed graph.*

Definition 9.14. *A* **tournament** *on n vertices is a directed graph obtained from the clique K_n by assigning a direction to each edge of K_n. For any arc $i \to j$, we say i is the* **head** *and j is the* **tail** *of the arc.*

Definition 9.15. *A tournament T satisfies the* **property S_k** *if for any subset A of size k, there exists a vertex $u \in V(T) \setminus A$ such that $u \to x$ for any $x \in A$.*

Question 9.16. *For any $k \in Z^+$, can we find a tournament satisfying the property S_k?*

Theorem 9.17. *For any $k \in Z^+$, if $\binom{n}{k}(1 - \frac{1}{2^k})^{n-k} < 1$, then there exists a tournament on n vertices satisfying the property S_k.*

Proof. We prove this by considering a random tournament T on $[n]$, that is, for any pair $\{i, j\}$, the arc $i \to j$ occurs with probability $\frac{1}{2}$, independent of other choices. Let B be the event that T does not satisfy the property S_k. For $A \in \binom{[n]}{k}$, let B_A be the event that for every vertex $x \in [n] \setminus A$ there exists some $u \in A$ with $u \to x$. So

$$B = \bigcup_{A \in \binom{[n]}{k}} B_A.$$

For $x \in [n]\setminus A$, let $B_{A,x}$ be the event that there exists some $u \in A$ with $u \to x$. So

$$B_A = \bigcap_{x\in[n]\setminus A} B_{A,x}.$$

It is easy to see that for any $x \in [n] \setminus A$

$$P(B_{A,x}) = 1 - \left(\frac{1}{2}\right)^k.$$

Note that only the arcs between x and A will effect the event $B_{A,x}$, and these arcs for distinct vertices x's are disjoint. This explains that all events $B_{A,x}$ for all $x \in [n] \setminus A$ are independent. So

$$P(B_A) = P(\bigcap_{x\notin A} B_{A,x}) = \prod_{x\notin A} P(B_{A,x}) = \left(1 - \left(\frac{1}{2}\right)^k\right)^{n-k}.$$

Therefore,

$$P(B) \le \sum_{A\in\binom{[n]}{k}} P(B_A) \le \binom{n}{k}\left(1 - \left(\frac{1}{2}\right)^k\right)^{n-k} < 1.$$

Thus, $P(B^c) > 0$, i.e., there exists a tournament on $[n]$ satisfying property S_k.

∎

Corollary 9.18. *For any $k \in Z^+$, there exists a minimal $f(k)$ such that there exists a tournament on $f(k)$ vertices satisfying the property S_k.*

Example 9.19. *We have $f(3) \le 91$, as $\binom{91}{3}(\frac{7}{8})^{88} < 1$.*

9.3 The Linearity of Expectation

- For any two variables X, Y, we have $E[X+Y] = E[X] + E[Y]$.
- $P(X \geq E[X]) > 0$.
- $P(X \leq E[X]) > 0$.

Definition 9.20. *A set A is* **sum-free**, *if for any $x, y \in A, x + y \notin A$, i.e., $x + y = z$ has no solutions in A.*

Example: Both $\{\lfloor \frac{n}{2} \rfloor + 1, \lfloor \frac{n}{2} \rfloor + 2, ..., n\}$ and $\{\text{all odd integers in } [n]\}$ are two sum-free sets in $[n]$ of size $\lceil \frac{n}{2} \rceil$.

Exercise 9.21. *Show that the maximum size of a sum-free subset A in $[n]$ is $\lceil \frac{n}{2} \rceil$.*

Theorem 9.22. *For any set A of non-zero integers, there exists a sum-free subset $B \subseteq A$ with $|B| \geq \frac{|A|}{3}$.*

Proof. We choose a prime p large enough such that $p > |a|$ for any $a \in A$. Consider $Z_p = \{0, 1, ..., p-1\}$ and $Z_p^* = \{1, 2, ..., p-1\}$. We note that there is a large sum-free subset under $Z_p($ mod $p)$:

$$S = \Big\{ \lceil \frac{p}{3} \rceil + 1, \lceil \frac{p}{3} \rceil + 2, ..., \lceil \frac{2p}{3} \rceil \Big\}.$$

Claim: For any $x \in Z_p^*$, $A_x = \{a \in A : ax \pmod{p} \in S\}$ is sum-free.

