Combinatorics @USTC

Lectured by Jie Ma and Xiande Zhang
Notes compiled by Xiaoshuo Lin

https://xiaoshuo-lin.github.io

Contents

Chapter 1 Enumeration

Throughout these notes, we use the following notational conventions.

- ◇ We use $[n]$ to denote the set $\{1, 2, \ldots, n\}$ for any positive integer n.
- ◇ We use $|X|$ or $\#X$ to denote the **cardinality** of a set X.
- ◇ The **factorial** of a non-negative integer n, denoted by $n!$, is the product of all positive integers less than or equal to n. The value of $0!$ is 1, according to the convention for an empty product.
- ◇ We use $(n)_k$ to denote the **falling factorial**, defined as $(n)_k = n(n-1)(n-2)\cdots(n-k+1)$.

1.1 Binomial Coefficients

Let X be a set with n elements. The power set of X, denoted by 2^X, is the set of all subsets of X. It has cardinality $\left|2^X\right| = 2^{|X|} = 2^n$.

We use $\binom{X}{k}$ to denote the set of all k-element subsets of X, and let $\binom{n}{k}$ denote its cardinality.

Fact 1.1.1 We have $\left|\binom{X}{k}\right| = \binom{n}{k} = \frac{n!}{k!(n-k)!}$ for any positive integer n and any non-negative integer k such that $k \leqslant n$.

Proof If $k = 0$, then $\binom{X}{0}$ is the set containing only the empty set, so $\left|\binom{X}{0}\right| = 1 = \binom{n}{0}$. For $k \geqslant 1$, note that the choices for an ordered k-tuple of distinct elements from X can be made in $(n)_k$ ways, and each k-element subset of X corresponds to exactly $k!$ such ordered tuples. Thus, we have

$$\left|\binom{X}{k}\right| = \frac{(n)_k}{k!} = \frac{n!}{k!(n-k)!}.$$

□

Fact 1.1.2 For $0 \leqslant k \leqslant n$, we have the following properties of binomial coefficients:

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

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

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

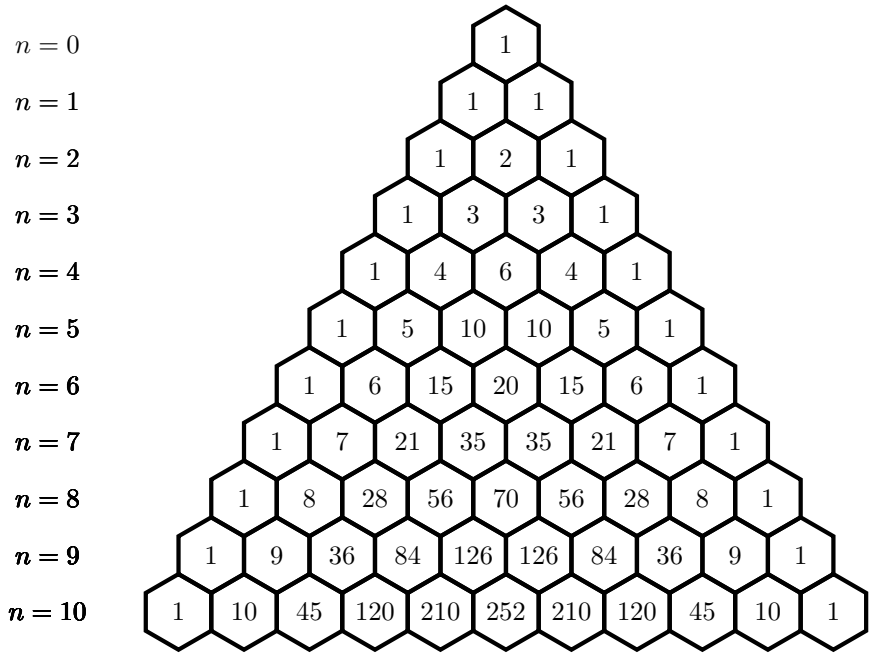
Proof Besides the algebraic proofs using Fact 1.1.1, we can also provide combinatorial interpretations for these properties:

(1) When you select k elements from a set of n elements, you are simultaneously leaving behind $n-k$ elements.

(2) The sum $\sum_{k=0}^{n} \binom{n}{k}$ counts the total number of subsets of an n-element set, which is 2^n.

(3) The last property follows from considering whether a particular element is included in a k-element subset or not. □

Remark 1.1.3 The last property is known as **Pascal's rule**. With this rule, one sees that in the n-th row of Pascal's triangle, the k-th entry is $\binom{n}{k}$.



Fact 1.1.4 The number of integer solutions to the equation $x_1 + \cdots + x_n = k$ under various constraints on x_i is given as follows:

Equation	Condition on x_i	# Solutions
$x_1 + \cdots + x_n = k$	$x_i \in \{0, 1\}$	$\binom{n}{k}$
	$x_i > 0$	$\binom{k-1}{n-1}$
	$x_i \geqslant 0$	$\binom{n+k-1}{k}$

Proof (1) This counts how many ways we can select exactly k of the n variables to be 1.

(2) Imagine k identical apples arranged in a row. We want to distribute these apples to n children so that each child gets at least one apple. To do this, place $n-1$ dividers in the gaps between apples to split the k apples into n nonempty groups. Each group corresponds to the apples given to one child.

- ◇ Since each child must get at least one apple, no group can be empty, so no two dividers can be adjacent, and no divider can be placed before the first apple or after the last apple.
- ◇ The dividers can only be placed in the $k-1$ spaces between apples.

Choosing which $n-1$ of the $k-1$ spaces to place dividers determines the distribution uniquely.

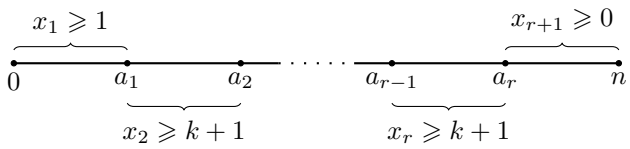
(3) Distribute k identical apples to n children, allowing some children to receive zero apples. Add n extra apples first, and give one extra apple to each child to ensure everyone has at least one. Now the total apples are $n+k$. By (2), the number of ways to distribute these $n+k$ apples so that each child has at least one is $\binom{n+k-1}{n-1} = \binom{n+k-1}{k}$. Removing the extra apples corresponds exactly to allowing children to have zero apples originally. □

The proof of (3) in Fact 1.1.4 can be rephrased as follows. Let $y_i = x_i + 1$ for all i. Then the original equation is transformed into $y_1 + \cdots + y_n = k + n$, where $y_i > 0$. Now the result follows from (2). We can apply the same idea to solve Exercise 1.1.5.

Exercise 1.1.5 Given $n \geqslant r > 1$ and $k \geqslant 0$, count the cardinality of the set

$$A = \{(a_1, \ldots, a_r) \in [n]^r : a_{i+1} - a_i \geqslant k+1, \forall i \in [r-1]\}.$$

Solution Define $a_0 = 0$, $a_{r+1} = n$, and consider $x_i \coloneqq a_i - a_{i-1}$ for $i \in [r+1]$.



The original problem is equivalent to counting the number of integer solutions to the equation

$$x_1 + \cdots + x_{r+1} = n$$

subject to the constraints $x_1 \geqslant 1$, $x_i \geqslant k+1$ for $2 \leqslant i \leqslant r$, and $x_{r+1} \geqslant 0$. If we let

$$y_i = \begin{cases} x_1, & \text{if } i = 1, \\ x_i - k, & \text{if } 2 \leqslant i \leqslant r, \\ x_{r+1} + 1, & \text{if } i = r+1, \end{cases}$$

then the equation becomes

$$y_1 + \cdots + y_{r+1} = n - k(r-1) + 1$$

with the constraints $y_i \geqslant 1$ for all i. By Fact 1.1.4 (2), the number of solutions is

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

□

Exercise 1.1.6 Give combinatorial proofs of the following identities:

(1) $\displaystyle\sum_{k=0}^{n} \binom{n}{k}\binom{k}{m} = \binom{n}{m} 2^{n-m}.$

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

Proof (1) Both sides count the number of ways to select pairs (A,B) where $A\in\binom{[n]}{m}$ and $B\supset A$.

(2) We consider the following process: from m women and n men, we first choose a subset of the women to form a women's representative team, and then select a final team of exactly m people from the union of this women's team and all the men. Each valid outcome consists of a pair

$$(\text{women's team},\ \text{final team}),$$

and we count the total number of such pairs in two ways:

- ◇ By fixing the size k of the women's team, there are $\binom{m}{k}$ ways to choose them and then $\binom{n+k}{m}$ ways to form the final team.
- ◇ First choose k men from the n men and $m-k$ women from the m women to form the final team, in $\binom{n}{k}\binom{m}{k}$ ways. Then, the remaining k women (not in the final team) independently decide whether to join the women's representative team or not, giving 2^k possibilities. □

Proof of (2) using the binomial theorem Since

$$\binom{n+k}{m}=[x^m](1+x)^{n+k}=[x^m](1+x)^n(1+x)^k,$$

the left-hand side becomes

$$\begin{aligned}\sum_{k=0}^{m}\binom{m}{k}[x^m](1+x)^n(1+x)^k&=[x^m](1+x)^n\sum_{k=0}^{m}\binom{m}{k}(1+x)^k\\&=[x^m](1+x)^n(2+x)^m\\&=[x^m](1+x)^n\sum_{\ell=0}^{m}\binom{m}{\ell}2^{m-\ell}x^\ell\\&=\sum_{\ell=0}^{m}\binom{m}{\ell}2^{m-\ell}\binom{n}{m-\ell}\\&\overset{k:=m-\ell}{=\!=}\sum_{k=0}^{m}\binom{m}{m-k}2^k\binom{n}{k}\\&=\sum_{k=0}^{m}\binom{n}{k}\binom{m}{k}2^k.\end{aligned}$$

□

1.2 Counting Mappings

We denote the set of all mappings from Y to X by X^Y. The cardinality of this set is given by $|X|^{|Y|}$.

Fact 1.2.1 The number of injective maps from $[r]$ to $[n]$ is $(n)_r$.

Proof To construct an injective map $f\colon [r] \to [n]$, assign distinct values to $f(1), \ldots, f(r)$: there are n choices for $f(1)$, $n-1$ for $f(2)$, continuing down to $n-r+1$ for $f(r)$, giving a total of

$$n(n-1)(n-2)\cdots(n-r+1) = (n)_r. \qquad \square$$

Definition 1.2.2 The **Stirling numbers of the second kind**, written $S(r,n)$ or $\left\{ {r \atop n} \right\}$, count the number of ways to partition the set $[r]$ into n nonempty unlabeled subsets.

Exercise 1.2.3 Show that $\left\{ {r \atop 2} \right\} = 2^{r-1} - 1$.

Proof By Fact 1.1.2 (2), we have $\left\{ {r \atop 2} \right\} = \dfrac{1}{2}\displaystyle\sum_{k=1}^{r-1} \binom{r}{k} = \dfrac{2^r - 2}{2} = 2^{r-1} - 1.$ $\square$

Fact 1.2.4 The number of surjective maps from $[r]$ to $[n]$ is $n!S(r,n)$.

Proof The number of ways to partition the set $[r]$ into n nonempty labeled subsets is $n!S(r,n)$. $\square$

A map $f\colon X \to X$ is called a **permutation** if it is injective (and hence bijective). It can be viewed as the act of changing the linear order of an ordered set.

Cycle notation describes the effect of repeatedly applying the permutation on the elements of the set X, with an orbit being called a **cycle**. The permutation is written as a list of cycles; since distinct cycles involve disjoint sets of elements, this is referred to as "decomposition into disjoint cycles".

Definition 1.2.5 The **unsigned Stirling numbers of the first kind**, written $\left[{r \atop n} \right]$, count the number of permutations of $[r]$ with n disjoint cycles. The **(signed) Stirling numbers of the first kind** $s(r,n)$ are defined as $(-1)^{r-n}\left[{r \atop n} \right]$.

Exercise 1.2.6 Give a combinatorial proof of the following identities:

(1) $\left\{ {n \atop k} \right\} = \left\{ {n-1 \atop k-1} \right\} + k\left\{ {n-1 \atop k} \right\}$.

(2) $\left[{n \atop k} \right] = \left[{n-1 \atop k-1} \right] + (n-1)\left[{n-1 \atop k} \right]$.

Proof (1) The right-hand side can be interpreted as considering whether the element 1 in $[n]$ is in a subset by itself or in a subset with other elements.

(2) The right-hand side can be interpreted as considering whether the element 1 is a fixed point of a permutation or not. $\square$

Table 1.1: Triangular array of values for the Stirling numbers of the second kind

$\left\{ {n \atop k} \right\}$ n \ k	0	1	2	3	4	5	6	7	8	9	10
0	1										
1	0	1									
2	0	1	1								
3	0	1	3	1							
4	0	1	7	6	1						
5	0	1	15	25	10	1					
6	0	1	31	90	65	15	1				
7	0	1	63	301	350	140	21	1			
8	0	1	127	966	1701	1050	266	28	1		
9	0	1	255	3025	7770	6951	2646	462	36	1	
10	0	1	511	9330	34105	42525	22827	5880	750	45	1

Table 1.2: Triangular array of unsigned values for the Stirling numbers of the first kind

$\left[{n \atop k} \right]$ n \ k	0	1	2	3	4	5	6	7	8	9	10
0	1										
1	0	1									
2	0	1	1								
3	0	2	3	1							
4	0	6	11	6	1						
5	0	24	50	35	10	1					
6	0	120	274	225	85	15	1				
7	0	720	1764	1624	735	175	21	1			
8	0	5040	13068	13132	6769	1960	322	28	1		
9	0	40320	109584	118124	67284	22449	4536	546	36	1	
10	0	362880	1026576	1172700	723680	269325	63273	9450	870	45	1

1.3 Binomial Theorem

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

Fact 1.3.1 For $j \in [n]$, let $f_j(x) = \sum_{i \in I_j} x^i$, where I_j is a finite set of non-negative integers. If we define $f(x) = f_1(x) \cdots f_n(x)$, then

$$[x^k]f = \#\{(i_1, \ldots, i_n) : i_1 + \cdots + i_n = k,\ i_j \in I_j\}.$$

Fact 1.3.2 Let $f_1, \ldots, f_n$ be polynomials in x and define $f = f_1 \cdots f_n$. Then

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

Theorem 1.3.3 (Binomial theorem) For any real x and any positive integer n, we have

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

Proof Let $f(x) = (1+x)^n$. By Fact 1.3.1, we have

$$[x^k]f = \#\{(i_1, \ldots, i_n) : i_1 + \cdots + i_n = k,\ i_j \in \{0,1\}\} = \binom{n}{k}.$$

□

Fact 1.3.4 $\displaystyle\binom{2n}{n} = \sum_{k=0}^{n} \binom{n}{k}^2$.

Proof 1 By Fact 1.3.2, we have

$$\begin{aligned}
\binom{2n}{n} &= [x^n](1+x)^{2n} \\
&= \sum_{\substack{i+j=n \\ i,j \geqslant 0}} \left([x^i](1+x)^n\right)\left([x^j](1+x)^n\right) \\
&= \sum_{k=0}^{n} \binom{n}{k}\binom{n}{n-k} \\
&= \sum_{k=0}^{n} \binom{n}{k}^2.
\end{aligned}$$

□

Proof 2 Consider a set of $2n$ elements split into two disjoint subsets A and B, each of size n; the left-hand side $\binom{2n}{n}$ counts the number of ways to choose n elements from $A \cup B$. For each $k = 0, 1, \ldots, n$, the right-hand side counts the number of such subsets that include exactly k elements from A and $n-k$ from B, which is $\binom{n}{k}\binom{n}{n-k} = \binom{n}{k}^2$; summing over k gives the identity. □

This combinatorial double counting technique can be applied to prove the following identity.

Fact 1.3.5 (Vandermonde's convolution) $\displaystyle\binom{n+m}{k} = \sum_{j=0}^{k} \binom{n}{j}\binom{m}{k-j}$.

Exercise 1.3.6 Show that $\displaystyle\binom{n+m}{r+m} = \sum_{i-j=r} \binom{n}{i}\binom{m}{j}$.

Proof If we let $k = m - j$, then the right-hand side becomes

$$\sum_{i+k=m+r} \binom{n}{i}\binom{m}{k} = \binom{n+m}{r+m}$$

by Fact 1.3.5. □

Fact 1.3.7 (1) $\sum\limits_{k \text{ even}} \binom{n}{k} = \sum\limits_{k \text{ odd}} \binom{n}{k} = 2^{n-1}$.

(2) $\sum\limits_{k=0}^{n} k\binom{n}{k} = n2^{n-1}$.

Proof (1) Denote by E_n and O_n the sums of the even and odd binomial coefficients, respectively. Substituting $x = 1$ and $x = -1$ into Theorem 1.3.3 gives

$$(1+1)^n = 2^n = E_n + O_n$$

and

$$(1-1)^n = 0 = E_n - O_n.$$

Solving these two equations yields $E_n = O_n = 2^{n-1}$.

(2) If we let $f(x) = (1+x)^n$, then by Theorem 1.3.3,

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

Substituting $x = 1$ gives the desired result. □

Definition–Theorem 1.3.8 For any positive integer m and any non-negative integer n, the **multinomial theorem** describes how a sum with m terms expands when raised to the n-th power:

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{\substack{k_1+k_2+\cdots+k_m=n \\ k_1,k_2,\ldots,k_m \geqslant 0}} \binom{n}{k_1, k_2, \ldots, k_m} x_1^{k_1} x_2^{k_2} \cdots x_m^{k_m}$$

where

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

is a **multinomial coefficient**.

Proof This proof of the multinomial theorem uses Theorem 1.3.3 and induction on m.

First, for $m = 1$, both sides equal x_1^n. For the induction step, suppose the multinomial theorem holds for m. Then

$$\begin{aligned}
&(x_1 + x_2 + \cdots + x_m + x_{m+1})^n \\
&= [x_1 + x_2 + \cdots + (x_m + x_{m+1})]^n \\
&= \sum_{k_1+k_2+\cdots+k_{m-1}+K=n} \binom{n}{k_1, k_2, \ldots, k_{m-1}, K} x_1^{k_1} x_2^{k_2} \cdots x_{m-1}^{k_{m-1}} (x_m + x_{m+1})^K \\
&= \sum_{k_1+k_2+\cdots+k_{m-1}+K=n} \binom{n}{k_1, k_2, \ldots, k_{m-1}, K} x_1^{k_1} x_2^{k_2} \cdots x_{m-1}^{k_{m-1}} \sum_{k_m+k_{m+1}=K} \binom{K}{k_m, k_{m+1}} x_m^{k_m} x_{m+1}^{k_{m+1}} \\
&= \sum_{k_1+k_2+\cdots+k_{m-1}+k_m+k_{m+1}=n} \binom{n}{k_1, k_2, \ldots, k_{m-1}, k_m, k_{m+1}} x_1^{k_1} x_2^{k_2} \cdots x_{m-1}^{k_{m-1}} x_m^{k_m} x_{m+1}^{k_{m+1}},
\end{aligned}$$

which completes the induction. The last step follows because

$$\binom{n}{k_1,k_2,\ldots,k_{m-1},K}\binom{K}{k_m,k_{m+1}}=\binom{n}{k_1,k_2,\ldots,k_{m-1},k_m,k_{m+1}},$$

as can easily be seen by writing the three coefficients using factorials as follows:

$$\frac{n!}{k_1!k_2!\cdots k_{m-1}!K!}\cdot\frac{K!}{k_m!k_{m+1}!}=\frac{n!}{k_1!k_2!\cdots k_{m+1}!}.$$ □

Remark 1.3.9 The multinomial coefficient $\binom{n}{k_1,k_2,\ldots,k_m}$ can be interpreted in several ways:

(1) The number of ways of depositing n distinct objects into m distinct bins, with k_1 objects in the first bin, k_2 objects in the second bin, and so on.