Proof. Suppose that there are $a, b, c \in A_x$ satisfying $a + b = c$. But we also have $ax \pmod p \in S$, $bx \pmod p \in S$, $cx \pmod p \in S$ and $ax \pmod p + bx \pmod p = cx \pmod p$ in Z_p. This is a contradiction to that S is sum-free in Z_p. ∎

Next, we want to find some $x \in Z_p^*$ such that $|A_x| \geq \frac{|A|}{3}$. We choose $x \in Z_p^*$ uniformly at random, and we compute, $E\big[|A_x|\big]$, the expectation of $|A_x|$.

Note that $|A_x| = \sum\limits_{a \in A} 1_{\{ax \pmod p \in S\}}$. So

$$E\big[|A_x|\big] = E\Big[\sum_{a \in A} 1_{\{ax \pmod p \in S\}}\Big] = \sum_{a \in A} E\big[1_{\{ax \pmod p \in S\}}\big] = \sum_{a \in A} P\big(ax \pmod p \in S\big).$$

We observe that for a fixed $a \in A$, $\{ax : x \in Z_p^*\} = Z_p^*$. So $P\big(ax \pmod p \in S\big) = \frac{|S|}{|Z_p^*|} \geq \frac{1}{3}$. And thus, $E\big[|A_x|\big] \geq \sum\limits_{a \in A} \frac{1}{3} = \frac{|A|}{3}$. Then, we know that there exists a choice of $x \in Z_p^*$ such that $|A_x| \geq E\big[|A_x|\big] \geq \frac{|A|}{3}$. ∎

Definition 9.23. *Given a graph G, a* **dominating set** *A in G is a subset of $V(G)$ such that any $u \in V(G) \backslash A$ has a neighbor in A.*

Theorem 9.24. *Let G be a graph on n vertices and with minimum degree $\delta > 1$. Then G contains a dominating set of at most $\dfrac{1 + \ln(1+\delta)}{1+\delta} n$ vertices.*

Proof. Take $p \in (0,1)$, whose value will be determined later. We pick each vertex in $V(G)$ with probability p uniformly at random. Let A be the set of those chosen vertices. Let B be the set of vertices $b \in V(G) \setminus A$, which has no neighbors in A. Then we can see that

- $A \cup B$ is a dominating set in G.
- $b \in B$ if and only if $\big(\{b\} \cup N_G(b)\big) \cap A = \emptyset$.

That is, $b \in B$ if and only if b and all neighbors of b are not picked. So

$$P(b \in B) = (1-p)^{1+d_G(b)} \leq (1-p)^{1+\delta} \leq e^{-p(1+\delta)},$$

where the last inequality holds since $1 + x \leq e^x$. Then, we have

$$E\big[|B|\big] = E\Big[\sum_{b \in V(G)} 1_{\{b \in B\}}\Big] = \sum_{b \in V(G)} P(b \in B) \leq n \cdot e^{-p(1+\delta)}.$$

We also have $E\big[|A|\big] = np$. Thus,

$$E\big[|A \cup B|\big] \leq E\big[|A| + |B|\big] = E\big[|A|\big] + E\big[|B|\big] \leq n\big(p + e^{-p(1+\delta)}\big).$$