(2) The number of m-dimensional lattice paths from $(0,0,\ldots,0)$ to $(k_1,k_2,\ldots,k_m)$.

(3) The number of distinct ways to permute a multiset of n elements, where k_i is the multiplicity of each of the i-th element. For example, the number of distinct permutations of the letters of the word MISSISSIPPI, which has 1 M, 4 Is, 4 Ss, and 2 Ps, is

$$\binom{11}{1,4,4,2}=\frac{11!}{1!4!4!2!}=34650.$$

(4) Given a number distribution $\{k_i\}$ on a set of n total items, k_i represents the number of items to be given the label i. The number of arrangements is found by

- ◇ Choosing k_1 of the total n to be labeled 1. This can be done $\binom{n}{k_1}$ ways.
- ◇ From the remaining $n-k_1$ items choose k_2 to label 2. This can be done $\binom{n-k_1}{k_2}$ ways.
- ◇ From the remaining $n-k_1-k_2$ items choose k_3 to label 3. Again, this can be done $\binom{n-k_1-k_2}{k_3}$ ways.

Multiplying the number of choices at each step results in:

$$\binom{n}{k_1}\binom{n-k_1}{k_2}\binom{n-k_1-k_2}{n_3}\cdots=\frac{n!}{k_1!(n-k_1)!}\cdot\frac{(n-k_1)!}{k_2!(n-k_1-k_2)!}\cdot\frac{(n-k_1-k_2)!}{k_3!(n-k_1-k_2-k_3)!}\cdots.$$

Cancellation results in $\binom{n}{k_1,k_2,\ldots,k_m}$.

The following exercise is a generalization of Pascal's rule (see Fact 1.1.2 (3)).

Exercise 1.3.10 Suppose $k_1+\cdots+k_m=n$ where $k_i\geqslant 1$ for all $i\in[m]$. Show that

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

Proof We can use the interpretations given in Remark 1.3.9 to prove this identity.

- ◇ If we use interpretation (1), then the right-hand side can be interpreted as considering which bin the first item goes into.
- ◇ If we use interpretation (2), each path ends with a step in some direction i. Removing that last step, the path reaches $(k_1,\ldots,k_i-1,\ldots,k_m)$, which has $\binom{n-1}{k_1,\ldots,k_i-1,\ldots,k_m}$ paths.

◇ If we use interpretation (3), the multinomial coefficient counts the number of words of length n using k_i copies of symbol i. Each such word begins with some symbol i. Removing that first letter, we get a word of length $n-1$ with $k_i - 1$ copies of symbol i, and the rest unchanged.

◇ If we use interpretation (4), the first item can be labeled with any of the m labels. After removing that first item, we have $n-1$ items left, and the number of arrangements is given by the right-hand side. □

1.4 Estimating Binomial Coefficients

Theorem 1.4.1 For any integer $n \geqslant 1$, we have

$$e\Big(\frac{n}{e}\Big)^n \leqslant n! \leqslant en\Big(\frac{n}{e}\Big)^n.$$

Proof 1 On the one hand, we have

$$\ln(n!) = \sum_{k=1}^{n} \ln k \leqslant \int_1^{n+1} \ln x \,\mathrm{d}x = (n+1)\ln(n+1) - n,$$

which implies

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

and thus

$$n! = n(n-1)! \leqslant n\frac{n^n}{e^{n-1}} = en\Big(\frac{n}{e}\Big)^n.$$

On the other hand, we have

$$\ln(n!) = \sum_{k=1}^{n} \ln k \geqslant \int_1^{n} \ln x \,\mathrm{d}x = n\ln n - (n-1),$$

which implies

$$n! \geqslant \frac{n^n}{e^{n-1}} = e\Big(\frac{n}{e}\Big)^n.$$

□

Proof 2 Theorem 1.4.1 also follows by induction on n. The base case $n = 1$ is trivial. For the induction step, it suffices to show that

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

or equivalently,

$$n\ln\left(1+\frac{1}{n}\right) \leqslant 1 \leqslant (n+1)\ln\left(1+\frac{1}{n}\right).$$

This follows from $\frac{2x}{2+x} \leqslant \ln(1+x) \leqslant x$ for all $x \geqslant 0$ (see Remark 1.4.3 for the left inequality). □

Exercise 1.4.2 Show that the right-hand side of Theorem 1.4.1 can be improved to

$$n! \leqslant e\sqrt{n}\Big(\frac{n}{e}\Big)^n.$$

Proof We are going to show that

$$\ln(n!) \leqslant n\ln n - n + \frac{1}{2}\ln n + 1.$$

For this, we let

$$f(n) = \ln(n!) - n\ln n + n - \frac{1}{2}\ln n - 1.$$

Then

$$\begin{aligned} f(n+1) - f(n) &= \ln(n+1) - (n+1)\ln(n+1) + n\ln n + 1 - \frac{1}{2}\ln(n+1) + \frac{1}{2}\ln n \\ &= -\left(n+\frac{1}{2}\right)\ln\left(1+\frac{1}{n}\right) + 1 \\ &\leqslant -\left(n+\frac{1}{2}\right)\frac{\frac{2}{n}}{2+\frac{1}{n}} + 1 \\ &= 0, \end{aligned}$$

where the last inequality follows from the fact that $\ln(1+x) \geqslant \frac{2x}{2+x}$ for all $x \geqslant 0$ (see Remark 1.4.3). Hence, $f(n) \leqslant f(1) = 0$, which implies the desired inequality. □

Remark 1.4.3 To prove that $\ln(1+x) \geqslant \frac{2x}{2+x}$ for all $x \geqslant 0$, we consider the function $g(x) = \ln(1+x) - \frac{2x}{2+x}$. Since

$$g'(x) = \frac{1}{1+x} - \frac{4}{(2+x)^2} = \frac{x^2}{(1+x)(2+x)^2} \geqslant 0$$

for all $x \geqslant 0$, we have $g(x) \geqslant g(0) = 0$.

Theorem 1.4.4 (Stirling's approximation) For large n, the factorial satisfies $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$. Here the sign $\sim$ means that the two quantities are asymptotic, that is, their ratio tends to 1 as n tends to infinity.

Fact 1.4.5 The binomial coefficients increase up to the middle term and then decrease:

$$\binom{n}{0} \leqslant \binom{n}{1} \leqslant \cdots \leqslant \binom{n}{\lfloor n/2\rfloor} = \binom{n}{\lceil n/2\rceil} \geqslant \cdots \geqslant \binom{n}{n-1} \geqslant \binom{n}{n}.$$

As a consequence of this unimodality and Fact 1.1.2 (2), we have

$$\frac{2^n}{n+1} \leqslant \binom{n}{\lfloor n/2\rfloor} \leqslant 2^n.$$

Exercise 1.4.6 (Central binomial coefficient) For any positive even integer n, show that

$$\frac{2^n}{\sqrt{2n}} \leqslant \binom{n}{n/2} \leqslant \frac{2^n}{\sqrt{n}}.$$

Proof It suffices to show for any positive integer n that

$$\frac{4^n}{2\sqrt{n}} \leqslant \binom{2n}{n} \leqslant \frac{4^n}{\sqrt{2n}}.$$

(1) To prove the left-hand side, we note that

$$\frac{1}{4^n}\binom{2n}{n} = \frac{(2n)!}{4^n(n!)^2} = \frac{(2n)!!(2n-1)!!}{(2^n n!)^2} = \frac{(2n-1)!!}{(2n)!!} = \frac{1}{2n}\prod_{k=1}^{n-1}\left(1+\frac{1}{2k}\right),$$

from which we can deduce that

$$\left\{\frac{1}{4^n}\binom{2n}{n}\right\}^2 = \frac{1}{4n^2}\prod_{k=1}^{n-1}\left(1+\frac{1}{2k}\right)^2 \geqslant \frac{1}{4n^2}\prod_{k=1}^{n-1}\left(1+\frac{1}{k}\right) = \frac{1}{4n},$$

and thus

$$\frac{1}{4^n}\binom{2n}{n} \geqslant \frac{1}{2\sqrt{n}}.$$

(2) For the right-hand side, we can prove a stronger inequality

$$\binom{2n}{n} < \frac{4^n}{\sqrt{2n+1}}$$

via induction on n. The base case $n = 1$ is trivial. Since

$$\binom{2n+2}{n+1} = \frac{(2n+2)(2n+1)}{(n+1)^2}\binom{2n}{n},$$

if we assume the inequality holds for n, then it suffices to show that

$$\frac{(2n+2)(2n+1)}{(n+1)^2} < \frac{4\sqrt{2n}}{\sqrt{2n+2}},$$

which holds trivially after simplifying both sides. □

Remark 1.4.7 Using Theorem 1.4.4, one has $\binom{n}{n/2} \sim \frac{2^n}{\sqrt{\frac{\pi}{2}n}}$ for large even n.

Fact 1.4.8 $\displaystyle\binom{n}{k} \leqslant \frac{n^k}{k!}$.

Proof We have $\binom{n}{k} = \frac{(n)_k}{k!} \leqslant \frac{n^k}{k!}$. □

Fact 1.4.9 For any $1 \leqslant k \leqslant n$, we have

$$\left(\frac{n}{k}\right)^k \leqslant \binom{n}{k} \leqslant \left(\frac{en}{k}\right)^k. \tag{1–1}$$

Proof Since $\frac{n-j}{k-j} \geqslant \frac{n}{k}$ for $0 \leqslant j \leqslant k-1$, we have

$$\binom{n}{k} = \frac{n}{k}\cdot\frac{n-1}{k-1}\cdots\frac{n-(k-1)}{k-(k-1)} \geqslant \left(\frac{n}{k}\right)^k.$$

By Theorem 1.4.1, $k! \geqslant e\left(\frac{k}{e}\right)^k > \left(\frac{k}{e}\right)^k$. Hence, by Fact 1.4.8,

$$\binom{n}{k} \leqslant \frac{n^k}{k!} \leqslant \left(\frac{en}{k}\right)^k.$$

□

In fact, the right-hand side of (1–1) can be improved to the following.

Theorem 1.4.10 For any $1 \leqslant k \leqslant n$, we have

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k} \leqslant \left(\frac{en}{k}\right)^k.$$

Proof By Theorem 1.3.3, for $0 < x \leqslant 1$,

$$\binom{n}{0} + \binom{n}{1}x + \cdots + \binom{n}{k}x^k \leqslant (1+x)^n,$$

and thus

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k} \leqslant \frac{\binom{n}{0}}{x^k} + \frac{\binom{n}{1}}{x^{k-1}} + \cdots + \frac{\binom{n}{k}}{1} \leqslant \frac{(1+x)^n}{x^k}.$$

Setting $x = \frac{k}{n} \in (0, 1]$ gives

$$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{k} \leqslant \frac{\left(1+\frac{k}{n}\right)^n}{\left(\frac{k}{n}\right)^k} \leqslant \frac{\left(e^{\frac{k}{n}}\right)^n}{\left(\frac{k}{n}\right)^k} = \left(\frac{en}{k}\right)^k.$$

□

1.5 Inclusion–Exclusion Principle

Let Ω be a ground set and let $A_1, A_2, \ldots, A_n$ be subsets of Ω. Denote by A_i^c the complement of A_i in Ω. We adopt the following notations:

- $A_\varnothing \coloneqq \Omega$.
- $A_I \coloneqq \bigcap_{i \in I} A_i$ for any nonempty subset $I \subset [n]$.
- $S_k \coloneqq \sum_{I \in \binom{[n]}{k}} |A_I|$ any $0 \leqslant k \leqslant n$.

Theorem 1.5.1 (Inclusion–exclusion principle, union form)

$$|A_1 \cup \cdots \cup A_n| = \sum_{\varnothing \neq I \subset [n]} (-1)^{|I|+1} |A_I| = \sum_{k=1}^{n} (-1)^{k+1} S_k.$$

Proof 1 Let $A = A_1 \cup \cdots \cup A_n$. Observe that

$$(\mathbb{1}_A - \mathbb{1}_{A_1})(\mathbb{1}_A - \mathbb{1}_{A_2}) \cdots (\mathbb{1}_A - \mathbb{1}_{A_n})(x) \equiv 0, \quad \forall x \in \Omega.$$

Expanding this product gives

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

Summing over all $x \in \Omega$ gives

$$|A| + \sum_{\varnothing \neq I \subset [n]} (-1)^{|I|} |A_I| = 0.$$

□

Proof 2 It suffices to show that

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

For a fixed $x\in\Omega$, let ℓ be the number of sets among $A_1,\ldots,A_n$ that contain x. If $\ell=0$, then both sides are 0. If $\ell\geqslant 1$, then the left-hand side is 1 and the right-hand side is

$$\sum_{k=1}^{\ell}(-1)^{k+1}\binom{\ell}{k} = 1-(1-1)^{\ell} = 1.$$

□

For some applications the following form of the inclusion–exclusion principle is more convenient.

Theorem 1.5.2 (Inclusion–Exclusion principle, complement form)

$$\left|\Omega\setminus\bigcup_{i=1}^{n}A_i\right| = |A_1^c\cap\cdots\cap A_n^c| = \sum_{I\subset[n]}(-1)^{|I|}|A_I| = \sum_{k=0}^{n}(-1)^k S_k. \tag{1–2}$$

Proof The left-hand side of (1–2) is $|A_\varnothing|$ minus $|A_1\cup\cdots\cup A_n|$. By Theorem 1.5.1 this number is

$$|A_\varnothing| - \sum_{\varnothing\neq I\subset[n]}(-1)^{|I|+1}|A_I| = \sum_{I\subset[n]}(-1)^{|I|}|A_I|.$$

□

Let $\varphi(n)$ denote **Euler's totient function**, which counts the positive integers up to a given integer n that are relatively prime to n.

Theorem 1.5.3 (Euler's product formula) For any positive integer n, we have

$$\varphi(n) = n\prod_{p\mid n}\left(1-\frac{1}{p}\right),$$

where the product is taken over the distinct prime divisors of n.

Proof Let $n = p_1^{a_1}\cdots p_t^{a_t}$ be the prime factorization of n, where $p_1,\ldots,p_t$ are distinct primes and $a_1,\ldots,a_t$ are positive integers. Define $\Omega=[n]$ and $A_i=\{m\in[n]: p_i\mid m\}$ for $i\in[t]$. By Theorem 1.5.2, we have

$$\varphi(n) = \left|\Omega\setminus\bigcup_{i=1}^{t}A_i\right| = \sum_{I\subset[t]}(-1)^{|I|}|A_I|.$$

Since A_I is the set of integers in $[n]$ that are divisible by all primes in I, we have

$$|A_I| = \frac{n}{\prod_{i\in I}p_i},$$

and thus

$$\varphi(n) = \sum_{I\subset[t]}(-1)^{|I|}\frac{n}{\prod_{i\in I}p_i} = n\prod_{i=1}^{t}\left(1-\frac{1}{p_i}\right).$$

□

Definition 1.5.4 A **derangement** is a permutation that has no fixed points.

Theorem 1.5.5 The number of derangements of $[n]$ is equal to $n!\sum_{k=0}^{n}\frac{(-1)^k}{k!}$.

Proof Let Ω be the set of all permutations of $[n]$, and A_i the set of permutations fixing the point i. Then $|A_I| = (n-|I|)!$ for any $I \subset [n]$, and thus by Theorem 1.5.2, the number of derangements is

$$\left|\Omega \setminus \bigcup_{i=1}^{n} A_i\right| = \sum_{I\subset[n]} (-1)^{|I|}(n-|I|)! = \sum_{k=0}^{n}(-1)^k\binom{n}{k}(n-k)! = n!\sum_{k=0}^{n}\frac{(-1)^k}{k!}.$$

□

Remark 1.5.6 The sum $\sum_{k=0}^{n}\frac{(-1)^k}{k!}$ is the initial part of the Taylor expansion of e^{-1}; so about an e^{-1} fraction of all permutations are derangements.

Theorem 1.5.7 Stirling numbers of the second kind can be calculated using a one-sum formula:

$$\left\{{n \atop k}\right\} = \frac{1}{k!}\sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i}i^n.$$

Proof Let Ω be the set of all functions from $[n]$ to $[k]$, and A_i the set of functions whose image does not contain i. Then $|A_I| = (k-|I|)^n$ for any $I \subset [k]$, and thus by Theorem 1.5.2, the number of surjective functions from $[n]$ to $[k]$ is

$$\left|\Omega \setminus \bigcup_{i=1}^{k} A_i\right| = \sum_{I\subset[k]} (-1)^{|I|}(k-|I|)^n = \sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^n.$$

By Fact 1.2.4, this number is also equal to $k!\left\{{n \atop k}\right\}$. Combined, we have

$$\left\{{n \atop k}\right\} = \frac{1}{k!}\sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^n = \frac{1}{k!}\sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i}i^n,$$

□

1.6 Ordinary Generating Functions

Definition 1.6.1 The **(ordinary) generating function (OGF)** of a sequence $(a_n)_{n\geqslant 0}$ is $\sum_{n=0}^{\infty} a_n x^n$.

Remark 1.6.2 Unlike an ordinary series, the formal power series is not required to converge: in fact, the generating function is not actually regarded as a function, and the "variable" remains an indeterminate. These expressions in terms of the indeterminate x may involve arithmetic operations, differentiation with respect to x and composition with (i.e., substitution into) other generating functions; since these operations are also defined for functions, the result looks like a function of x. Indeed, the closed form expression can often be interpreted as a function that can be evaluated at (sufficiently small) concrete values of x, and which has the formal series as its series expansion; this explains the designation "generating functions". However such interpretation is not required to be possible, because formal series are not required to give a convergent series when a nonzero numeric value is substituted for x.

Example 1.6.3 The generating function of the constant sequence $1, 1, 1, \ldots$ is the geometric series

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

Since its derivative is given by

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

we see that the generating function of the sequence $1, 2, 3, \ldots$ is $\frac{1}{(1-x)^2}$.

Example 1.6.4 Suppose $a_0 = 1$ and $a_n = 2a_{n-1}$ for $n \geqslant 1$. Then the generating function $f(x)$ satisfies

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

This implies that

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

and thus $a_n = 2^n$.

Exercise 1.6.5 Let A_n denote the set of strings of length n formed from the alphabet $\{a, b, c\}$, in which the substring "aa" does not appear. Determine $|A_n|$ for $n \geqslant 1$.

Solution Let $a_n = |A_n|$. We first observe that $a_1 = 3$, and $a_2 = 3^2 - 1 = 8$. For $n \geqslant 3$, we determine a_n recursively as follows:

- ◇ If the first character is 'a', the second character must be either 'b' or 'c', and the remaining $n-2$ characters can be any string in A_{n-2}. This gives $2a_{n-2}$ strings.
- ◇ If the first character is 'b' or 'c', the remaining $n-1$ characters can be any string in A_{n-1}. This gives $2a_{n-1}$ strings.

Thus, we have the recurrence relation

$$a_n = 2a_{n-1} + 2a_{n-2} \tag{1–3}$$

for $n \geqslant 3$. Set $a_0 = 1$. Then (1–3) holds for all $n \geqslant 2$. Let $f(x)$ be the generating function of the sequence (a_n). Then

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + \sum_{n=2}^{\infty} (2a_{n-1} + 2a_{n-2})x^n = 1 + 3x + 2x[f(x) - 1] + 2x^2 f(x).$$

Solving for $f(x)$ gives

$$\begin{aligned} f(x) &= \frac{1+x}{1-2x-2x^2} \\ &= \frac{1-\sqrt{3}}{2\sqrt{3}} \frac{1}{\sqrt{3}+1+2x} + \frac{1+\sqrt{3}}{2\sqrt{3}} \frac{1}{\sqrt{3}-1-2x}. \end{aligned}$$

This implies that

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

Exercise 1.6.6 The **Fibonacci sequence** is defined by the recurrence relation $F_0 = 1$, $F_1 = 1$, and $F_n = F_{n-1} + F_{n-2}$ for $n \geqslant 2$. Find a formula for F_n.

Solution Let $f(x)$ be the generating function of the Fibonacci sequence. Then

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} F_n x^n = 1 + x + \sum_{n=2}^{\infty} F_n x^n, \\ -xf(x) &= -\sum_{n=0}^{\infty} F_n x^{n+1} = -\sum_{n=1}^{\infty} F_{n-1} x^n = -x - \sum_{n=2}^{\infty} F_{n-1} x^n, \\ -x^2 f(x) &= -\sum_{n=0}^{\infty} F_n x^{n+2} = -\sum_{n=2}^{\infty} F_{n-2} x^n. \end{aligned}$$

Adding these three equations gives

$$\left(1 - x - x^2\right) f(x) = 1 + \sum_{n=2}^{\infty} (F_n - F_{n-1} - F_{n-2}) x^n = 1.$$

Hence,

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

Thus, we have

$$\begin{aligned} F_n &= \frac{2}{\sqrt{5}}\frac{1}{\sqrt{5}-1}\left(\frac{2}{\sqrt{5}-1}\right)^n + \frac{2}{\sqrt{5}}\frac{1}{\sqrt{5}+1}\left(\frac{-2}{\sqrt{5}+1}\right)^n \\ &= \frac{1}{\sqrt{5}}\left\{\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right\}. \end{aligned} \qquad \square$$

Around 1665, Isaac Newton generalized the binomial theorem to allow real exponents other than non-negative integers. (The same generalization also applies to complex exponents.) In this generalization, the finite sum is replaced by an infinite series. In order to do this, one needs to give meaning to binomial coefficients with an arbitrary upper index, which cannot be done using the usual formula with factorials. However, for an arbitrary number r, one can define

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

Theorem 1.6.7 (Newton's generalized binomial theorem) If x and y are real numbers with $|x| > |y|$, and r is any complex number, one has

$$\begin{aligned} (x+y)^r &= \sum_{k=0}^{\infty} \binom{r}{k} x^{r-k} y^k \\ &= x^r + r x^{r-1} y + \frac{r(r-1)}{2!} x^{r-2} y^2 + \frac{r(r-1)(r-2)}{3!} x^{r-3} y^3 + \cdots. \end{aligned}$$

Proof Applying the Taylor expansion of the function $f(t) = t^r$ at $t = 1$, we obtain

$$(x+y)^r = x^r\Big(1+\frac{y}{x}\Big)^r = x^r\sum_{k=0}^{\infty}\frac{(r)_k}{k!}\Big(\frac{y}{x}\Big)^k = \sum_{k=0}^{\infty}\binom{r}{k}x^{r-k}y^k.$$

□

Corollary 1.6.8 When $r = -s$, we have by Theorem 1.6.7, for $|x| < 1$:

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

Replacing x with $-x$ yields

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

In particular, if $s = n$ is a positive integer, we have

$$\frac{1}{(1-x)^n} = \sum_{k=0}^{\infty}\binom{n+k-1}{k}x^k. \tag{1–4}$$

According to Fact 1.1.4 (3), the binomial coefficient $\binom{n+k-1}{k}$ counts the number of non-negative integer solutions to the equation $x_1 + \cdots + x_n = k$. This agrees with Fact 1.3.1 when we take $f_j(x) = \frac{1}{1-x}$ for each $j \in [n]$.

Remark 1.6.9 We can also prove (1–4) by taking the $(n-1)$st derivative of $\frac{1}{1-x}$.

Example 1.6.10 For any non-negative integer k, we have

$$\begin{aligned}\binom{\frac{1}{2}}{k} &= \frac{\frac{1}{2}\big(-\frac{1}{2}\big)\big(-\frac{3}{2}\big)\cdots\big(\frac{1}{2}-k+1\big)}{k!}\\ &= \frac{(-1)^{k-1}(2k-3)!!}{2^k k!}\cdot\frac{(2k-2)!!}{(2k-2)!!}\\ &= \frac{(-1)^{k-1}(2k-2)!}{2^k k!2^{k-1}(k-1)!}\\ &= \frac{(-1)^{k-1}\cdot 2}{4^k}\cdot\frac{(2k-2)!}{k!(k-1)!}.\end{aligned}$$

Exercise 1.6.11 Let a_n be the number of ways to pay n dollars using only one-dollar, two-dollar, and five-dollar bills. Find the generating function of the sequence (a_n).

Solution Since

$$a_n = \#\{(i_1,i_2,i_3) : i_1+i_2+i_3 = n,\ i_j \in I_j\},$$

where $I_1 = \{0,1,2,\ldots\}$, $I_2 = \{0,2,4,\ldots\}$, and $I_3 = \{0,5,10,\ldots\}$, we have by Fact 1.3.1:

$$\sum_{n=0}^{\infty}a_n x^n = \prod_{j=1}^{3}\sum_{k\in I_j}x^k = \frac{1}{(1-x)(1-x^2)(1-x^5)}.$$

□

Exercise 1.6.12 Let S_n denote the number of length-n *selections* over the alphabet $\{a,b,c\}$ with unlimited repetitions, such that the numbers of occurrences of both a and b are even. Determine S_n.

Solution Since

$$S_n = \#\big\{(i_1, i_2, i_3) \in \mathbb{N}^3 : i_1 + i_2 + i_3 = n,\ 2 \mid i_1,\ 2 \mid i_2\big\},$$

by Fact 1.3.1, the generating function of (S_n) is given by

$$f(x) = \sum_{\substack{i_1 \geqslant 0 \\ 2 \mid i_1}} x^{i_1} \sum_{\substack{i_2 \geqslant 0 \\ 2 \mid i_2}} x^{i_2} \sum_{i_3 \geqslant 0} x^{i_3} = \frac{1}{(1-x^2)^2(1-x)} = \frac{1}{(1-x)^3(1+x)^2}.$$

Therefore,

$$S_n = [x^n] f(x) = \frac{\left\lceil \frac{n+1}{2} \right\rceil \left(\left\lceil \frac{n+1}{2} \right\rceil + 1 \right)}{2}.$$

□

1.7 Integer Partitions

Definition 1.7.1 Let n be a non-negative integer.

(1) A **composition** of n is a way of writing n as the sum of a *sequence* of positive integers.

(2) A **partition** of n is a way of writing n as a sum of positive integers.

Remark 1.7.2 (1) Two sequences that differ in the order of their terms define different compositions of their sum, while they are considered to define the same integer partition of that number.

(2) The integer 0 has one composition (the empty sequence) and one partition (the empty sum).

(3) By Fact 1.1.4 (2), each positive integer n has $\displaystyle\sum_{k=1}^{n} \binom{n-1}{k-1} = 2^{n-1}$ distinct compositions.

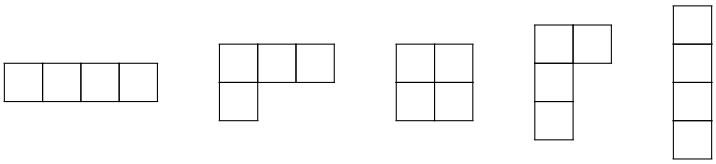


Figure 1.1: Young diagrams associated to the partitions of 4

The **partition function** $p(n)$ counts the partitions of a non-negative integer n. For instance, we have

$$p(0) = 1, \quad p(1) = 1, \quad p(2) = 2, \quad p(3) = 3, \quad p(4) = 5, \quad p(5) = 7.$$

Theorem 1.7.3 The generating function of p is

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

Proof Let $I_j = \{0, j, 2j, 3j, \ldots\}$ for each $j \geqslant 1$. Since

$$p(n) = \#\{(i_1, \ldots, i_n) : i_1 + \cdots + i_n = n,\ i_j \in I_j\},$$

by Fact 1.3.1 we have

$$P(x) = \prod_{j=1}^{\infty} \sum_{\ell \in I_j} x^{\ell} = \prod_{j=1}^{\infty} \frac{1}{1-x^j}.$$

□

Remark 1.7.4 An asymptotic expression for $p(n)$ is given by

$$p(n) \sim \frac{1}{4\sqrt{3}n} e^{\pi\sqrt{\frac{2n}{3}}} \quad \text{as } n \to \infty.$$

This asymptotic formula was first obtained by G. H. Hardy and Ramanujan in 1918 and independently by J. V. Uspensky in 1920.

Exercise 1.7.5 Show that $p(n) \geqslant \frac{1}{e^5 n^2} e^{2\sqrt{n}}$.

Proof Since each partition of n with k parts corresponds to at most $k!$ distinct compositions of n with k parts, we have by Remark 1.7.2 (3):

$$p(n) \geqslant \sum_{k=1}^{n} \frac{1}{k!} \binom{n-1}{k-1} \geqslant \frac{1}{\lfloor\sqrt{n}\rfloor!} \binom{n-1}{\lfloor\sqrt{n}\rfloor - 1}.$$

It is enough to prove that

$$\frac{1}{\lfloor\sqrt{n}\rfloor!} \binom{n-1}{\lfloor\sqrt{n}\rfloor - 1} \geqslant \frac{1}{e^5 n^2} e^{2\sqrt{n}}.$$

Setting $m = \lfloor\sqrt{n}\rfloor$, this reduces to showing that

$$\frac{1}{m!} \binom{m^2-1}{m-1} \geqslant \frac{e^{2m-5}}{m^4}. \tag{1–5}$$

To prove (1–5), observe first that

$$\frac{1}{m!} \binom{m^2-1}{m-1} = \frac{(m^2-1)(m^2-2)\cdots(m^2-m+1)}{m!(m-1)!} \geqslant \frac{(m^2-m+1)^{m-1}}{m!(m-1)!}.$$

Applying the upper bound from Exercise 1.4.2 to each factorial yields

$$m!(m-1)! \leqslant \frac{m^{m+\frac{1}{2}}(m-1)^{m-\frac{1}{2}}}{e^{2m-3}},$$

and therefore

$$\begin{aligned}
\frac{1}{m!} \binom{m^2-1}{m-1} &\geqslant \frac{e^{2m-3}(m^2-m+1)^{m-1}}{m^{m+\frac{1}{2}}(m-1)^{m-\frac{1}{2}}} \\
&\geqslant \frac{e^{2m-3}(m^2-m)^{m-1}}{m^{m+1}(m-1)^{m-1}} \\
&= \frac{e^{2m-3}}{m^2} \\
&\geqslant \frac{e^{2m-5}}{m^4}.
\end{aligned}$$

□

1.8 Catalan Numbers

Definition 1.8.1 A **polygon triangulation** is a subdivision of a given polygon into triangles meeting edge-to-edge, with the property that the set of triangle vertices coincides with the set of vertices of the polygon.

Definition–Theorem 1.8.2 The total number of polygon triangulations of a convex $(n+2)$-gon is the n-th **Catalan number**, given by

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

Proof For each $n \geqslant 3$, let b_{n-1} denote the number of ways to triangulate a convex n-gon whose vertices are labeled $1, 2, \ldots, n$ counterclockwise. Clearly, $b_2 = 1$ and $b_3 = 2$. We aim to determine the recurrence relation satisfied by this sequence.

Consider a triangulation of this n-gon, and let T be the triangle that contains vertices 1 and 2. Let k be the label of the third vertex of T.

- ◇ If $k = 3$ or $k = n$, then T divides the polygon into the triangle T itself and an $(n-1)$-gon, which can be triangulated in b_{n-2} ways.

- ◇ If $4 \leqslant k \leqslant n-1$, then T divides the polygon into the triangle T and two smaller polygons: one with vertices $2, 3, \ldots, k$ and the other with vertices $k, \ldots, n-1, n, 1$. The first can be triangulated in b_{k-2} ways and the second in b_{n-k+1} ways, giving a total of $b_{k-2}b_{n-k+1}$ triangulations of the original polygon.

Therefore, the recurrence relation is

$$b_{n-1} = 2b_{n-2} + \sum_{k=4}^{n-1} b_{k-2}b_{n-k+1} = 2b_{n-2} + \sum_{k=2}^{n-3} b_k b_{n-k-1}.$$

Setting $b_0 = 0$ and $b_1 = 1$, we can rewrite this as

$$b_{n-1} = \sum_{k=0}^{n-1} b_k b_{n-1-k}, \quad \forall n \geqslant 3,$$

or equivalently,

$$b_n = \sum_{k=0}^{n} b_k b_{n-k}, \quad \forall n \geqslant 2.$$

If $f(x)$ is the generating function of (b_n), then

$$f^2(x) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} b_k b_{n-k}\right)x^n,$$

which gives

$$f(x) = x + \sum_{n=2}^{\infty} b_n x^n = x + \sum_{n=2}^{\infty}\left(\sum_{k=0}^{n} b_k b_{n-k}\right)x^n$$

$$
\begin{aligned}
&= x + \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} b_k b_{n-k} \right) x^n \\
&= x + f^2(x).
\end{aligned}
$$

Solving this quadratic equation yields

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

Since $f(0) = 0$, we must take the minus sign:

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

Using Theorem 1.6.7 and Example 1.6.10, we have

$$
\begin{aligned}
f(x) &= \frac{1}{2} - \frac{1}{2} \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (-4x)^n \\
&= \sum_{n=1}^{\infty} \frac{(-1)^{n+1} 4^n}{2} \binom{\frac{1}{2}}{n} x^n \\
&= \sum_{n=1}^{\infty} \frac{(-1)^{n+1} 4^n}{2} \cdot \frac{(-1)^{n-1} \cdot 2}{4^n} \cdot \frac{(2n-2)!}{n!(n-1)!} x^n \\
&= \sum_{n=1}^{\infty} \frac{(2n-2)!}{n!(n-1)!} x^n \\
&= \sum_{n=1}^{\infty} \frac{1}{n} \binom{2n-2}{n-1} x^n.
\end{aligned}
$$

Hence,

$$b_{n+1} = \frac{1}{n+1} \binom{2n}{n}.$$

□

1.9 Random Walk

Consider the real axis marked with integer points $0, \pm 1, \pm 2, \pm 3, \ldots$. A frog moves along these points according to the following rules:

- ◇ The frog starts at position 1.
- ◇ At each step, it either
  - jumps two units to the right, or
  - jumps one unit to the left,

  each with probability $\frac{1}{2}$, independently of previous steps.

Question What is the probability that the frog ever reaches position 0?

Solution Let A be the event that the frog reaches position 0 at least once. For $i \geqslant 1$, let A_i be the event that the frog reaches 0 for the first time at step i. Clearly,

$$P(A) = \sum_{i=1}^{\infty} P(A_i).$$

To compute $P(A_i)$, let a_i be the number of length-i trajectories starting at 1 and reaching 0 for the first time at step i. Then

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

Let $f(x)$ be the generating function of $(a_i)_{i\geqslant 0}$, with $a_0 = 0$. Then

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

To determine $f(x)$, we also consider related problems starting from different positions:

- ◇ Let b_i be the number of length-i trajectories starting at 2 and reaching 0 for the first time at step i.
- ◇ Let c_i be the number of length-i trajectories starting at 3 and reaching 0 for the first time at step i.
- ◇ Let $g(x)$ and $h(x)$ be the generating functions of $(b_i)_{i\geqslant 0}$ and $(c_i)_{i\geqslant 0}$, respectively.

A length-i trajectory from 1 to 0 begins with one of two possible moves:

- ◇ Jump left: The frog moves from 1 to 0 in a single step, so $a_1 = 1$.
- ◇ Jump right: The frog moves from 1 to 3. From there, it must eventually reach 0. The number of such trajectories in $i-1$ steps is c_{i-1}.

Hence,

$$a_i = \begin{cases} 1, & i = 1, \\ c_{i-1}, & i \geqslant 2. \end{cases}$$

This yields

$$f(x) = a_1 x + \sum_{i=2}^{\infty} a_i x^i = x + \sum_{i=2}^{\infty} c_{i-1} x^i = x + xh(x). \tag{1–6}$$

From 2 to 0, the frog must first reach 1 (in j steps, counted by a_j), then from 1 reach 0 (in $i-j$ steps, counted by a_{i-j}). Thus,

$$b_i = \sum_{j=1}^{i-1} a_j a_{i-j} = \sum_{j=0}^{i} a_j a_{i-j},$$

which gives

$$g(x) = f(x)^2.$$

Similarly, a trajectory from 3 to 0 must first reach 2, then from 2 reach 0. The generating function from 3 to 2 is $f(x)$ (by symmetry of the step sizes), and from 2 to 0 is $g(x) = f(x)^2$. Therefore,

$$h(x) = f(x)g(x) = f(x)^3.$$

Substituting into (1–6):

$$f(x) = x + xf^3(x). \tag{1–7}$$

Let $\alpha = P(A) = f(1/2) \in [0,1]$. Then from (1–7):

$$\alpha = \frac{1}{2} + \frac{\alpha^3}{2} \implies (\alpha - 1)\left(\alpha^2 + \alpha - 1\right) = 0.$$

Thus, the possible values for $P(A)$ are 1 and $\frac{\sqrt{5}-1}{2}$.

Finally, note that (1–7) defines $f(x)$, whose inverse is $g(x) = \frac{x}{1+x^3}$. Since $f(x)$ is the generating function of a sequence of non-negative integers, it is increasing within its radius of convergence, and so must be its inverse $g(x)$. Examining $g(x)$, we observe that it is increasing near $\frac{\sqrt{5}-1}{2}$ but decreasing near 1 (see Figure 1.2). Therefore, the only valid solution is

$$P(A) = \frac{\sqrt{5}-1}{2}.$$

□

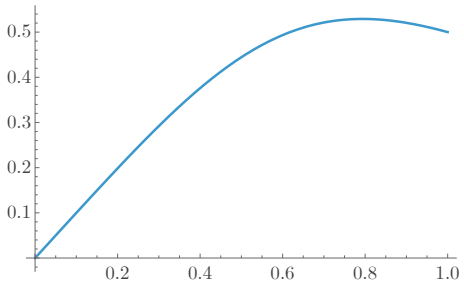


Figure 1.2: Graph of $g(x)$

1.10 Exponential Generating Functions

Definition 1.10.1 The **exponential generating function (EGF)** of a sequence $(a_n)_{n\geqslant 0}$ is $\displaystyle\sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$.

As corollaries and analogues of Facts 1.3.1 and 1.3.2, we have the following results.

Fact 1.10.2 For $j \in [n]$, let $g_j(x) = \displaystyle\sum_{i\in I_j} \frac{x^i}{i!}$, where I_j is a finite set of non-negative integers. If we let

$$b_k = \sum_{\substack{i_1+\cdots+i_n=k \\ i_j\in I_j}} \frac{k!}{i_1!\cdots i_n!},$$

then

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

Fact 1.10.3 For $j \in [n]$, let $g_j(x) = \displaystyle\sum_{i=0}^{\infty} \frac{a_i^{(j)}}{i!} x^i$ and define $g(x) = g_1(x)\cdots g_n(x)$. Then

$$g(x) = \sum_{k=0}^{\infty} \frac{A_k}{k!} x^k \iff A_k = \sum_{\substack{i_1+\cdots+i_n=k \\ i_j\geqslant 0}} \left(\frac{k!}{i_1!\cdots i_n!} \prod_{j=1}^{n} a_{i_j}^{(j)} \right).$$

Exponential generating functions are generally more convenient than ordinary generating functions for combinatorial enumeration problems that involve *labeled* objects. For example, Exercise 1.10.4 contrasts with Exercise 1.6.12 in this regard.