By calculus, we see that when $p = \dfrac{\ln(1+\delta)}{1+\delta}$, $p + e^{-p(1+\delta)}$ is minimized with value $\dfrac{1+\ln(1+\delta)}{1+\delta}$. So we pick $p = \dfrac{\ln(1+\delta)}{1+\delta}$ to get $E\big[|A \cup B|\big] \leq \dfrac{1+\ln(1+\delta)}{1+\delta} n$. Therefore there exists a choice of $A \cup B$ such that $|A \cup B| \leq E\big[|A \cup B|\big] \leq \dfrac{1+\ln(1+\delta)}{1+\delta} n$, where $A \cup B$ is a dominating set of G. ∎

Definition 9.25. *Let $\boldsymbol{\alpha(G)}$ be the maximum size of an independent set in G.*

Theorem 9.26. *For any graph G, $\alpha(G) \geqslant \sum_{v \in V(G)} \dfrac{1}{d(v)+1}$ where $d(v)$ denotes the degree of v in G.*

Proof. Let $V(G) = [n]$. For $i \in [n]$, let N_i be the neighborhood of i in G. Let S_n be the family of all permutations $\pi : [n] \to [n]$.

Given a permutation $\pi \in S_n$, we say a vertex $i \in [n]$ is $\pi-$good, if $\pi(i) < \pi(j)$ for any $j \in N_i$. Let M_π be the set of all π-good vertices.
Claim: For any $\pi \in S_n$, M_π is an independent set in G.

Proof. Suppose that there are two vertices $i, j \in M_\pi$ with $ij \in E(G)$. Let $\pi(i) < \pi(j)$. Then j is not π-good, a contradiction. ∎

We pick an $\pi \in S_n$ uniformly at random, and compute $E\big[|M_\pi|\big]$. Since $|M_\pi| = \sum_{i \in [n]} 1_{\{i \text{ } is \text{ } \pi\text{-good}\}}$, we have $E\big[|M_\pi|\big] = \sum_{i \in [n]} P(i \text{ } is \text{ } \pi\text{-good}) = \sum_{i \in [n]} \dfrac{1}{d(i)+1}$. Thus there exists a permutation $\pi \in S_n$ such that $|M_\pi| \geq \sum_{i \in [n]} \dfrac{1}{d(i)+1}$. Then by the definition of $\alpha(G)$ and our claim, we can get that $\alpha(G) \geq \sum_{v \in V(G)} \dfrac{1}{d(v)+1}$ which completes the proof. ∎

Corollary 9.27. *For any graph G with n vertices and m edges, we have $\alpha(G) \geq \frac{n^2}{2m+n}$.*

Proof. Exercise. ∎

Corollary 9.28. *For any graph G with n vertices and average degree d (i.e., $d = \frac{2m}{n}$), then $\alpha(G) \geq \frac{n}{1+d}$.*

Definition 9.29. Turán graph $T_r(n)$ *on r parts is an n-vertex graph G such that $V(G) = V_1 \cup V_2 \cup ... \cup V_r$ and $|V_1| \leq |V_2| \leq ... \leq |V_r| \leq |V_1| + 1$, where $ab \in E(G)$ if and only if $a \in V_i$ and $b \in V_j$ for some $i \neq j$.*

$T_r(n)$ is a balanced complete r-partite graph.

Theorem 9.30 (Turán's Theorem approximate form)**.** *If G is K_{r+1}-free, then $e(G) \leqslant \dfrac{r-1}{2r}n^2$.*

Second proof. Using Corollary 9.28 (Exercise). ∎

9.4 The Deletion Method

Earlier, we often define an appropriate probability space and show the random event occurs with positive probability.

Today, we extend this idea and consider situations where random events do not always have the desired property, and may have very few "blemishes". The point that we want to make here is that after deleting all blemishes, we will obtain the desired property.

Theorem 9.31. *Let G be a graph on n vertices and with average degree $d \geq 1$. Then $\alpha(G) \geq \frac{n}{2d}$.*

Proof. Let $S \subseteq V(G)$ be a random subset, where for any $v \in V(G)$, $\mathbb{P}(v \in S) = p$. Here the value of p will be determined later.

Let X denote the number of vertices in S, and let Y denote the number of edges of G, both ends of which lie in S. Then $\mathbb{E}[X] = np$, and $\mathbb{E}[Y] = e(G)p^2 = \frac{nd}{2}p^2$. By taking $p = \frac{1}{d}$, we have

$$\mathbb{E}[X - Y] = \mathbb{E}[X] - \mathbb{E}[Y] = \frac{n}{d} - \frac{n}{2d} = \frac{n}{2d}.$$

So there is a subset $S \subseteq V(G)$ such that

$$|S| - e(G[S]) \geq \mathbb{E}[X - Y] = \frac{n}{2d}.$$

Now we delete one vertex for each edge of $G[S]$. This leaves a subset $S^* \subseteq S$ of size at least $|S| - e(G[S]) \geq \frac{n}{2d}$. Since all edges of $G[S]$ are destroyed, S^* must be an independent set. ∎

Recall: If $\binom{n}{k}2^{1-\binom{k}{2}} < 1$, then the Ramsey number $R(k,k) > n$. So $R(k,k) > \frac{1}{e\sqrt{2}}k2^{\frac{k}{2}}$.

Theorem 9.32. *For all integer n, we have $R(k,k) > n - \binom{n}{k}2^{1-\binom{k}{2}}$.*

Proof. Consider a random 2-edge-coloring of K_n, where each edge is colored by red or blue with probability $\frac{1}{2}$, independent of other choices. For $A \in \binom{[n]}{k}$, let X_A be the indicator random variable of the event that A induces a monochromatic K_k. Then $\mathbb{E}[X_A] = 2 \cdot \left(\frac{1}{2}\right)^{\binom{k}{2}}$.

Let $X = \sum_{A \in \binom{[n]}{k}} X_A$ be the number of monochromatic k-subsets. Then we have

$$\mathbb{E}[X] = \sum_{A \in \binom{[n]}{k}} \mathbb{E}[X_A] = \binom{n}{k} 2^{1-\binom{k}{2}}.$$

So there exists a 2-edge-coloring of K_n whose number of monochromatic k-subsets is at most $\mathbb{E}[X] = \binom{n}{k}2^{1-\binom{k}{2}}$. Next, we remove one vertex from each monochromatic k-subset. This will delete at most $X \leq \mathbb{E}[X] \leq \binom{n}{k}2^{1-\binom{k}{2}}$ vertices and will destroy all monochromatic k-subsets. So it remains at least $n - \binom{n}{k}2^{1-\binom{k}{2}}$ vertices, which contains NO monochromatic k-subsets. ∎

Corollary 9.33.

$$R(k,k) > \frac{1}{e}(1+o(1))k2^{\frac{k}{2}}.$$

Proof. We leave the proof of this corollary as an exercise. ∎

Exercise 9.34. *Prove Corollary 9.33 by maximizing $n - \binom{n}{k}2^{1-\binom{k}{2}}$ for a fixed k.*

9.5 Markov's inequality

Theorem 9.35 (Markov's Inequality)**.** *Let $X \geq 0$ be a random variable and $t > 0$, then*

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}.$$

Corollary 9.36. *Let $X_n \geq 0$ be an integer-value random variable for $n \in \mathbb{N}^+$ in (Ω_n, P_n). If*

$$\mathbb{E}[X_n] \to 0 \text{ as } n \to +\infty,$$

then

$$\mathbb{P}(X_n = 0) \to 1 \text{ as } n \to +\infty.$$

That is, $X_n = 0$ almost surely occurs.

Lemma 9.37. *For a random graph $G(n,p)$ where $p \in (0,1)$, we have*

$$\mathbb{P}\left(\alpha(G(n,p)) \leq \lceil \frac{2\ln n}{p} \rceil \right) \to 1 \quad as \quad n \to +\infty.$$

Note that here p can be a function of n.