Exercise 1.10.4 Let T_n denote the number of length-n *words* over the alphabet $\{a, b, c\}$ with unlimited repetitions, such that the numbers of occurrences of both a and b are even. Determine T_n.

Solution A length-n *selection* with i_1 letters a, i_2 letters b, and i_3 letters c can be arranged into $\frac{n!}{i_1!i_2!i_3!}$ distinct *words*. Therefore,

$$T_n = \sum_{\substack{i_1+i_2+i_3=n \\ 2|i_1,\, 2|i_2}} \frac{n!}{i_1!i_2!i_3!}.$$

Consider the exponential generating function

$$\begin{aligned} g(x) &:= \sum_{\substack{i_1 \geqslant 0 \\ 2|i_1}} \frac{x^{i_1}}{i_1!} \sum_{\substack{i_2 \geqslant 0 \\ 2|i_2}} \frac{x^{i_2}}{i_2!} \sum_{i_3 \geqslant 0} \frac{x^{i_3}}{i_3!} = \left(\frac{e^x + e^{-x}}{2}\right)^2 e^x = \frac{e^{3x} + 2e^x + e^{-x}}{4} \\ &= \sum_{n=0}^{\infty} \frac{3^n + 2 + (-1)^n}{4} \cdot \frac{x^n}{n!}. \end{aligned}$$

By Fact 1.10.2,

$$T_n = n![x^n]g(x) = \frac{3^n + 2 + (-1)^n}{4}.$$

□

Exercise 1.10.5 Find the number a_n of ways to assign n students to four distinct classes so that each class has at least one student.

Solution Since

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

we have by Fact 1.10.2:

$$\sum_{n=0}^{\infty} \frac{a_n}{n!} x^n = \left(\sum_{k=1}^{\infty} \frac{x^k}{k!}\right)^4 = (e^x - 1)^4 = e^{4x} - 4e^{3x} + 6e^{2x} - 4e^x + 1.$$

Extracting the coefficient, we find that

$$a_n = 4^n - 4 \cdot 3^n + 6 \cdot 2^n - 4, \quad \forall n \geqslant 1.$$

□

Exercise 1.10.5 can be extended to the following more general setting, which is closely connected to the Stirling numbers of the second kind:

Example 1.10.6 Let n be a positive integer and fix $1 \leqslant k \leqslant n$. By Fact 1.2.4, the quantity

$$a_n^{(k)} := k! \left\{ {n \atop k} \right\}$$

counts the number of surjective maps from $[n]$ to $[k]$. Each surjection $f : [n] \to [k]$ is uniquely determined

by the ordered k-tuple of its nonempty preimage sets:

$$\big(f^{-1}(1), f^{-1}(2), \ldots, f^{-1}(k)\big).$$

This characterization implies that

$$a_n^{(k)} = \sum_{\substack{i_1+\cdots+i_k=n \\ i_j \geqslant 1}} \frac{n!}{i_1! \cdots i_k!}.$$

Using Fact 1.10.2, the exponential generating function of the sequence $\left(a_n^{(k)}\right)$ is

$$g_k(x) = \sum_{n=0}^{\infty} \frac{a_n^{(k)}}{n!} x^n = \left(\sum_{j=1}^{\infty} \frac{x^j}{j!} \right)^k = (e^x - 1)^k.$$

From this, we deduce

$$\begin{aligned} \left\{ n \atop k \right\} = \frac{a_n^{(k)}}{k!} &= \frac{n!}{k!} [x^n](e^x - 1)^k \\ &= \frac{n!}{k!} [x^n] \left\{ \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} \sum_{n=0}^{\infty} \frac{i^n}{n!} x^n \right\} \\ &= \frac{1}{k!} \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} i^n, \end{aligned}$$

which recovers the explicit formula in Theorem 1.5.7.

Exercise 1.10.7 Let a_n be the number of arrangements of n people according to rule A, and let b_n be the number of arrangements according to rule B. Define a new arrangement rule C for n people as follows:

- ◇ Divide the n people into two groups (say, the 1st group and the 2nd group).
- ◇ Arrange the 1st group according to rule A, and arrange the 2nd group according to rule B.

Let c_n be the number of arrangements of n people according to rule C. Let $A(x)$, $B(x)$, and $C(x)$ denote the exponential generating functions of the sequences (a_n), (b_n), and (c_n), respectively. Prove that

$$C(x) = A(x)B(x).$$

Proof Note that for each n,

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

By Fact 1.10.3, we have $C(x) = A(x)B(x)$. □

Chapter 2 Preliminaries in Graph Theory

2.1 Basic Definitions

Definition 2.1.1 A **graph** is an ordered pair $G = (V, E)$ comprising:

- V, a set of **vertices** (also called **nodes** or **points**);
- $E \subset \binom{V}{2} = \{(x, y) : x, y \in V \text{ and } x \neq y\}$, a set of **edges** (also called **links** or **lines**), which are unordered pairs of vertices (that is, an edge is associated with two distinct vertices).

In the edge $\{x, y\}$, the vertices x and y are called the **endpoints** of the edge. The edge is said to **join** x and y and to be **incident** on x and on y. The endpoints x and y are said to be **adjacent** to one another, denoted by $x \sim_G y$, $x \sim y$, or $xy \in E$.

Remark 2.1.2 (1) This type of object may be called an **undirected simple graph**.

(2) A vertex may exist in a graph and not belong to an edge. Under this definition, **multiple edges**, in which two or more edges connect the same vertices, are not allowed.

(3) A **loop** is an edge that joins a vertex to itself. Graphs as defined above cannot have loops.

Definition 2.1.3 A **subgraph** of a graph G is another graph formed from a subset of the vertices and edges of G. The vertex subset must include all endpoints of the edge subset, but may also include additional vertices.

Definition 2.1.4 Let $G = (V, E)$ be any graph, and let $S \subset V$ be any subset of vertices of G. Then the **induced subgraph** $G[S]$ is the graph whose vertex set is S and whose edge set consists of all of the edges in E that have both endpoints in S. That is, for any two vertices $u, v \in S$, u and v are adjacent in $G[S]$ if and only if they are adjacent in G.

Definition 2.1.5 Let $G = (V, E)$ be any graph.

(1) The **order** of G is $|V|$, its number of vertices.

(2) The **size** of G is $|E|$, its number of edges.

(3) The **degree** or **valency** of a vertex v, denoted $d_G(v)$, is the number of edges that are incident to v, where a loop is counted twice. The **degree** of a graph is the maximum of the degrees of its vertices. A simple graph is said to be **regular** of degree r, or simply r**-regular**, if all vertex degrees are the same number r.

(4) The **neighborhood** of a vertex v in G, denoted $N_G(v)$, is the subgraph of G induced by all vertices adjacent to v, i.e., the graph composed of the vertices adjacent to v and all edges connecting vertices adjacent to v.

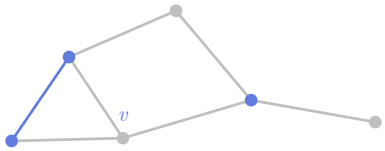


Figure 2.1: Neighborhood of v highlighted in the graph

Remark 2.1.6 (1) The same neighborhood notation may also be used to refer to sets of adjacent vertices rather than the corresponding induced subgraphs.

(2) The neighborhood described above does not include v itself.

Definition 2.1.7 The **complement** or **inverse** of a graph G is a graph H on the same vertices such that two distinct vertices of H are adjacent if and only if they are not adjacent in G.

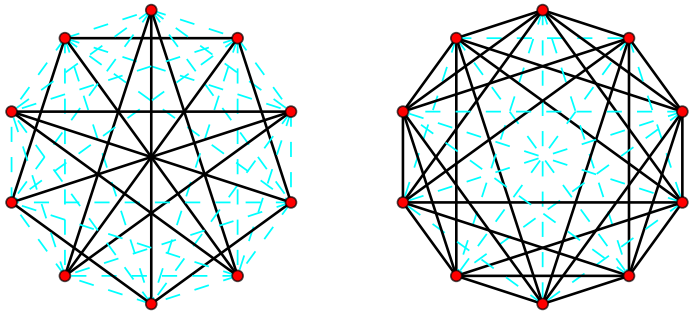


Figure 2.2: The Petersen graph (on the left) and its complement graph (on the right)

Definition 2.1.8 An **isomorphism of graphs** G and H is a bijection between the vertex sets of G and H

$$f\colon V(G) \to V(H)$$

such that any two vertices u and v of G are adjacent in G if and only if $f(u)$ and $f(v)$ are adjacent in H.

Definition 2.1.9 A **complete graph** is a simple undirected graph in which every pair of distinct vertices is connected by a unique edge. The complete graph on n vertices is denoted by K_n.[1]

Fact 2.1.10 K_n has $\frac{n(n-1)}{2}$ edges.

[1]Some sources claim that the letter K in this notation stands for the German word *komplett*, but the German name for a complete graph, *vollständiger Graph*, does not contain the letter K, and other sources state that the notation honors the contributions of Kazimierz Kuratowski to graph theory.

Definition 2.1.11 A **clique** is a subset of vertices of an undirected graph such that every two distinct vertices in the clique are adjacent. That is, a clique of a graph G is an induced subgraph of G that is complete.

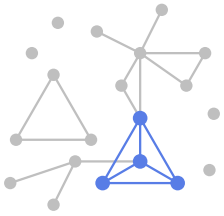


Figure 2.3: Maximum clique highlighted in the graph

Definition 2.1.12 An **independent set**, **stable set**, **coclique** or **anticlique** is a set of vertices in a graph, no two of which are adjacent. Equivalently, each edge in the graph has at most one endpoint in S. The independent set on n vertices is denoted by I_n.

Remark 2.1.13 A set is independent if and only if it is a clique in the graph's complement.

Definition 2.1.14 A **maximum independent set** is an independent set of largest possible size for a given graph G. The cardinality of a maximum independent set is called the **independence number** of G and is denoted by $\alpha(G)$.

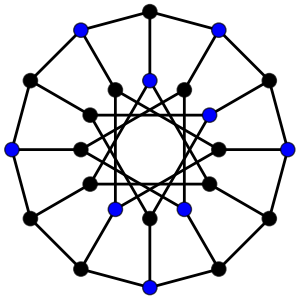


Figure 2.4: Nine blue vertices forming a maximum independent set

Definition 2.1.15 Given an undirected graph, a **degree sequence** is a monotonic nonincreasing sequence of the vertex degrees (valencies) of its graph vertices.

Definition 2.1.16 A **walk** is a finite or infinite sequence of edges which joins a sequence of vertices. The **length** of a walk is its number of edges.

Definition 2.1.17 A **trail** is a walk in which all edges are distinct.

Definition 2.1.18 A **path** is a trail in which all vertices (and therefore also all edges) are distinct.

Definition 2.1.19 A **cycle** is a non-empty trail in which only the first and last vertices are equal. The cycle graph with n vertices is called C_n.

Definition 2.1.20 In an undirected graph G, two vertices u and v are called **connected** if G contains a path from u to v. Otherwise, they are called **disconnected**.

Definition 2.1.21 A graph is said to be **connected** if every pair of vertices in the graph is connected. An undirected graph that is not connected is called **disconnected**.

Definition 2.1.22 A **component** of a graph is a connected subgraph that is not part of any larger connected subgraph.

Definition 2.1.23 Let G be a simple graph with vertex set $\{v_1, \ldots, v_n\}$ and edge set $\{e_1, \ldots, e_m\}$.

(1) The **adjacency matrix** $A(G)$ is a square $n \times n$ matrix whose (i, j)-th entry is 1 if v_i and v_j are adjacent in G, and 0 otherwise. The diagonal elements of the matrix are all 0, since edges from a vertex to itself (loops) are not allowed in simple graphs.

(2) The **incidence matrix** $M(G)$ is an $n \times m$ matrix whose (i, j)-th entry is 1 if vertex v_i is incident with edge e_j, and 0 otherwise.

Definition 2.1.24 A **planar graph** is a graph that can be embedded in the plane, i.e., it can be drawn on the plane in such a way that its edges intersect only at their endpoints.

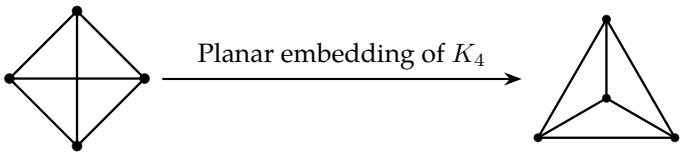


Figure 2.5: K_4 is planar

Definition 2.1.25 A **bipartite graph** (or **bigraph**) is a graph whose vertices can be divided into two disjoint and independent sets U and V that is, every edge connects a vertex in U to one in V. Vertex sets U and V are usually called the **parts** of the graph. Equivalently, a bipartite graph is a graph that does not contain any odd-length cycles. One often writes $G = (U, V, E)$ to denote a bipartite graph whose partition has the parts U and V, with E denoting the edges of the graph.

Definition 2.1.26 A **complete bipartite graph** or **biclique** is a special kind of bipartite graph where every vertex of the first set is connected to every vertex of the second set. A complete bipartite graph with partitions of size m and n is denoted as $K_{m,n}$.

Definition 2.1.27 A **biregular graph** or **semiregular bipartite graph** is a bipartite graph $G = (U, V, E)$ for which every two vertices on the same side of the given bipartition have the same degree as each other. If the degree of the vertices in U is x and the degree of the vertices in V is y, then the graph is said to be (x, y)-biregular.

2.2 Basic Graph Theorems

Theorem 2.2.1 (Euler's formula) If a finite, connected, planar graph is drawn in the plane without any edge intersections, and v is the number of vertices, e is the number of edges and f is the number of faces (regions bounded by edges, including the outer, infinitely large region), then

$$v - e + f = 2.$$

Corollary 2.2.2 If a connected simple planar graph has $v \geqslant 3$ vertices and e edges, then $e \leqslant 3v - 6$.

Proof Any face (except possibly the outer one) is bounded by at least three edges, and every edge touches at most two faces. Therefore,

$$3f \leqslant 2e.$$

By Theorem 2.2.1,

$$2 = v - e + f \leqslant v - e + \frac{2}{3}e,$$

which simplifies to

$$e \leqslant 3v - 6.$$

□

Example 2.2.3 K_5 has 5 vertices and 10 edges, and thus by Corollary 2.2.2 it is not planar.

Exercise 2.2.4 Prove that $K_{3,3}$ is not planar.

Proof Suppose to the contrary that $K_{3,3}$ is planar. Since $K_{3,3}$ is bipartite, it contains no 3-cycles (see Definition 2.1.25). So each face of the planar embedding of $K_{3,3}$ must be bounded by at least 4 edges. Moreover, each edge is counted twice among the boundaries for faces. Hence, we must have

$$f \leqslant \frac{2e}{4} = 4.5.$$

Now, by Theorem 2.2.1,

$$2 = v - e + f \leqslant 6 - 9 + 4.5 = 1.5,$$

a contradiction. Therefore, $K_{3,3}$ is not planar. □

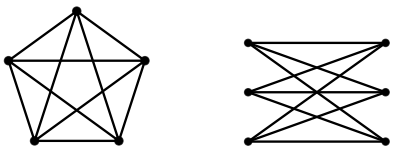


Figure 2.6: K_5 and $K_{3,3}$ are not planar

Lemma 2.2.5 (Degree sum formula) For any graph $G = (V, E)$,

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

Proof We count the number of incident pairs (v, e), where e is an edge and v is one of its endpoints, in two different ways. First, each vertex v belongs to $d_G(v)$ incident pairs, so the total number of incident

pairs is

$$\sum_{v \in V} d_G(v).$$

On the other hand, each edge in the graph belongs to exactly two incident pairs, one for each of its endpoints. Therefore, the number of incident pairs is $2|E|$. □

Theorem 2.2.6 (Handshaking lemma) In any finite graph, the number of vertices with odd degree is even. For example, if there is a party of people who shake hands, the number of people who shake an odd number of other people's hands is even.

Proof Since one side of Lemma 2.2.5 is the even number $2|E|$, the sum on the other side must have an even number of odd terms; that is, there must be an even number of odd-degree vertices. □

Chapter 3 Sperner's Lemma and Theorem

3.1 Sperner's Lemma and Brouwer's Fixed Point Theorem

In one dimension, Sperner's lemma can be regarded as a discrete version of the intermediate value theorem.

Fact 3.1.1 (Sperner's lemma, one-dimensional case) If a discrete function takes only the values 0 and 1, begins at the value 0 and ends at the value 1, then it must switch values an odd number of times.

Figure 3.1: One-dimensional case example

Theorem 3.1.2 (Sperner's lemma, two-dimensional case) Subdivide a triangle ABC arbitrarily into a triangulation consisting of smaller triangles meeting edge to edge. Then a **Sperner coloring** of the triangulation is defined as an assignment of three colors to the vertices of the triangulation such that

- Each of the three vertices A, B, and C of the initial triangle has a distinct color.
- The vertices that lie along any edge of triangle ABC have only two colors, the two colors at the endpoints of the edge. For example, each vertex on AC must have the same color as A or C.

Then every Sperner coloring of every triangulation has at least one "**rainbow triangle**", a smaller triangle in the triangulation that has its vertices colored with all three different colors. More precisely, there must be an odd number of rainbow triangles.

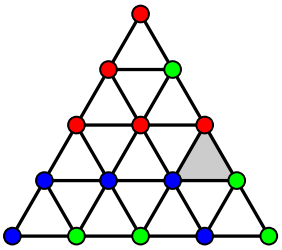


Figure 3.2: Two-dimensional case example: a Sperner coloring, with its 3-colored triangle shaded

Proof Consider a graph G built from the triangulation T as follows. The vertices of G are the members of T plus the area outside the triangle. Two vertices are connected with an edge if their corresponding areas share a common border with one endpoint colored 1 and the other colored 2.

By Fact 3.1.1, on the interval AB there is an odd number of borders colored 1-2. On the intervals BC and CA, there are no borders colored 1-2 at all. Therefore, the vertex of G corresponding to the outer area has an odd degree. By Theorem 2.2.6, the remaining graph, excluding the outer area, has an odd number of vertices with odd degree corresponding to members of T.

It can be easily seen that the only possible degree of a triangle from T is 0, 1, or 2, and that the degree 1 corresponds to a triangle colored with the three colors 1, 2, and 3.

Thus we have obtained a slightly stronger conclusion, which says that in a triangulation T there is an odd number (and at least one) of full-colored triangles. □

Theorem 3.1.3 (Sperner's lemma, multidimensional case) Consider an n-dimensional simplex

$$\mathcal{A} = A_1 A_2 \cdots A_{n+1},$$

and any triangulation T, a disjoint division of $\mathcal{A}$ into smaller n-dimensional simplices, meeting face-to-face. Denote the coloring function as $f\colon S \to [n+1]$, where S is the set of vertices of T. A coloring function defines a **Sperner coloring** when:

- ◇ The vertices of the large simplex are colored with different colors, that is, without loss of generality, $f(A_i) = i$ for $i \in [n+1]$.
- ◇ Vertices of T located on any k-dimensional subface of the large simplex $A_{i_1} A_{i_2} \cdots A_{i_{k+1}}$ are colored only with the colors $i_1, i_2, \ldots, i_{k+1}$.

Then every Sperner coloring of every triangulation of the n-dimensional simplex has an odd number of instances of a **rainbow n-simplex**, meaning an n-simplex whose vertices are colored with all $n+1$ colors. In particular, there must be at least one rainbow simplex.