Proof. Let $t = \lceil \frac{2\ln n}{p} \rceil$. Let X_n be the number of independent sets of size $t+1$ in $G(n,p)$. For any $S \in \binom{[n]}{t+1}$, let X_S be the indicator random variable of the event that S is an independent set in $G(n,p)$. Then

$$X_n = \sum_{S \in \binom{[n]}{t+1}} X_S$$

and

$$\begin{aligned}\mathbb{E}[X_n] &= \sum_{S\in\binom{[n]}{t+1}} \mathbb{E}[X_S] = \sum_{S\in\binom{[n]}{t+1}} (1-p)^{\binom{t+1}{2}} = \binom{n}{t+1}(1-p)^{\binom{t+1}{2}} \\ &\le \frac{n^{t+1}}{(t+1)!}e^{-p\binom{t+1}{2}} = \frac{1}{(t+1)!}(ne^{-p\cdot\frac{t}{2}})^{t+1} \\ &\le \frac{1}{(t+1)!} \to 0 \text{ as } n \to +\infty.\end{aligned}$$

By Corollary 9.36, we see that $\mathbb{P}(\alpha(G(n,p)) \le t) = \mathbb{P}(X_n = 0) \to 1$ as $n \to +\infty$. ∎

Definition 9.38. *For a graph G, the* **chromatic number** *$\chi(G)$ is the minimum integer k such that $V(G)$ can be partitioned into k independent sets.*

Fact 9.39. *(1) $\chi(K_n) = n$;*
(2) $\chi(C_{2n+1}) = 3$;
(3) $\chi(G) \le 2$ if and only if G is bipartite.

Proposition 9.40. *For any graph G on n vertices, we have $\chi(G)\cdot\alpha(G) \ge n$.*

Definition 9.41. *The* **girth** *$g(G)$ of a graph G is the length of the shortest cycle in G.*

Theorem 9.42 (Erdős)**.** *For any fixed $k \in \mathbb{N}^+$, there exists a graph G with $\chi(G) \ge k$ and $g(G) \ge k$.*

Proof. Consider a random graph $G = G(n,p)$, where p will be determined later. Let $t = \lceil \frac{2\ln n}{p} \rceil$, then by Lemma 9.37 we have $\mathbb{P}(\alpha(G) \le t) \to 1$ as $n \to +\infty$.

Let X_n be the total number of cycles of length less than k in G. Then

$$\mathbb{E}[X_n] = \sum_{i=3}^{k-1} \frac{n(n-1)\cdots(n-i+1)}{2i} p^i,$$

where $\frac{n(n-1)\cdots(n-i+1)}{2i}$ is the number of $C_i's$ in K_n. So

$$\mathbb{E}[X_n] \le \sum_{i=3}^{k-1} (np)^i = \frac{(np)^k - 1}{np - 1}.$$

By Markov's inequality,

$$\mathbb{P}(X_n > \frac{n}{2}) \le \frac{E[X_n]}{n/2} \le \frac{2((np)^k - 1)}{n(np-1)}.$$

Let $p = n^{-\frac{k-1}{k}}$. So $np = n^{\frac{1}{k}}$. Then

$$\mathbb{P}(X_n > \frac{n}{2}) \le \frac{2(n-1)}{n(n^{\frac{1}{k}} - 1)} \to 0 \quad as \quad n \to +\infty.$$

Let n be sufficiently large then there exists a graph G on n vertices with $X_n \le \frac{n}{2}$ and $\alpha(G) \le t = \lceil \frac{2\ln n}{p} \rceil \le 3n^{\frac{k-1}{k}} \ln n$.

By deleting one vertex from each cycle of length at most $k-1$, we can find an induced subgraph G^* of G, which has at least $\frac{n}{2}$ vertices and has NO cycles of length at most $k-1$. Hence $g(G^*) \geq k$. As G^* is an induced subgraph of G, we have

$$\alpha(G^*) \leq \alpha(G) \leq 3n^{\frac{k-1}{k}} \ln n.$$

By Proposition 9.40, we have

$$\chi(G^*) \geq \frac{|V(G^*)|}{\alpha(G^*)} \geq \frac{\frac{n}{2}}{3n^{\frac{k-1}{k}} \ln n} = \frac{n^{\frac{1}{k}}}{6 \ln n} \geq k \text{ (as } n \text{ is large enough).}$$

Therefore G^* is the desired graph.

∎

9.6 Lovász Local Lemma

Consider "bad events" $A_1, \ldots, A_n$. We want to avoid them all.

- If all $\mathbb{P}(A_i)$ small, say $\sum_i \mathbb{P}(A_i) < 1$, then we can avoid all bad events.
- If they are all independent, then the probability that none of A_i occurs is $\prod_{i=1}^n (1 - \mathbb{P}(A_i)) > 0$ (provided that all $\mathbb{P}(A_i) < 1$). We can avoid all bad events.

The Lovász Local Lemma deals with the case when each bad event A_i is independent with most of the other events, but it could possibly be dependent with a small number of events.

Definition 9.43. *An event A_0 is* **independent** *from $\{A_1, \ldots, A_m\}$ if for any $B_i \in \{A_i, \overline{A_i}\}$*

$$\mathbb{P}(A_0 \cap B_1 \cap \cdots \cap B_m) = \mathbb{P}(A_0)\mathbb{P}(B_1 \cap \cdots \cap B_m).$$

Theorem 9.44 (Lovász Local Lemma in symmetric form)**.** *Let $\{A_1, \ldots, A_n\}$ be events with $\mathbb{P}(A_i) = p$, $\forall i \in [n]$.Suppose that each A_i is independent from a set of all other A_j except for at most d of them. If $ep(d+1) \leq 1$ (where e is the natural constant), then there is a positive probability such that none of the events A_i occur.*

Theorem 9.45 (Lovász Local Lemma in general form)**.** *Let $\{A_1, \ldots, A_n\}$ be events. For $i \in [n]$, let $N(i) \subseteq [n]$ be the set such that A_i is independent from $\{A_j : j \notin N(i) \cup \{i\}\}$. If $x_1, \ldots, x_n \in [0, 1)$ satisfy*

$$\mathbb{P}(A_i) \leq x_i \prod_{j \in N(i)} (1 - x_j) \quad \forall i \in [n],$$

then with probability no less than $\prod_{i=1}^n (1 - x_i)$ that none of the events A_i occur.

Proof of symmetric form by the general form. Suppose we have $ep(d+1) \leq 1$. Let $x_i = \frac{1}{d+1} < 1$ for all $i \in [n]$. Then

$$x_i \prod_{j \in N(i)} (1 - x_j) = \frac{1}{d+1} \prod_{j \in N(i)} (1 - \frac{1}{d+1}) \geq \frac{1}{d+1}(1 - \frac{1}{d+1})^d > \frac{1}{e(d+1)} \geq p \geq \mathbb{P}(A_i),$$

so the condition of the general form holds. Thus there is a positive probability such that none of the events A_i occur. ∎

Proof of the general form. We will prove that for any $i \notin S \subseteq [n]$,

$$\mathbb{P}(A_i | \cap_{j\in S} \overline{A_j}) \le x_i.$$

Once this is proved, we have

$$\begin{aligned}\mathbb{P}(\overline{A_1} \cap \overline{A_2} \cap, \dots, \cap \overline{A_n}) &= \mathbb{P}(\overline{A_1})\mathbb{P}(\overline{A_2}|\overline{A_1})\mathbb{P}(\overline{A_3}|\overline{A_1} \cap \overline{A_2}) \cdots \mathbb{P}(\overline{A_n}|\overline{A_1} \cap \overline{A_2}, \dots, \cap \overline{A_n - 1}) \\ &\ge (1-x_1)(1-x_2)\cdots(1-x_n).\end{aligned}$$