Proof Assume the statement holds for dimension $n-1$ with $n \geqslant 3$. Define

$$\begin{aligned} R &\coloneqq \{\text{rainbow } n\text{-simplices in } T\}, \\ Q &\coloneqq \{n\text{-simplices in } T \text{ using exactly the colors } 1, 2, \ldots, n\}, \\ X &\coloneqq \{\text{boundary } (n-1)\text{-faces of } T \text{ using colors } 1, 2, \ldots, n\}, \\ Y &\coloneqq \{\text{interior } (n-1)\text{-faces of } T \text{ using colors } 1, 2, \ldots, n\}. \end{aligned}$$

We count, in two ways, the set

$$\mathcal{P} = \{(F, \Delta) : \Delta \text{ is an } n\text{-simplex of } T,\ F \text{ is an } (n-1)\text{-face of } \Delta \text{ using colors } 1, 2, \ldots, n\}.$$

- ◇ Each simplex of type R contributes exactly one $(n-1)$-face colored $\{1, 2, \ldots, n\}$, and each simplex of type Q contributes exactly two such faces. Hence $|\mathcal{P}| = |R| + 2|Q|$.
- ◇ Every interior $(n-1)$-face colored $\{1, 2, \ldots, n\}$ is shared by two n-simplices, while each boundary such face belongs to exactly one n-simplex. Therefore $|\mathcal{P}| = |X| + 2|Y|$.

Equating the two counts gives

$$|R| + 2|Q| = |X| + 2|Y|. \tag{3–1}$$

By the Sperner condition, the only boundary $(n-1)$-faces that can use the colors $1, 2, \ldots, n$ are those lying in the facet $A_1A_2 \cdots A_n$. The restriction of T and f to this facet forms a Sperner coloring of an $(n-1)$-simplex, so by the induction hypothesis it contains an odd number of rainbow $(n-1)$-simplices. Hence $|X|$ is odd, and by (3–1) $|R|$ is odd as well. □

Theorem 3.1.4 (Brouwer's fixed-point theorem) Every continuous function f from a closed ball in $\mathbb{R}^n$ to itself has a fixed point.

Proof It is convenient to work with an n-simplex instead of a closed ball, since the two are homeomorphic. Specifically, consider the standard n-simplex

$$\Delta^n = \left\{x \in \mathbb{R}^{n+1} : x_0 + x_1 + \cdots + x_n = 1 \text{ and } x_i \geqslant 0 \text{ for all } i\right\}.$$

For every $x \in \Delta^n$, we also have $f(x) \in \Delta^n$, so the sum of the coordinates satisfies

$$x_0 + x_1 + \cdots + x_n = 1 = f(x)_0 + f(x)_1 + \cdots + f(x)_n.$$

By the pigeonhole principle, there exists at least one index $j \in \{0, 1, \ldots, n\}$ such that $x_j \geqslant f(x)_j$. Moreover, if x lies on a k-dimensional face of Δ^n, then by the same argument, the index j can be chosen from the $k+1$ coordinates corresponding to that face.

We now use this observation to define a Sperner coloring. Let $\Delta_1^n, \Delta_2^n, \ldots$ be a sequence of increasingly finer subdivisions of Δ^n, with Δ_k^n refining Δ_{k-1}^n and the diameters of the simplices in Δ_k^n tending to zero as $k \to \infty$. Using the above rule, assign to each vertex x of Δ_k^n a color $j \in \{0, 1, \ldots, n\}$ such that $f(x)_j \leqslant x_j$. This gives a Sperner coloring on each subdivision Δ_k^n.

By Theorem 3.1.3, each Δ_k^n contains an n-simplex $\mathcal{B}_k^n$ whose vertices are colored with all $n+1$ colors. Since Δ^n is compact, the sequence of simplices $\mathcal{B}_k^n$ has a convergent subsequence whose limit is a point $x^* \in \Delta^n$. By continuity of f, we have $f(x^*)_j \leqslant (x^*)_j$ for all j. Because the sum of the coordinates of x^* and $f(x^*)$ are equal, all these inequalities must actually be equalities. Hence $f(x^*) = x^*$. □

3.2 Double Counting Example in Number Theory

We have seen how double counting can be used to prove Sperner's lemma in §3.1. Here is another application of double counting from number theory. If $t(n)$ is the number of divisors of n, then the behavior of this function is rather non-uniform: $t(p) = 2$ for every prime number, whereas $t(2^m) = m+1$. It is therefore interesting that the *average* number

$$\tau(n) := \frac{t(1) + t(2) + \cdots + t(n)}{n}$$

of divisors is quite stable: It is about $\ln n$.

Proposition 3.2.1 $|\tau(n) - \ln n| \leqslant 1$.

Proof To apply the double counting principle, consider the binary $n \times n$ matrix $M = (m_{i,j})$ with $m_{i,j} = 1$ if and only if $i \mid j$:

	1	2	3	4	5	6	7	8	9	10	11	12
1	1	1	1	1	1	1	1	1	1	1	1	1
2		1		1		1		1		1		1
3			1			1			1			1
4				1				1				1
5					1					1		
6						1						1
7							1					
8								1				

The number of 1s in the j-th column is exactly the number $t(j)$ of divisors of j. So, summing over columns we see that the total number of 1s in the matrix is $T(n) = t(1) + \cdots + t(n)$.

On the other hand, the number of 1s in the i-th row is the number of multipliers $i, 2i, 3i, \ldots, ri$ of i such that $ri \leqslant n$. Hence, we have exactly $\lfloor n/i \rfloor$ ones in the i-th row. Summing over rows, we obtain that

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

Since $x - 1 < \lfloor x \rfloor \leqslant x$ for every real number x, we have

$$H_n - 1 \leqslant \tau(n) = \frac{T_n}{n} \leqslant H_n,$$

where

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} = \ln n + \gamma_n, \quad 0 \leqslant \gamma_n \leqslant 1$$

is the n-th harmonic number. □

3.3 Sperner's Theorem

Definition 3.3.1 A family of sets in which none of the sets is a strict subset of another is called a **Sperner family**, also known as an **antichain of sets**, a **clutter**, an **independent system**, or an **irredundant set**.

Example 3.3.2 The family of k-element subsets of an n-element set is a Sperner family. No set in this family can contain any of the others, because a containing set has to be strictly bigger than the set it contains, and in this family all sets have equal size.

The value of k that makes Example 3.3.2 have as many sets as possible is $\lfloor n/2 \rfloor$. For this choice, the number of sets in the family is $\binom{n}{\lfloor n/2 \rfloor}$. Sperner's theorem states that these examples are the largest possible Sperner families over an n-element set.

Theorem 3.3.3 (Sperner's theorem) For every Sperner family $\mathcal{S}$ whose union has a total of n elements,

$$|\mathcal{S}| \leqslant \binom{n}{\lfloor n/2 \rfloor}.$$

To establish Theorem 3.3.3, we first develop some auxiliary results. Throughout, let X denote an n-element set.

Definition 3.3.4 A **chain** in the power set 2^X is a collection of subsets of X ordered by inclusion.

Definition 3.3.5 Let $\mathcal{C} = \{A_1, \ldots, A_k\}$ be a chain in 2^X, i.e., $A_1 \subset A_2 \subset \cdots \subset A_k$.

(1) This chain is **symmetric** if $|A_1| + |A_k| = n$ and $|A_{i+1}| = |A_i| + 1$ for all $i \in [k-1]$.

(2) Symmetric chains with $k = n + 1$ are called **maximal**.

Remark 3.3.6 (1) "Symmetric" here means symmetric positioned about the middle level $n/2$.

(2) Maximal chains are in one-to-one correspondence with the permutations of the underlying set $[n]$: every permutation $(x_1, \ldots, x_n)$ gives the maximal chain

$$\varnothing \subset \{x_1\} \subset \{x_1, x_2\} \subset \cdots \subset \{x_1, \ldots, x_n\}.$$

Hence the number of maximal chains in $2^{[n]}$ is $n!$.

The following theorem is due to de Bruijn, Tengbergen, and Kruyswijk (1952).

Theorem 3.3.7 The power set 2^X can be partitioned into $\binom{n}{\lfloor n/2 \rfloor}$ mutually disjoint symmetric chains.

Proof 1 Assume for a moment that we already have some partition of 2^X into symmetric chains. Every such chain contains exactly one set from the middle level; hence there are $\binom{n}{\lfloor n/2 \rfloor}$ chains in that partition.

Let us now prove that such a partition is possible at all. We argue by the induction on $n = |X|$. Clearly the result holds for the one point set X. So, suppose that it is true for all sets with fewer points than n. Pick a point $x \in X$, and let $Y = X \setminus \{x\}$. By induction, we can partition 2^Y into symmetric chains $\mathcal{C}_1, \ldots, \mathcal{C}_r$. Each of these chains over Y

$$\mathcal{C}_i: \quad A_1 \subset A_2 \subset \cdots \subset A_k$$

produces the following two chains over the whole set X:

$$\begin{aligned} \mathcal{C}_i' &: \quad A_1 \subset A_2 \subset \cdots \subset A_{k-1} \subset A_k \subset A_k \cup \{x\}, \\ \mathcal{C}_i'' &: \quad A_1 \cup \{x\} \subset A_2 \cup \{x\} \subset \cdots \subset A_{k-1} \cup \{x\}. \end{aligned}$$

These chains are symmetric since

$$|A_1| + |A_k \cup \{x\}| = (|A_1| + |A_k|) + 1 = (n-1) + 1 = n$$

and

$$|A_1 \cup \{x\}| + |A_{k-1} \cup \{x\}| = (|A_1| + |A_{k-1}|) + 2 = (n-2) + 2 = n.$$

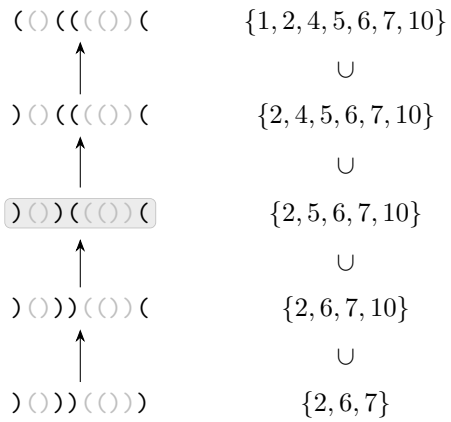
This is indeed a partition:

- ◇ If $A \subset Y$, then only $\mathcal{C}_i'$ contains A, where $\mathcal{C}_i$ is the chain in 2^Y containing A.
- ◇ If $A = B \cup \{x\}$ where $B \subset Y$, then $B \in \mathcal{C}_i$ for some i. If B is the maximal element of $\mathcal{C}_i$, then $\mathcal{C}_i'$ is the only chain containing A, otherwise A is contained only in $\mathcal{C}_i''$. □

Proof 2 Let $X = [n]$. For each subset $A \subset X$, associate a word $w(A)$ whose i-th letter is an open or closed parenthesis:

$$w(A)_i = \begin{cases} (, & \text{if } i \in A, \\ ), & \text{if } i \notin A. \end{cases}$$

For example, if $n = 10$ and $A = \{2, 5, 6, 7, 10\}$, then $w(A)$ is)()((()(. Here the matched parentheses are shown in gray. To obtain the symmetric chain containing A, we define how to "go up" and "go down" along the chain. To go up, take the rightmost unmatched) and flip it to (; to go down, flip the leftmost unmatched (to).



This clearly defines a symmetric chain. It remains to check that any two chains constructed this way are either identical or disjoint. But this is immediate: flipping the rightmost) or the leftmost (can never produce a new matched pair ().

□

A considerably sharper result than Theorem 3.3.3, Theorem 3.3.8 below, is due to Lubell (1966). The same result was discovered by Meshalkin (1963) and (not so explicitly) by Yamamoto (1954).

Theorem 3.3.8 (LYM inequality) Let $\mathcal{S}$ be a Sperner family over a set X of n elements. Then

$$\sum_{A\in\mathcal{S}} \binom{n}{|A|}^{-1} \leqslant 1.$$

Proof For each $A \in \mathcal{S}$, exactly $|A|!(n-|A|)!$ maximal chains over X contain A (cf. Remark 3.3.6 (2)). Since none of the $n!$ maximal chains meet $\mathcal{S}$ more than once, we have

$$\sum_{A\in\mathcal{S}} |A|!(n-|A|)! \leqslant n!.$$

Dividing this inequality by $n!$ gives the desired result. □

Proof 1 of Sperner's theorem 3.3.3 By Fact 1.4.5 and Theorem 3.3.8,

$$|\mathcal{S}| \cdot \binom{n}{\lfloor n/2\rfloor}^{-1} \leqslant \sum_{A\in\mathcal{S}} \binom{n}{|\mathcal{S}|}^{-1} \leqslant 1.$$

□

Proof 2 of Sperner's theorem 3.3.3 By Theorem 3.3.7, the power set of the union of $\mathcal{S}$ can be partitioned into $\binom{n}{\lfloor n/2\rfloor}$ symmetric chains, none of which meets $\mathcal{S}$ more than once. Hence,

$$|\mathcal{S}| \leqslant \binom{n}{\lfloor n/2\rfloor}.$$

□

3.4 Littlewood–Offord Problem

Question Given a finite subset $S = \{v_1, \dots, v_n\} \subset \mathbb{R}$ and a convex subset $A \subset \mathbb{R}$, what is the number of subsets of S whose summation is in A?

The first upper bound for this problem was proven in 1938 by John Edensor Littlewood and A. Cyril Offord. This Littlewood–Offord lemma states that if S is a set n real numbers of absolute value at least one and A is any disc of radius one, then not more than $(c \log n/\sqrt{n})2^n$ of the 2^n possible subsums of S fall into the disc.

In 1945 Paul Erdős improved the upper bound to $\binom{n}{\lfloor n/2\rfloor} \approx \frac{2^n}{\sqrt{n}}$ using Sperner's theorem. This bound is sharp; equality is attained when all numbers in S are equal.

By subtracting $\frac{1}{2}(v_1 + \cdots + v_n)$ from each possible subsum (that is, by changing the origin and then scaling by a factor of 2), the Littlewood–Offord problem reduces to determining the number of sums of the form $\varepsilon_1 v_1 + \cdots + \varepsilon_n v_n$ that fall in the target set A, where ε_i takes the value 1 or -1.

Theorem 3.4.1 Let $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ with $|v_i| \geqslant 1$ for all i. Define

$$F(v) = \{\varepsilon \in \{-1, 1\}^n : \varepsilon \cdot v \in (-1, 1)\}.$$

Then

$$|F(v)| \leqslant \binom{n}{\lfloor n/2\rfloor}.$$

Proof For each $\varepsilon \in F(v)$, define $V_\varepsilon = \{i \in [n] : \varepsilon_i v_i > 0\}$, and let $\mathcal{S} = \{V_\varepsilon : \varepsilon \in F(v)\}$.

We claim that $\mathcal{S}$ is a Sperner family. Suppose to the contrary that $V_\varepsilon \subsetneq V_\mu$ for some $\varepsilon, \mu \in F(v)$. Since $\varepsilon \cdot v,\ \mu \cdot v \in (-1, 1)$, we have

$$|\varepsilon \cdot v - \mu \cdot v| < 2.$$

On the other hand,

$$\varepsilon \cdot v = \sum_{i \in V_\varepsilon} |v_i| - \sum_{i \notin V_\varepsilon} |v_i| = 2 \sum_{i \in V_\varepsilon} |v_i| - \sum_{i=1}^{n} |v_i|,$$
$$\mu \cdot v = \sum_{i \in V_\mu} |v_i| - \sum_{i \notin V_\mu} |v_i| = 2 \sum_{i \in V_\mu} |v_i| - \sum_{i=1}^{n} |v_i|.$$

Thus,

$$\mu \cdot v - \varepsilon \cdot v = 2 \left(\sum_{i \in V_\mu} |v_i| - \sum_{i \in V_\varepsilon} |v_i| \right) \geqslant 2|v_k| \geqslant 2$$

for some $k \in V_\mu \setminus V_\varepsilon$, a contradiction. Hence $\mathcal{S}$ is indeed a Sperner family.

By Theorem 3.3.3, it follows that

$$|F(v)| = |\mathcal{S}| \leqslant \binom{n}{\lfloor n/2 \rfloor}.$$

□

Chapter 4 Forbidden Subgraph Problem

4.1 Basic Definitions

Definition 4.1.1 A graph is called G**-free** if it contains no subgraph isomorphic to G. The **extremal number** $\mathrm{ex}(n, G)$ is the maximum number of edges in an n-vertex G-free graph.

Remark 4.1.2 In this context, G is called a **forbidden subgraph**.

Example 4.1.3 The complete bipartite graph $K_{m,n}$ is K_3-free (cf. Definition 2.1.25).

4.2 Upper Bounds

Theorem 4.2.1 (Reiman's theorem) For $n \geqslant 4$, $\mathrm{ex}(n, C_4) \leqslant \frac{n}{4}\left(1 + \sqrt{4n-3}\right)$.

Proof Let G be a C_4-free graph with n vertices and m edges. We count pairs $(c, \{a, b\})$ where a, b, c are vertices of G such that c is adjacent to both a and b. In other words, we double count **cherries**.

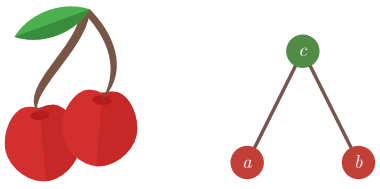


Figure 4.1: A cherry is a path on three vertices

If c is the middle of a cherry it contributes $\binom{d_G(c)}{2}$. Hence the total number of cherries is

$$\frac{1}{2}\sum_{c\in V(G)} d_G(c)^2 - \frac{1}{2}\sum_{c\in V(G)} d_G(c).$$

By the Cauchy–Schwarz inequality and Lemma 2.2.5, this number is

$$\left(\frac{1}{2n}\sum_{c\in V(G)} d_G(c)^2 \sum_{c\in V(G)} 1^2\right) - m \geqslant \frac{1}{2n}\left(\sum_{c\in V(G)} d_G(c)\cdot 1\right)^2 - m = \frac{2m^2}{n} - m.$$

Since there are no C_4 in G, for fixed a, b there is at most one vertex c that is adjacent to both a and b. Thus the number of cherries is at most $\binom{n}{2}$ and so

$$\frac{2m^2}{n} - m \leqslant \frac{n(n-1)}{2}.$$

This implies

$$m \leqslant \frac{n}{4}\left(1+\sqrt{4n-3}\right).$$

□

Exercise 4.2.2 Show that Theorem 4.2.1 can be improved to $\mathrm{ex}(n, C_4) < \frac{n}{4}\left(1+\sqrt{4n-3}\right)$.

Proof From the proof of Theorem 4.2.1 we know that equality in the bound requires:

- ◇ G is regular.
- ◇ Every pair of vertices shares exactly one common neighbor.

But a graph where every two vertices share exactly one neighbor is exactly the setting of the **Friendship theorem**, which states that such a graph must be a **Dutch windmill graph**, i.e., a collection of triangles sharing exactly one common vertex (see Figure 4.2 for example). However, Dutch windmill graphs are regular only in the trivial case K_3. Since $n \geqslant 4$, we cannot have equality. □

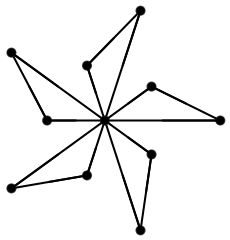


Figure 4.2: The Dutch windmill graph with 5 triangles

We now present a slight generalization of Theorem 4.2.1 (note that $C_4 \simeq K_{2,2}$):

Theorem 4.2.3 (Kővári–Sós–Turán theorem) For every pair of positive integers s, t with $t \geqslant s \geqslant 1$ and any positive integer n,

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

Proof Let G be a $K_{s,t}$-free graph with n vertices and m edges, and call $1, 2, \ldots, n$ its vertices. Let d_i be the degree of vertex i. We count pairs $(v, \{v_1, \ldots, v_s\})$, where $\{v_1, \ldots, v_s\}$ is a set of s distinct neighbors of v ("generalized cherries"). For each $v = i$, there are precisely $\binom{d_i}{s}$ sets $\{v_1, \ldots, v_s\}$ sharing a pair with v. Thus we find

$$\sum_{i=1}^{n} \binom{d_i}{s}$$

pairs in total. On the other hand, for each set of s vertices $\{v_1, \ldots, v_s\}$ there are at most $t-1$ vertices v adjacent to all of them, since G is $K_{s,t}$-free. Thus

$$\sum_{i=1}^{n} \binom{d_i}{s} \leqslant (t-1)\binom{n}{s}.$$

Since $f(x) \coloneqq \binom{x}{s}$ is convex[1], by Jensen's inequality and Lemma 2.2.5 we have

$$\sum_{i=1}^{n} \binom{d_i}{s} = \sum_{i=1}^{n} f(d_i) \geqslant nf\left(\frac{1}{n}\sum_{i=1}^{n} d_i\right) = n\binom{\frac{2m}{n}}{s}.$$

Hence we obtain the estimate

$$(t-1)n^s \geqslant n\left(\frac{2m}{n} - s + 1\right)^s,$$

from which the result follows. □

By Theorems 4.2.1 and 4.2.3, we obtain the following corollary.

Corollary 4.2.4 (1) $\mathrm{ex}(n, C_4) \leqslant \left(\frac{1}{2} + o(1)\right)n^{3/2}$.

(2) There exists some constant $C = C(s,t)$ (independent of n) such that $\mathrm{ex}(n, K_{s,t}) \leqslant Cn^{2-1/s}$ for every positive integer n.

4.3 Lower Bound on $\mathrm{ex}(n, C_4)$

We now present an algebraic construction that shows that the bound in Corollary 4.2.4 (1) is asymptotically tight.

Theorem 4.3.1 (Erdős–Rényi–Sós theorem) $\mathrm{ex}(n, C_4) \geqslant \left(\frac{1}{2} - o(1)\right)n^{3/2}$.

Proof First, suppose that $n = p^2 - 1$ for some prime p. Consider the graph G with vertices elements of $\mathbb{F}_p^2 \setminus \{(0,0)\}$ and edges between vertices (x,y) and (a,b) if and only if $ax + by = 1$ in $\mathbb{F}_p$. This graph is C_4-free because a system of two linear equations in $\mathbb{F}_p$ cannot have more than one solution. A vertex (a,b) (assume $b \neq 0$) is connected to is connected to $\big(x, b^{-1}(1-ax)\big)$ for any $x \in \mathbb{F}_p$, for a total of at least $p-1$ edges (subtracted 1 in case $(a,b) = \big(x, b^{-1}(1-ax)\big)$). So there are at least

$$\frac{1}{2}(p^2-1)(p-1) = \left(\frac{1}{2} - o(1)\right)p^3 = \left(\frac{1}{2} - o(1)\right)n^{3/2}$$

edges, as desired. For general n, we can take $p = (1-o(1))\sqrt{n}$ with $p \leqslant \sqrt{n+1}$ (which is possible because there exists a prime p in the interval $\left[k - k^{0.525}, k\right]$ for sufficiently large k) and construct a graph using such p, then adding $n - p^2 + 1$ isolated vertices, which do not affect the asymptotic value. □

[1] To make this completely formal, we can extend $\binom{x}{s}$ to all real $x \geqslant 0$ by setting $\binom{x}{s} = 0$ for $x < s-1$.

4.4 Mantel's Theorem

How many edges are possible in a triangle-free graph G on n vertices? Certainly, G can have $\lfloor n^2/4 \rfloor$ edges without containing a triangle: just let G be the complete bipartite graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ (cf. Example 4.1.3 and Definition–Theorem 4.5.1). Indeed, $\lfloor n^2/4 \rfloor$ turns out to be the maximum possible number of edges: if we take one more edge then the graph will have a triangle.

Theorem 4.4.1 (Mantel's theorem) $\mathrm{ex}(n, K_3) = \lfloor n^2/4 \rfloor$.

Proof 1 Let G be a graph with n vertices and $m > n^2/4$ edges. Assume that G has no triangles. Then adjacent vertices have no common neighbors, so $d_G(x) + d_G(y) \leqslant n$ for each edge $\{x, y\} \in E(G)$. Summing over all edges of G, we have

$$\sum_{x \in V(G)} d_G(x)^2 = \sum_{e \in E(G)} \sum_{x \in e} d_G(x) = \sum_{\{x,y\} \in E(G)} [d_G(x) + d_G(y)] \leqslant mn.$$

On the other hand, using the Cauchy–Schwarz inequality and Lemma 2.2.5, we obtain

$$\sum_{x \in V(G)} d_G(x)^2 = \frac{1}{n} \sum_{x \in V(G)} d_G(x)^2 \sum_{x \in V(G)} 1^2 \geqslant \frac{1}{n} \left(\sum_{x \in V(G)} d_G(x) \cdot 1 \right)^2 = \frac{4m^2}{n}.$$

These two inequalities imply that $m \leqslant n^2/4$, contradicting the hypothesis. □

Proof 2 We argue by induction on n. The base cases $n = 1, 2$ are trivial. Assuming the result for $n < k$ where $k \geqslant 3$, we now consider a K_3-free graph G with k vertices and m edges. Let x and y be adjacent vertices in G, and let H be the induced subgraph on the remaining $k - 2$ vertices. As before, we have $d_G(x) + d_G(y) \leqslant k$. Since H is also K_3-free, by the inductive hypothesis it has at most $\lfloor (k-2)^2/4 \rfloor$ edges. Thus

$$m \leqslant \frac{(k-2)^2}{4} + d_G(x) + d_G(y) - 1 \leqslant \frac{(k-2)^2}{4} + k - 1 = \frac{k^2}{4}.$$

This completes the proof. □

Proof 3 Let $G = (V, E)$ be a graph on a set V of n vertices and assume that G has no triangles. Let $A \subset V$ be a maximum independent set. Since G is K_3-free, the neighbors of a vertex $x \in V$ form an independent set, and we infer $d(x) \leqslant |A|$ for all x.

The set $B \coloneqq V \setminus A$ meets every edge of G. Counting the edges of G according to their end-vertices in B, we obtain

$$|E| \leqslant \sum_{x \in B} d(x).$$

By the AM–GM inequality,

$$|E| \leqslant \sum_{x \in B} d(x) \leqslant |A| \cdot |B| \leqslant \left(\frac{|A| + |B|}{2} \right)^2 = \frac{n^2}{4}.$$

□

Remark 4.4.2 $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ is the unique K_3-free graph with n vertices and $\lfloor n^2/4 \rfloor$ edges (see Theorem 4.5.4).

4.5 Turán's Theorem

Definition–Theorem 4.5.1 The **Turán graph**, denoted by $T(n,r)$, is the complete r-partite graph obtained by partitioning a set of n vertices into r subsets of sizes as equal as possible, and connecting two vertices by an edge if and only if they belong to different subsets. If $n = qr + s$ with $0 \leqslant s < r$, then $T(n,r)$ is of the form

$$K_{q+1,q+1,\ldots,q,q},$$

with s parts of size $q+1$ and $r-s$ parts of size q. The number of edges is given by

$$\begin{aligned} t(n,r) &\coloneqq \binom{n}{2} - s\binom{q+1}{2} - (r-s)\binom{q}{2} \\ &= \frac{n(n-1) - s(q+1)q - (r-s)q(q-1)}{2} \\ &= \frac{(qr+s)^2 - s - 2qs - q^2 r}{2} \\ &= \frac{q^2r^2 + 2qrs + s^2 - s - 2qs - q^2 r}{2} \\ &= \frac{(r-1)\big(q^2 r + 2qs\big)}{2} + \binom{s}{2} \\ &= \frac{(r-1)\big(q^2 r^2 + 2qrs\big)}{2r} + \binom{s}{2} \\ &= \left(1 - \frac{1}{r}\right)\frac{n^2 - s^2}{2} + \binom{s}{2}. \end{aligned}$$

Remark 4.5.2 For any n, if we delete one vertex from each of the r parts of $T(n,r)$, we obtain a copy of $T(n-r,r)$. Moreover, each non-deleted vertex is adjacent to exactly $r-1$ deleted vertices. So we delete

$$(r-1)(n-r) + \binom{r}{2}$$

edges to obtain $T(n-r,r)$ from $T(n,r)$. This shows that

$$t(n,r) = t(n-r,r) + (r-1)(n-r) + \binom{r}{2}. \tag{4–1}$$

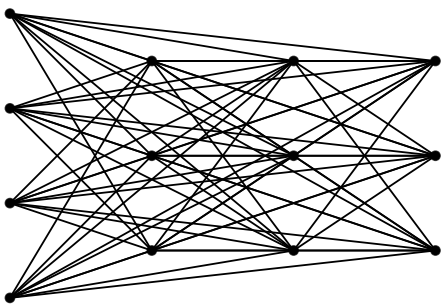


Figure 4.3: The Turán graph $T(13,4)$

Example 4.5.3 The Turán graph $T(n,r)$ is K_{r+1}-free (cf. Example 4.1.3).

Theorem 4.5.4 (Turán's theorem) For every $r \geqslant 2$, we have

$$\mathrm{ex}(n, K_{r+1}) = t(n, r).$$

Moreover, the unique K_{r+1}-free graph with n vertices and $t(n, r)$ edges is the Turán graph $T(n, r)$.

Proof We use induction on n, with steps of size r. So we need r base cases, corresponding to $n \in [r]$. But the theorem holds for such n, because for such n, any n-vertex graph is K_{r+1}-free. So $\mathrm{ex}(n, K_{r+1}) = \binom{n}{2}$ for $n \in [r]$. Moreover, $T(n, r)$ is exactly K_n in these cases. This proves the base cases of the induction.

Now let $n > r$, and assume the theorem is true for $n - r$. Let $G = (V, E)$ be a K_{r+1}-free graph with n vertices and with a maximal number of edges. This graph certainly contains K_r, since otherwise we could add edges. Let A be a subgraph of G isomorphic to K_r, and set $B = V \setminus A$.

Since each two vertices of A are joined by an edge, A contains $e_A = \binom{r}{2}$ edges. Let e_B be the number of edges joining the vertices of B and $e_{A,B}$ the number of edges between A and B. By induction, we have

$$e_B \leqslant \mathrm{ex}(n - r, K_{r+1}) = t(n - r, r).$$

Since G is K_{r+1}-free, every $x \in B$ is adjacent to at most $r - 1$ vertices in A, and we obtain

$$e_{A,B} \leqslant (r - 1)(n - r).$$

Summing up we conclude that

$$\begin{aligned} |E| &= e_A + e_B + e_{A,B} \\ &\leqslant \binom{r}{2} + t(n - r, r) + (r - 1)(n - r) \\ &= t(n, r) \end{aligned}$$

by (4–1).

If $|E| = t(n, r)$, then every inequality above must be an equality:

- ◇ The induction hypothesis implies that $B \simeq T(n - r, r)$.
- ◇ Each vertex in B must be adjacent to exactly $r - 1$ vertices in A.

By the second condition, given two adjacent vertices in B, they cannot be non-adjacent to the same vertex of A, for otherwise we could take the remaining $r - 1$ vertices and these two to get a K_{r+1}. This implies that each part of $B \simeq T(n - r, r)$ is associated to exactly one missed vertex. So by adding this missed vertex to its part, we see that $G \simeq T(n, r)$. □

Corollary 4.5.5 $\mathrm{ex}(n, K_{r+1}) \leqslant \left(1 - \frac{1}{r}\right)\frac{n^2}{2} = \binom{r}{2}\left(\frac{n}{r}\right)^2$.

Chapter 5 Trees

5.1 Basic Definitions

Definition 5.1.1 A **tree** is an undirected graph in which every pair of distinct vertices is connected by exactly one path, or equivalently, a connected graph without cycles.

Definition 5.1.2 In a tree, the following types of vertices are distinguished:

(1) An **internal vertex** (or **inner vertex**) is a vertex of degree at least 2.

(2) An **external vertex** (also called an **outer vertex**, **terminal vertex**, or **leaf**) is a vertex of degree 1.

(3) A **branch vertex** is a vertex of degree at least 3.

Fact 5.1.3 Let $T = (V, E)$ be a tree. Then $|V| = |E| + 1$.

Proof Every tree has at least one leaf. Otherwise, if all vertices had degree at least 2, one could follow edges indefinitely and obtain a cycle, contradicting acyclicity.

We proceed by induction on the number of vertices n. If $n = 2$, the tree consists of a single edge, so the statement holds. Assume the claim is true for all trees on $n - 1$ vertices, and let T be a tree on $n \geqslant 2$ vertices. Since T has a leaf, let v denote such a vertex. The graph $T \setminus \{v\}$, obtained by removing v and its incident edge, is still a tree on $n - 1$ vertices. By the induction hypothesis, $T \setminus \{v\}$ has $n - 2$ edges. Adding back v and its incident edge shows that T has $n - 1$ edges, as required. □

Fact 5.1.4 Every tree T with at least two vertices has at least two leaves.

Proof Suppose, for contradiction, that an n-vertex tree T has exactly one leaf v. Then every other vertex $u \in V(T) \setminus \{v\}$ satisfies $d_T(u) \geqslant 2$. By Fact 5.1.3 Lemma 2.2.5,

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

which is impossible. Hence T must have at least two leaves. □

Proposition 5.1.5 A tree is an undirected graph T that satisfies any of the following equivalent conditions:

(1) T is connected and acyclic.

(2) T is connected, but would become disconnected if any single edge is removed from T.

(3) T is acyclic, and a simple cycle is formed if any edge is added to T.

Proof $\boxed{(1) \Rightarrow (2)}$ Suppose (2) fails. Then there exists an edge $e = xy \in E(T)$ such that $T \setminus \{e\}$ remains connected. In particular, $T \setminus \{e\}$ contains a path P from x to y. Hence $P \cup \{e\}$ forms a cycle in T, contradicting (1).

$\boxed{(2) \Rightarrow (1)}$ Suppose (1) fails. Then T contains a cycle C. Deleting any edge e from C leaves $T \setminus \{e\}$ connected, contradicting (2).

$\boxed{(1) \Rightarrow (3)}$ Let $e = xy$ be any new edge. Since T is connected, there exists a path P from x to y in T. Then $P \cup \{e\}$ is a cycle, so (3) holds.

$\boxed{(3) \Rightarrow (1)}$ Suppose (1) fails, so T is disconnected. Then T has two components, say D_1 and D_2. Choose $x \in D_1$ and $y \in D_2$, and add the edge $e = xy$. The resulting graph $T \cup \{e\}$ is connected and still acyclic, contradicting (3). □

Definition 5.1.6 A **spanning tree** T of an undirected graph G is a subgraph that is a tree which includes all of the vertices of G.

Fact 5.1.7 A graph G is connected if and only if it has a spanning tree.

Proof The "if" direction is clear, since a spanning tree is connected. For the "only if" direction, suppose G is connected. By repeatedly deleting edges of G until condition (2) of Proposition 5.1.5 is satisfied, we obtain a spanning tree. □

Definition 5.1.8 Let G be a connected graph with labeled vertices. We denote by $t(G)$ the number of labeled spanning trees of G.

Example 5.1.9 (1) If G is itself a tree, then $t(G) = 1$.

(2) When G is the cycle graph C_n with n vertices, $t(G) = n$.

Theorem 5.1.10 (Cayley's formula) For every integer $n \geqslant 2$, we have $t(K_n) = n^{n-2}$.

The quantity $t(K_n)$ is precisely the number of labeled trees on n vertices. In the following three sections, we shall present three different proofs of this formula.

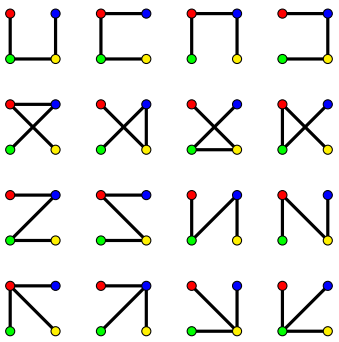


Figure 5.1: Labeled trees on 4 vertices

5.2 Proof 1 of Cayley's Formula

Let T be a spanning tree of K_n with vertex set $\{v_1, \ldots, v_n\}$. By Lemma 2.2.5 and Fact 5.1.3,

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

Lemma 5.2.1 Let $d_1, \ldots, d_n$ be positive integers such that $d_1 + \cdots + d_n = 2n-2$. The number of spanning trees in a complete graph K_n with a degree d_i specified for each vertex v_i is equal to the multinomial coefficient

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

Proof We prove by induction on n. In the base case where $n = 2$, we have $d_1 = d_2 = 1$, and there is exactly one spanning tree. The formula in the lemma also gives 1, so the base case holds.

For the induction step, suppose the lemma is true for all complete graphs with fewer than n vertices. Consider a complete graph on n vertices with a prescribed degree sequence $(d_1, \ldots, d_n)$. By Lemma 5.1.4, there exists a vertex, say v_n, with degree one.

Let v_m be the unique neighbor of v_n. If we remove v_n and the edge $v_m v_n$, then the remaining $n-1$ vertices form a spanning tree of K_{n-1}. In this smaller tree, the degree of v_m is reduced from d_m to $d_m - 1$, while the degrees of all other vertices remain the same. Thus the new degree sequence is

$$(d_1, \ldots, d_{m-1}, d_m - 1, d_{m+1}, \ldots, d_{n-1}).$$

By the induction hypothesis, the number of such trees equals

$$\frac{(n-3)!(d_m - 1)}{(d_1-1)!\cdots(d_{n-1}-1)!}.$$

Since v_m could be any of the $n-1$ vertices other than v_n, the total number of spanning trees with the original degree sequence is the sum of these expressions over all possible $m \in [n-1]$:

$$\begin{aligned}\frac{(n-3)!}{(d_1-1)!\cdots(d_{n-1}-1)!} \sum_{m=1}^{n-1} (d_m - 1) &= \frac{(n-2)!}{(d_1-1)!\cdots(d_{n-1}-1)!} \\ &= \frac{(n-2)!}{(d_1-1)!(d_2-1)!\cdots(d_n-1)!}.\end{aligned}$$

□

Proof 1 of Cayley's formula 5.1.10 By Lemma 5.2.1, the number of labeled spanning trees of K_n is

$$t(K_n) = \sum_{\substack{d_1+\cdots+d_n=2n-2 \\ d_i \geqslant 1}} \binom{n-2}{d_1 - 1, d_2 - 1, \ldots, d_n - 1}.$$

Setting $e_i = d_i - 1$ for each i transforms the condition into $e_1 + \cdots + e_n = n - 2$ with $e_i \geqslant 0$, so

$$t(K_n) = \sum_{\substack{e_1+\cdots+e_n=n-2 \\ e_i \geqslant 0}} \binom{n-2}{e_1, e_2, \ldots, e_n} = n^{n-2}$$

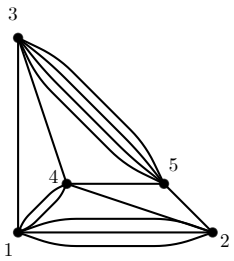
by Definition–Theorem 1.3.8. □

5.3 Proof 2 of Cayley's Formula

Definition 5.3.1 A **multigraph** is a graph which is permitted to have multiple edges (but no loops).

Definition 5.3.2 Let G be a multigraph on the vertex set $[n]$. The **Laplacian matrix** of G is the $n \times n$ matrix $Q = (q_{i,j})$ defined by

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



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

Figure 5.2: A multigraph and its Laplacian matrix

Remark 5.3.3 Let Q denote the Laplacian matrix of a multigraph G. Then:

(1) Q is symmetric.

(2) Every row sum and column sum of Q is zero.