Now we prove this by induction on $|S|$. The base case is when $|S| = 0$. It is trivial to get $\mathbb{P}(A_i) \le x_i$. When $|S| \ge 1$, consider $i \notin S$. Let $S_1 = S \cup N(i) = \{j_1, j_2, \dots, j_r\}$, $S_2 = S \backslash N(i)$. Let $B = \cap_{j\in S_2} \overline{A_j}$, we have

$$\mathbb{P}(A_i | \cap_{j\in S} \overline{A_j}) = \frac{\mathbb{P}(A_i \cap (\cap_{j\in S_1} \overline{A_j}) | B)}{\mathbb{P}(\cap_{j\in S_1} \overline{A_j} | B)}.$$

$$\mathbb{P}(A_i \cap (\cap_{j\in S_1} \overline{A_j}))|B) \le \mathbb{P}(A_i|B) = \mathbb{P}(A_i) \le x_i \prod_{j\in N(i)} (1-x_j).$$

$$\begin{aligned}\mathbb{P}(\cap_{j\in S_1} \overline{A_j})|B) &= \mathbb{P}(\overline{A_{j_1}}|B)\mathbb{P}(\overline{A_{j_2}}|B \cap \overline{A_{j_1}}) \dots \mathbb{P}(\overline{A_{j_r}}|B \cap \overline{A_{j_1}} \cap \dots \cap \overline{A_{j_{r-1}}}) \\ &\ge \prod_{j\in S_1} (1-x_j) \ge \prod_{j\in N(i)} (1-x_j).\end{aligned}$$

Therefore,

$$\mathbb{P}(A_i | \cap_{j\in S} \overline{A_j}) = \frac{\mathbb{P}(A_i \cap (\cap_{j\in S_1} \overline{A_j}) | B)}{\mathbb{P}(\cap_{j\in S_1} \overline{A_j} | B)} \le \frac{x_i \prod_{j\in N(i)} (1-x_j)}{\prod_{j\in N(i)} (1-x_j)} = x_i.$$

∎

Theorem 9.46 (Spencer 1977)**.** *If $e(\binom{k}{2}\binom{n}{k-2} + 1)2^{1-\binom{k}{2}} < 1$, then $R(k,k) > n$.*

Proof. Random 2-edge-coloring on K_n. For each $R \in \binom{[n]}{k}$, let E_R be the event that R induces a monochromatic K_k. $\mathbb{P}(E_R) = 2^{1-\binom{k}{2}}$. E_R is independent of all E_S that satisfies $|R \cap S| \le 1$. For each R, there are at most $\binom{k}{2}\binom{n}{k-2}$ choices S with $|S| = k$ and $|R \cap S| \ge 2$.

Since $e(\binom{k}{2}\binom{n}{k-2} + 1)2^{1-\binom{k}{2}} < 1$, by Lovász Local Lemma in symmetric form $\mathbb{P}(\cap \overline{E_R}) > 0$, thus $R(k,k) > n$. ∎

Remark 9.47. *By optimizing the choice of n, we obtain*

$$R(k,k) > (\frac{\sqrt{2}}{e} + o(1))k2^{k/2}$$

This is still the best known lower bound for $R(k,k)$.

Theorem 9.48. *A k-graph is 2-colorable if every edge intersects at most $d = e^{-1}2^{k-1} - 1$ other edges.*

Proof. For each edge f, let A_f be the event that f is monochromatic (for random 2-coloring on vertices), then $\mathbb{P}(A_f) = 2^{1-k}$. Each A_f is independent from all $A_{f'}$ where $f \cap f' = \emptyset$. By condition, at most $d = e^{-1}2^{k-1} - 1$ edges intersects with f and $ep(d+1) = e2^{1-k}(e^{-1}2^{k-1}) \le 1$. By Lovász Local Lemma there is a positive probability such that none of the event A_f occurs. Therefore this k-graph is 2-colorable. ∎