(3) The smallest eigenvalue of Q is 0, because the vector $v_0 = (1, 1, \ldots, 1)$ satisfies $Qv_0 = 0$. In particular, $\det Q = 0$.

Lemma 5.3.4 Let $Q = (q_{i,j})$ be the Laplacian matrix of a multigraph G with vertex set $[n]$. For $i, j \in [n]$, let $Q_{i,j}$ denote the matrix obtained by deleting the i-th row and j-th column of Q, and let $C_{i,j} = (-1)^{i+j} \det Q_{i,j}$ denote the corresponding cofactor. Then all cofactors $C_{i,j}$ are equal.

Proof By Remark 5.3.3 (2), we can transform any minor into any other minor by adding rows and columns, switching them, and multiplying a row or a column by –1. Thus the cofactors are the same up to sign, and it can be verified that, in fact, they have the same sign. □

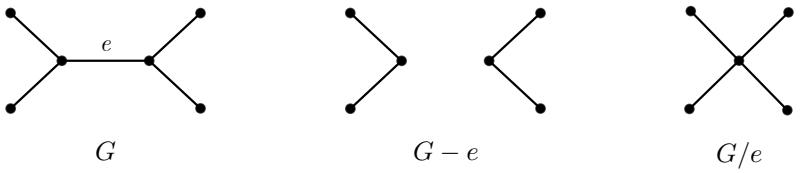
Theorem 5.3.5 (Kirchhoff's matrix tree theorem) Let G be a multigraph with vertex set $[n]$ and Laplacian matrix Q. Then $t(G)$ is equal to any cofactor of Q. That is,

$$t(G) = |\det Q_{i,j}|, \quad \forall i, j \in [n].$$

Proof By Lemma 5.3.4, it suffices to show that $t(G) = \det Q_{1,1}$.

◇ If G has an isolated vertex, say vertex 1, then clearly $t(G) = 0$. In this case, the first row and the first column of Q consist entirely of zeros, so the submatrix obtained by deleting them is itself a Laplacian matrix. By Remark 5.3.3 (3), $\det Q_{1,1} = 0$, as required.

◇ If G has no isolated vertex, we proceed by induction on the number of edges. When $e(G) = 1$, the graph G consists of a single edge between two vertices, and hence $t(G) = 1$. This establishes the base case. Now assume the claim holds for all multigraphs with fewer than $e(G)$ edges. Choose an edge e of G with endpoints i and j.



As illustrated above, there are two natural ways to modify G: either delete e, or contract it by merging i and j into a single vertex. We denote these graphs by $G - e$ and G/e, respectively. We then claim that $t(G)$ satisfies the recursive relation

$$t(G) = t(G - e) + t(G/e). \tag{5–1}$$

In fact, each spanning tree of G either contains the edge e or does not. Those that do not are spanning trees of $G - e$, while those that do correspond uniquely to spanning trees of G/e.

Let Q' and Q'' denote the Laplacian matrices of $G - e$ and G/e, respectively. We are free to reorder the vertices so that i and j are the first two. For the example in Figure 5.2, 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 $Q_{11,22}$ denote the matrix obtained from Q by deleting its first two rows and first two columns. By observing the above example, it is not hard to see that

$$\det Q_{1,1} = \det(Q')_{1,1} + \det Q_{11,22},$$
$$Q_{11,22} = (Q'')_{1,1}.$$

Hence,

$$\det Q_{1,1} = \det(Q')_{1,1} + \det(Q'')_{1,1}. \tag{5–2}$$

By the induction hypothesis,

$$\det(Q')_{1,1} = t(G - e), \quad \det(Q'')_{1,1} = t(G/e). \tag{5–3}$$

Combining (5–1), (5–2), and (5–3), we conclude that

$$t(G) = \det Q_{1,1}.$$

□

Proof 2 of Cayley's formula 5.1.10 The Laplacian matrix of K_n is the $n \times n$ matrix

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

Then

$$Q_{1,1} = nI_{n-1} - J_{n-1},$$

where I_{n-1} is the identity matrix and J_{n-1} is the all-ones matrix. Note that

$$J_{n-1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}.$$

By the matrix determinant lemma, we then have

$$t(K_n) = \det Q_{1,1} = n^{(n-1)-1}[n-(n-1)] = n^{n-2}. \qquad \square$$

5.4 Proof 3 of Cayley's Formula

The classical and most direct method is to find a bijection from the set of all trees on n vertices to another set whose cardinality is known to be n^{n-2}. Naturally, the set of all ordered sequences $(a_1, \ldots, a_{n-2})$ with $a_i \in [n]$ comes into mind. To this end, one seeks a way to uniquely encode each tree T as a sequence $(a_1, \ldots, a_{n-2})$. This idea—using what are now called Prüfer sequences—was introduced by Heinz Prüfer in 1918 to establish Cayley's formula.

Here we want to discuss another bijection proof, due to Joyal, which is less known but of equal elegance and simplicity. For this, we consider not just trees T on $[n]$ but trees together with two distinguished vertices, the *left end* ○ and the *right end* □, which may coincide. Let $\mathcal{T}_n = \{(T, ○, □)\}$ be this new set; then, clearly, $|\mathcal{T}_n| = n^2 t(K_n)$.

Our goal is thus to prove $|\mathcal{T}_n| = n^n$. Now there is a set whose size is known to be n^n, namely the set $[n]^{[n]}$ of all mappings from $[n]$ to $[n]$. Thus Theorem 5.1.10 is proved if we can find a bijection from $[n]^{[n]}$ to $\mathcal{T}_n$.

Let $f\colon [n] \to [n]$ be any map. We present f as a directed graph G_f by drawing arrows from i to $f(i)$. For example, the map

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 7 & 5 & 5 & 9 & 1 & 2 & 5 & 8 & 4 & 7 \end{pmatrix}$$

is represented by the directed graph in Figure 5.3.

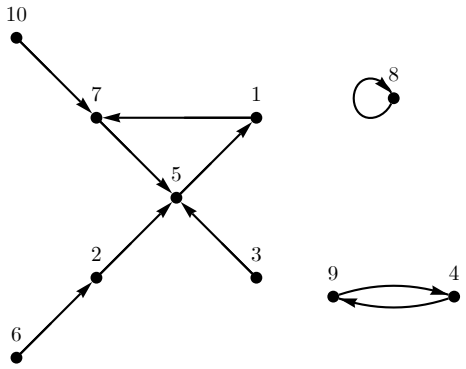


Figure 5.3: The directed graph G_f representing the map f

Look at a component of G_f. Since there is precisely one edge emanating from each vertex, the component contains equally many vertices and edges, and hence precisely one directed cycle. Let $M \subset [n]$ be the union of the vertex sets of these cycles. A moment's thought shows that M is the unique maximal subset of $[n]$ on which f restricts to a bijection. Write

$$f|_M = \begin{pmatrix} a & b & \cdots & z \\ f(a) & f(b) & \cdots & f(z) \end{pmatrix}$$

such that the numbers $a, b, \ldots, z$ in the first row appear in natural order. This gives us an ordering $f(a), f(b), \ldots, f(z)$ of M according to the second row. Now $f(a)$ is our left end and $f(z)$ is our right end.

The tree T corresponding to the map f is now constructed as follows: Draw $f(a), \ldots, f(z)$ in this order as a path from $f(a)$ to $f(z)$, and fill in the remaining vertices as in G_f (deleting the arrows).

In our example above we obtain $M = \{1, 4, 5, 7, 8, 9\}$, and

$$f|_M = \begin{pmatrix} 1 & 4 & 5 & 7 & 8 & 9 \\ 7 & 9 & 1 & 5 & 8 & 4 \end{pmatrix}.$$

Thus the tree T is depicted in Figure 5.4.

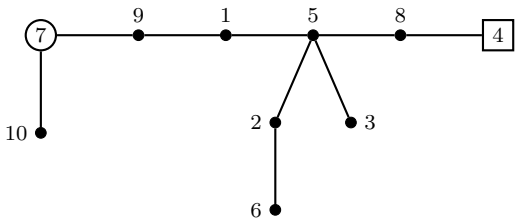


Figure 5.4: The tree T corresponding to the map f

It is immediate how to reverse this correspondence: Given a tree T, we look at the unique path P from the left end to the right end. This gives us the set M and the mapping $f|_M$. The remaining correspondences $i \mapsto f(i)$ are then filled in according to the unique paths from i to P.

Chapter 6 Systems of Distinct Representatives

6.1 Hall's Marriage Theorem

Definition 6.1.1 A **system of distinct representatives (SDR)** for a sequence of (not necessarily distinct) sets $S_1, S_2, \ldots, S_m$ is a sequence of distinct elements $x_1, x_2, \ldots, x_m$ such that $x_i \in S_i$ for all $i \in [m]$.

When does such a system exist? This problem is called the "marriage problem" because an easy reformulation of it asks whether we can marry each of m girls to a boy she knows; boys are the elements and S_i is the set of boys known to the i-th girl.

Clearly, if the sets $S_1, S_2, \ldots, S_m$ have a system of distinct representatives then the following **Hall's condition** is fulfilled: for every $k \in [m]$ the union of any k sets has at least k elements, i.e.,

$$\left|\bigcup_{i\in I} S_i\right| \geqslant |I| \quad \text{for all } I \subset [m]. \tag{6–1}$$

Surprisingly, this obvious necessary condition is also sufficient.

Theorem 6.1.2 (Hall's marriage theorem) The sets $S_1, S_2, \ldots, S_m$ have a system of distinct representatives if and only if (6–1) holds.

Proof 1 We prove the sufficiency of Hall's condition (6–1) by induction on m. The case $m = 1$ is clear. Assume that the claim holds for any collection with less than m sets.

Case 1: For each k, $1 \leqslant k < m$, the union of any k sets contains more than k elements. Take any of the sets, and choose any of its elements x as its representative, and remove x from all the other sets. The union of any $s \leqslant m-1$ of the remaining $m-1$ sets has at least s elements, and therefore the remaining sets have a system of distinct representatives, which together with x give a system of distinct representatives for the original family.

Case 2: The union of some k, $1 \leqslant k < m$, sets contains exactly k elements. By the induction hypothesis, these k sets have a system of distinct representatives. Remove these k elements from the re-

maining $m-k$ sets. Take any s of these sets. Their union contains at least s elements, since otherwise the union of these sets and the k sets would have less than $s+k$ elements. Consequently, the remaining $m-k$ sets also have a system of distinct representatives by the induction hypothesis. Together these two systems of distinct representatives give a system of distinct representatives for the original family. □

In general, Hall's condition (6–1) is hard to verify: we must check if the union of *any* k, $1 \leqslant k \leqslant m$, of the sets $S_1, \ldots, S_m$ contains at least k elements. But if we know more about these sets, then (sometimes) the situation is much better. Here is an example.

Corollary 6.1.3 Let $S_1, \ldots, S_m$ be r-element subsets of an n-element set such that each element belongs to the same number d of these sets. If $m \leqslant n$, then the sets $S_1, \ldots, S_m$ have a system of distinct representatives.

Proof By the double counting argument, $mr = nd$, and hence, $m \leqslant n$ implies that $d \leqslant r$. Now suppose that $S_1, \ldots, S_m$ do not have a system of distinct representatives. By Theorem 6.1.2, the union $Y = S_{i_1} \cup \cdots \cup S_{i_k}$ of some k ($1 \leqslant k \leqslant m$) sets contains strictly less than k elements. For $x \in Y$, let d_x be the number of these sets containing x. Then, we obtain

$$rk = \sum_{j=1}^{k} |S_{i_j}| = \sum_{x \in Y} d_x \leqslant d|Y| < dk,$$

a contradiction with $d \leqslant r$. □

6.2 Latin Rectangles

Definition 6.2.1 A **Latin rectangle** is an $r \times n$ matrix (where $r \leqslant n$) with entries in $[n]$ such that each of the numbers $1, 2, \ldots, n$ occurs once in each row and at most once in each column. An $n \times n$ Latin rectangle is called a **Latin square**.

In 1960, Trevor Evans raised the following question: if fewer than n entries in an $n \times n$ matrix are filled, can one then always complete it to obtain a Latin square? The assertion that a completion is always possible became known as the **Evans conjecture**, and was proved by Smetaniuk (1981) using a quite subtle induction argument.

1	2	3	?
			4

Figure 6.1: A partial Latin square showing that Evans' condition is sharp

On the other hand, it was long known that if a partial Latin square has no partially filled rows (that is, each row is either completely filled or completely free) then it can always be completed. That is, we can build Latin squares by adding rows one-by-one. And this can be easily derived from Hall's theorem.

Theorem 6.2.2 (Ryser's theorem) 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 those integers in $[n]$ which do not occur in the j-th column of R. It is sufficient to prove that the sets $S_1, \ldots, S_n$ have a system of distinct representatives. But this follows immediately from Corollary 6.1.3, because: every set S_j has precisely $n-r$ elements, and each element belongs to precisely $n-r$ sets S_j (since it appears in precisely r columns of the rectangle R). □

6.3 Decomposition of Doubly Stochastic Matrices

Definition 6.3.1 An $n \times n$ matrix $A = (a_{i,j})$ with real non-negative entries $a_{i,j} \geqslant 0$ is called **doubly stochastic** if the sum of entries along any row and any column equals 1. A **permutation matrix** is a doubly stochastic matrix with entries 0 and 1; such a matrix has exactly one 1 in each row and in each column.

Remark 6.3.2 Doubly stochastic matrices arise in the theory of Markov chains: $a_{i,j}$ is the transition probability from the state i to the state j.

Theorem 6.3.3 (Birkhoff–von Neumann theorem) Every doubly stochastic matrix A is a convex combination of permutation matrices. That is, there exist non-negative reals $\lambda_1, \ldots, \lambda_s$ and permutation matrices $P_1, \ldots, P_s$ such that

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

Proof We will prove a more general result that every $n \times n$ non-negative matrix $A = (a_{i,j})$ having all row and column sums equal to some positive value $\gamma > 0$ can be expressed as a linear combination $A = \lambda_1 P_1 + \cdots + \lambda_s P_s$ of permutation matrices $P_1, \ldots, P_s$, where $\lambda_1, \ldots, \lambda_s$ are non-negative reals such that $\lambda_1 + \cdots + \lambda_s = \gamma$.

To prove this, we apply induction on the number of non-zero entries in A. Since $\gamma > 0$, we have at least n such entries. If there are exactly n non-zero entries then $A = \gamma P$ for some permutation matrix P, and we are done. Now suppose that A has more than n non-zero entries and that the result holds for matrices with a smaller number of such entries. Define

$$S_i = \{j : a_{i,j} > 0\} \quad \text{for } i \in [n],$$

and observe that the sets $S_1, \ldots, S_n$ fulfill Hall's condition (6–1). Indeed, if the union of some k ($1 \leqslant k \leqslant n$) of these sets contained less than k elements, then all the non-zero entries of the corresponding k rows of A would occupy no more than k–1 columns; hence, the sum of these entries by columns would be at most $(k-1)\gamma$, whereas the sum by rows is $k\gamma$, a contradiction.

By Theorem 6.1.2, there is a system of distinct representatives $j_1 \in S_1, \ldots, j_n \in S_n$. Take the permutation matrix $P_1 = (p_{i,j})$ with entries $p_{i,j} = 1$ if and only if $j = j_i$. Let $\lambda_1 = \min\{a_{1,j_1}, \ldots, a_{n,j_n}\}$, and consider the matrix $A_1 = A - \lambda_1 P_1$. By the definition of the sets S_i, $\lambda_1 > 0$. So, this new matrix A_1 has less non-zero entries than A. Moreover, the matrix A_1 satisfies the condition of the theorem with $\gamma_1 = \gamma - \lambda_1$. We can therefore apply the induction hypothesis to A_1, which yields a decomposition $A_1 = \lambda_2 P_2 + \cdots + \lambda_s P_s$, and hence, $A = \lambda_1 P_1 + A_1 = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_s P_s$, as desired. □

6.4 Matching in Bipartite Graphs

Definition 6.4.1 Let G be a graph. Two edges of G are said to be **disjoint** if they share no common vertex. A **matching** in G is a set of pairwise disjoint edges, and a **perfect matching** is a matching that covers every vertex of G.

Definition 6.4.2 Let G be a bipartite graph with parts A and B. A matching which matches all the vertices from A is called a **matching of A into B**. In this case, we have $|A| \leqslant |B|$.

We may ask whether $G = (A, B, E)$ has a matching of A into B. The answer is given by Hall's marriage theorem. Let S_x be the set of all neighbors of x in G. Observing that there is a matching of A into B if and only if the sets S_x with $x \in A$ have a system of distinct representatives, Theorem 6.1.2 immediately yields the following:

Theorem 6.4.3 If G is a bipartite graph with parts A and B, then G has a matching of A into B if and only if, for every $k = 1, 2, \ldots, |A|$, every subset of k vertices from A has at least k neighbors.

In terms of bipartite graphs, Corollary 6.1.3 can be restated as follows:

Corollary 6.4.4 If $G = (A, B, E)$ is (d, r)-biregular with $|B| \leqslant |A|$, then G has a matching of B into A.

6.5 König's Min–Max Theorem

The early results of Frobenius and König have given rise to a large number of **min–max theorems** in combinatorics, in which the minimum of one quantity equals the maximum of another.

By Hall's theorem, we know whether each of the girls can be married to a boy she knows. If so, all are happy (except for the boys not chosen ...). But what if not? In this sad situation it would be nice to make as many happy marriages as possible. So, given a sequence of sets $S_1, S_2, \ldots, S_m$, we try to find a system of distinct representatives for as many of these sets as possible. In terms of 0-1 matrices this problem is solved by the following result.

Definition 6.5.1 Let A be an $m \times n$ matrix, all whose entries have value 0 or 1. Two 1s are **dependent** if they are on the same row or on the same column; otherwise, they are **independent**. The size of the largest set of independent 1s is also known as the **term rank** of A.

Theorem 6.5.2 (König's theorem) Let A be an $m \times n$ 0-1 matrix. The maximum number of independent 1s is equal to the minimum number of rows and columns required to cover all the 1s in A.

Proof Let r denote the maximum number of independent 1s and R the minimum number of rows and columns required to cover all the 1s. Clearly, $R \geqslant r$, because we can find r independent 1s in A, and any row or column covers at most one of them.

We need to prove that $r \geqslant R$. Assume that some a rows and b columns cover all the 1s and $a+b = R$. Because permuting the rows and columns changes neither r nor R, we may assume that the first a rows and the first b columns cover the 1s. Write A in the form

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

We know that there are no 1s in E. We will show that there are a independent 1s in C. The same argument shows, by symmetry, that there are b independent 1s in D. Since altogether these $a+b$ 1s are independent, this shows that $r \geqslant a+b = R$, as desired.

Define

$$S_i = \{j : c_{i,j} = 1\} \subset [n-b],$$

as the set of locations of the 1s in the i-th row of $C = (c_{i,j})$. We claim that the sequence $S_1, S_2, \ldots, S_a$ has a system of distinct representatives, i.e., we can choose a 1 from each row, no two in the same column. Otherwise, Theorem 6.1.2 tells us that the 1s in some k ($1 \leqslant k \leqslant a$) of these rows can all be covered by less than k columns. But then we obtain a covering of all the 1s in A with fewer than $a+b$ rows and columns, a contradiction. □

In terms of bipartite graphs, König's theorem is as follows.

Definition 6.5.3 A **vertex cover** of a graph is a set of vertices that includes at least one endpoint of every edge of the graph.

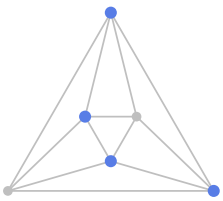


Figure 6.2: A vertex cover highlighted in the graph

Theorem 6.5.4 The maximum size of a matching in a bipartite graph equals the minimum size of a vertex cover.

Proof Let $G = (A, B, E)$ be a bipartite graph and $M = (m_{a,b})$ the $|A| \times |B|$ matrix defined by $m_{a,b} = 1$ if and only if $ab \in E$. There is a one-to-one correspondence between matchings in G and sets of independent 1s in M, and between vertex covers in G and sets of rows and columns covering all the 1s in M. The result now follows from Theorem 6.5.2. □

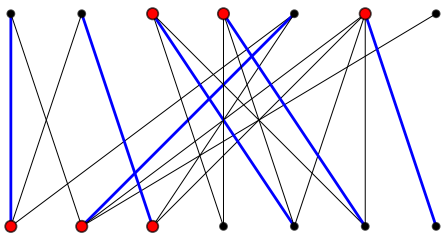


Figure 6.3: Bipartite graph with a maximum matching and minimum vertex cover both of size six

Chapter 7 Designs and Finite Geometry

7.1 Block Designs

Definition 7.1.1 A (v, k, λ)**-design**, where $v > k \geqslant 2$ and $\lambda \geqslant 1$, consists of a set X of v **points** (also called **varieties**) together with a collection $\mathcal{D}$ of distinct subsets of X (called **blocks**) such that:

- ⋄ each block in $\mathcal{D}$ contains exactly k points (the design is therefore **uniform** or **proper**);
- ⋄ every pair of distinct points in X is contained in exactly λ blocks (the design is **pairwise balanced**).

The number of blocks is usually denoted by b. A design in which $b = v$ (that is, the number of blocks equals the number of points) is often called **symmetric**.

Remark 7.1.2 Designs are usually said (or assumed) to be **incomplete**, meaning that the collection of blocks is not all possible k-subsets, thus ruling out a trivial design. In this case, a (v, k, λ)-design is referred to as a **balanced incomplete block design (BIBD)**.

Example 7.1.3 The complete graph K_n is an $(n, 2, 1)$-design: the vertices are the points, and the edges (2-element subsets) are the blocks.

Example 7.1.4 The **Fano plane** is a symmetric $(7, 3, 1)$-design. The points of the design are the points of the plane, and the blocks of the design are the lines $\{123, 145, 167, 246, 257, 347, 356\}$ of the plane.

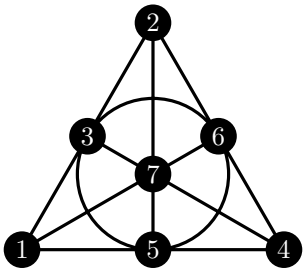


Figure 7.1: The Fano plane with 7 lines, each containing 3 points

Definition 7.1.5 A design is **r-regular** if each point occurs in exactly r blocks. Here r is called the **replication number** of the design.

Theorem 7.1.6 Let $\mathcal{D}$ be a (v,k,λ)-design containing b blocks. Then $\mathcal{D}$ is r-regular with the replication number r satisfying the equations

$$r(k-1) = \lambda(v-1)$$

and

$$bk = vr.$$

Proof Let $a \in X$ be fixed and assume that a occurs in r_a blocks. We count in two ways the number of pairs (x, B) where B is a block containing a, and x is a point in B other than a.

- ◇ For each of the $v-1$ possibilities for x there are exactly λ blocks B containing both a and x, giving a total of $\lambda(v-1)$ pairs.
- ◇ For each of the r_a blocks B containing a, the elements of B other than a can be chosen in $|B|-1 = k-1$ ways, giving a total of $r_a(k-1)$ pairs.

Equating these two expressions gives $r_a(k-1) = \lambda(v-1)$, which shows that r_a is independent of the choice of a and proves the first equation.

To prove the second claim, we count in two ways the number of pairs (x, B) where x is a point and B is a block containing x.

- ◇ For each $x \in X$ the block B can be chosen in r ways.
- ◇ For each of the b blocks B the element $x \in B$ can be chosen in k ways.

Hence $vr = bk$, as desired. □

Corollary 7.1.7 If a (v,k,λ)-design exists, then

$$\lambda(v-1) \equiv 0 \pmod{k-1} \quad \text{and} \quad \lambda v(v-1) \equiv 0 \pmod{k(k-1)}.$$

Proof By Theorem 7.1.6, the replication number r is an integer satisfying

$$r = \frac{\lambda(v-1)}{k-1} = \frac{\lambda v(v-1)}{k(k-1)}.$$

□

Remark 7.1.8 In the early 1970s, Richard M. Wilson showed that these trivial necessary conditions for the existence of a (v,k,λ)-design are *asymptotically* sufficient: for fixed k and λ, there exists a constant $v_0(k,\lambda)$ such that a (v,k,λ)-design exists for all $v > v_0(k,\lambda)$ satisfying the above divisibility conditions.

Theorem 7.1.9 (Fisher's inequality) Let $\mathcal{D}$ be a (v,k,λ)-design. Then $b \geqslant v$.

Proof Let the incidence matrix M be the $v \times b$ matrix whose (i,j)-entry is 1 if the i-th point is contained in the j-th block, and 0 otherwise. Note that

- ◇ Each row of M contains exactly r ones.
- ◇ Any two distinct rows of M have ones in exactly λ common columns.

Thus $B \coloneqq MM^{\mathsf{T}}$ is a $v \times v$ matrix whose (i,j)-entry is r if $i = j$ and λ if $i \neq j$. Since $v > k$, by Theorem 7.1.6, we have $r > \lambda$. By the matrix determinant lemma, we then have

$$\begin{aligned}\det B &= \det\left((r-\lambda)I_v + \begin{pmatrix}\lambda & \cdots & \lambda\end{pmatrix}^{\mathsf{T}}\begin{pmatrix}1 & \cdots & 1\end{pmatrix}\right)\\ &= \left\{1 + \begin{pmatrix}1 & \cdots & 1\end{pmatrix}\frac{1}{r-\lambda}I_v\begin{pmatrix}\lambda & \cdots & \lambda\end{pmatrix}^{\mathsf{T}}\right\}(r-\lambda)^v\\ &= (r-\lambda)^{v-1}[r + (v-1)\lambda] > 0.\end{aligned}$$

This implies that $b \geqslant \operatorname{rank} M \geqslant \operatorname{rank} B = v$. □

7.2 Finite Linear Spaces

Definition 7.2.1 A **finite linear space** over a finite set X is a family $\mathcal{L}$ of subsets of X, called **lines**, such that:

- ◇ Every line contains at least two points.
- ◇ Any two points are on exactly one line.

Theorem 7.2.2 (de Bruijn–Erdős theorem) If $\mathcal{L}$ is a finite linear space over X with $|\mathcal{L}| \geqslant 2$, then $|\mathcal{L}| \geqslant |X|$, with equality if and only if any two lines share exactly one point.

Proof Let $b = |\mathcal{L}| \geqslant 2$ and $v = |X|$. For a point $x \in X$, let r_x be the number of lines in $\mathcal{L}$ containing x. If $x \notin L$ then $r_x \geqslant |L|$ because there are $|L|$ lines joining x to the points on L.

Suppose $b \leqslant v$. So, for $x \notin L$, we have

$$b(v - |L|) \geqslant v(b - r_x). \tag{7–1}$$

Hence

$$\begin{aligned}b &= \sum_{L\in\mathcal{L}} 1 = \sum_{L\in\mathcal{L}}\sum_{x:x\notin L}\frac{1}{v-|L|} \leqslant \frac{b}{v}\sum_{L\in\mathcal{L}}\sum_{x:x\notin L}\frac{1}{b-r_x}\\ &= \frac{b}{v}\sum_{x\in X}\sum_{L:x\notin L}\frac{1}{b-r_x} = \frac{b}{v}\sum_{x\in X}1 = b,\end{aligned}$$

and this implies that all inequalities are equalities so that $b = v$, and $r_x = |L|$ whenever $x \notin L$.

Therefore, we must always have $b \geqslant v$. Now consider the case of equality $b = v$:

(⇒) No two lines can be disjoint: if $L_1 \cap L_2 = \varnothing$ and $x \in L_1$, then $x \notin L_2$ forces $r_x = |L_2|$. But every point of L_2 determines a distinct line through x, giving strictly more than r_x lines, a contradiction.

(⇐) If instead $b > v$, then the inequality in (7–1) is reversed since $r_x = |L|$ for $x \notin L$, and the same double counting argument yields $b > b$, impossible. □

Remark 7.2.3 Theorem 7.2.2 generalizes Theorem 7.1.9 in the case $\lambda = 1$. In particular, any two blocks of a symmetric $(v,k,1)$-design share exactly one point.

7.3 Difference Sets

In this section we will construct symmetric designs using difference sets. Throughout this section, let G be an additive abelian group of order v.

Definition 7.3.1 Let $2 \leqslant k < v$ and $\lambda \geqslant 1$. A (v, k, λ)**-difference set** in G is a k-element subset $D = \{d_1, d_2, \ldots, d_k\} \subset G$ such that the collection of values $d_i - d_j$ $(i \neq j)$ contains every element in $G \setminus \{0\}$ exactly λ times. In the case $G = \mathbb{Z}_v$, we call D a **cyclic difference set**.

Example 7.3.2 (1) $D = \{0, 1, 3\}$ is a cyclic $(7, 3, 1)$-difference set in $\mathbb{Z}_7$.

(2) $D = \{0, 1, 3, 9\}$ is a cyclic $(13, 4, 1)$-difference set in $\mathbb{Z}_{13}$.

(3) $D = \{(0,0,0,0), (1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1), (1,1,1,1)\}$ is a $(16, 6, 2)$-difference set in $\mathbb{Z}_2^4$, which is a non-cyclic example.

Since the number of pairs (i, j) with $i \neq j$ equals $k(k-1)$ and these give each of the $v - 1$ nonzero elements λ times as a difference, it follows that

$$\lambda(v - 1) = k(k - 1). \tag{7–2}$$

If D is a difference set, we call the set

$$a + D \coloneqq \{a + d_1, a + d_2, \ldots, a + d_k\}$$

a **translate** of D. Notice that our assumption $k < \lambda$ together with (7–2) implies that all the translates of a difference set are different. Indeed, if $a + D = D$ for some $a \in G \setminus \{0\}$, then there is a permutation π of $[k]$ so that $\pi(i) \neq i$ and $a + d_i = d_{\pi(i)}$ for all i. Hence, a can be expresses as a difference $d_{\pi(i)} - d_i$ in k ways; but $\lambda < k$ by (7–2) and our assumption that $k < \lambda$.

Theorem 7.3.3 If $D = \{d_1, d_2, \ldots, d_k\}$ is a (v, k, λ)-difference set then the translates

$$D, 1 + D, \ldots, (v - 1) + D$$

are the blocks of a symmetric (v, k, λ)-design.

Proof We have v blocks over v points. Since, clearly, every one of the translates contains k points, it remains to show that every pair of distinct points is contained in exactly λ blocks.

Let $x, y \in G$ with $x \neq y$. Then $x - y$ can be expressed as a difference $d_i - d_j$ in exactly λ ways. For each such pair (i, j), there is exactly one a for which $x, y \in a + D$, namely, $a = x - d_i = y - d_j$. □

Let us now describe one construction of difference sets. Denote by $\mathbb{F}_v$ the finite field of order v, where v is a prime power. **Squares** (or **quadratic residues**) in $\mathbb{F}_v$ are the elements a^2 for $a \in \mathbb{F}_v$.

Theorem 7.3.4 If v is a prime power and $v \equiv 3 \pmod 4$, then the nonzero squares in $\mathbb{F}_v$ form a (v, k, λ)-difference set in $(\mathbb{F}_v, +)$ with $k = (v - 1)/2$ and $\lambda = (v - 3)/4$.

Proof Let α be a primitive element of $\mathbb{F}_v$. Then

$$\mathbb{F}_v = \{0, \alpha^0, \alpha^1, \ldots, \alpha^{v-2}\}.$$

Let D denote the set of nonzero squares in $\mathbb{F}_v$:

$$D = \{\alpha^0, \alpha^2, \ldots, \alpha^{v-3}\}.$$

Thus, $k = |D| = (v-1)/2$. Since $v \equiv 3 \pmod 4$, the integer $(v-1)/2$ is odd. In particular, $-1 = \alpha^{\frac{v-1}{2}}$ cannot be a square in $\mathbb{F}_v$. Hence the set of nonsquares is exactly

$$-D = \{-d : d \in D\},$$

and we have

$$\mathbb{F}_v \setminus \{0\} = D \cup (-D).$$

For any $d \in D$, the pair $(x, y) \in D \times D$ satisfies the equation $x - y = 1$ if and only if the pair $(dx, dy) \in D \times D$ satisfies the equation $dx - dy = d$, or equivalently, if and only if the pair $(dy, dx) \in D \times D$ satisfies the equation $dy - dx = -d$. This shows that all nonzero squares $d \in D$ and all nonsquares $-d \in -D$ have the same number λ of representations as a difference of two nonzero squares. We can compute λ from (7–2), which gives

$$\lambda = \frac{k(k-1)}{v-1} = \frac{v-3}{4}.$$

□

7.4 Projective Planes

Let $\mathcal{L} \subset 2^X$ be a finite linear space with $|\mathcal{L}| = b$ and $|X| = v$. By Theorem 7.2.2, $b \geqslant v$. In this section we will consider finite linear spaces with $b = v$ and with an additional requirement that every line has the same number, say $q+1$, of points. Then $\mathcal{L}$ turns into a symmetric (v, k, λ)-design with $\lambda = 1$ and $k = q+1$. Such a design is known as a **projective plane of order** q and is denoted by $\mathrm{PG}(q)$, where PG stands for projective geometry. By Theorem 7.1.6, we have

$$v = b = \frac{k(k-1)}{\lambda} + 1 = q^2 + q + 1.$$

Formally, we make the following definition.

Definition 7.4.1 A **projective plane of order** q consists of a set X of q^2+q+1 elements called **points**, and a family $\mathcal{L}$ of X called **lines**, having the following properties:

- ◇ Every line has $q+1$ points.
- ◇ Every two points lie on a unique line.

Example 7.4.2 (1) The only possible projective plane of order $q = 1$ is a triangle.

(2) The Fano plane in Example 7.1.4 is a projective plane of order $q = 2$.

Proposition 7.4.3 A projective plane of order q has the following properties:

(1) Any point lies on exactly $q+1$ lines.

(2) There are exactly q^2+q+1 lines.

(3) Any two distinct lines intersect in exactly one point.

Proof (1) Take a point x. There are $q(q+1)$ points different from x; each line through x contains q further points, and there are no overlaps between these lines (apart from x). So, there must be $q+1$ lines through x.

(2) Counting in two ways the pairs (x, L) with $x \in L$, we obtain

$$|\mathcal{L}|(q+1) = \big(q^2+q+1\big)(q+1),$$

so $|\mathcal{L}| = q^2+q+1$.

(3) This follows from (2) and Theorem 7.2.2. □

Corollary 7.4.4 A projective plane of order q is a symmetric $\big(q^2+q+1, q+1, 1\big)$-design.

A standard way to construct a projective plane of prime power order q is as follows.

Let V be the set of all nonzero vectors $(x_0, x_1, x_2) \in \mathbb{F}_q^3$. The points of our plane are sets

$$[x_0 : x_1 : x_2] = \{(cx_0, cx_1, cx_2) : c \in \mathbb{F}_q \setminus \{0\}\}$$

of $q-1$ vectors in V. There are $\big(q^3-1\big)/(q-1) = q^2+q+1$ such sets, and hence, so many points. The line $L(a_0, a_1, a_2)$, where $(a_0, a_1, a_2) \in V$, is defined to be the set of all those points $[x_0 : x_1 : x_2]$ for which

$$a_0x_0 + a_1x_1 + a_2x_2 = 0. \tag{7–3}$$

Since $(a_0, a_1, a_2) \in V$, this vector has at least one nonzero component, say $a_0 \neq 0$. Therefore, (7–3) has exactly q^2-1 solutions $(x_0, x_1, x_2) \in V$: for arbitrary x_1, x_2, not both zero, this equation uniquely determines x_0. Since each $[x_0 : x_1 : x_2]$ consists of $q-1$ vectors, there are exactly $\big(q^2-1\big)/(q-1) = q+1$ points $[x_0 : x_1 : x_2]$ satisfying (7–3). In other words, there are exactly $q+1$ points on each line. So, it remains to verify that any two points lie on a unique line.

To show this, let $[x_0 : x_1 : x_2]$ and $[y_0 : y_1 : y_2]$ be two distinct points. For each line $L(a_0, a_1, a_2)$ containing both these points,

$$\begin{aligned} a_0x_0 + a_1x_1 + a_2x_2 &= 0, \\ a_0y_0 + a_1y_1 + a_2y_2 &= 0. \end{aligned}$$

Without loss of generality $x_0 \neq 0$. Then $a_0 = -a_1x_1/x_0 - a_2x_2/x_0$, and we can replace the second equation by

$$a_1\left(y_1 - \frac{y_0}{x_0}x_1\right) + a_2\left(y_2 - \frac{y_0}{x_0}x_2\right) = 0. \tag{7–4}$$

If

$$y_1 - \frac{y_0}{x_0}x_1 = y_2 - \frac{y_0}{x_0}x_2 = 0$$

then $(y_0, y_1, y_2) = (cx_0, cx_1, cx_2)$ with $c = y_0/x_0$, and hence, $[y_0 : y_1 : y_2] = [x_0 : x_1 : x_2]$, which is impossible since we consider distinct points. Therefore, at least one of them, say $y_1 - (y_0/x_0)x_1$, is nonzero. Then for arbitrary nonzero a_2, both a_1 and a_0 are uniquely determined by (7–4) and the first equation; and if (a_0, a_1, a_2) is a solution then (ca_0, ca_1, ca_2) for $c \neq 0$ are all the solutions. Consequently, every two different points $[x_0 : x_1 : x_2]$ and $[y_0 : y_1 : y_2]$ are contained in a unique line, as desired.

By the above construction and Corollary 7.4.4, we have the following result.

Corollary 7.4.5 For every prime power q, there exists a symmetric $\big(q^2+q+1, q+1, 1\big)$-design.

There is a famous conjecture that the order of a projective plane is always a prime power. However, experts in the field have tried to prove this conjecture now for several decades, and the status of the conjecture remains open. A priori there is no reason why for $n > 1$ there should not be a projective plane of order n. The sharpest result that rules out several cases is the theorem of Bruck and Ryser, which was first proven in 1949 (which we quote without proof here).

Theorem 7.4.6 (Bruck–Ryser theorem) If a projective plane of order q exists and $q \equiv 1$ or $2 \pmod 4$, then q is the sum of two squares.

Example 7.4.7 (1) Since $6 \equiv 2 \pmod 4$ and 6 is not the sum of two squares, there is no projective plane of order 6.

(2) Since $14 \equiv 2 \pmod 4$ and 14 is not the sum of two squares, there is no projective plane of order 14.

7.5 Resolvable Designs

Suppose $\mathcal{D}$ is a (v, k, λ)-design over a set X. A **parallel class** in $\mathcal{D}$ is a subset of disjoint blocks from $\mathcal{D}$ whose union is X. Observe that a parallel class contains $b = v/k$ blocks, and every point of X appears in exactly one of these blocks. Moreover, by Theorem 7.1.6, we have

$$r = \frac{bk}{v} = \frac{\lambda(v-1)}{k-1}$$

such classes, where r is the replication number of $\mathcal{D}$. A partition of $\mathcal{D}$ into r parallel classes is called a **resolution**, and a design admitting such a resolution is called **resolvable**.

Example 7.5.1 Let us consider the following example from Anderson and Honkala (1997). We have a football league of $2n$ teams and each team plays exactly once against every other team. We wish to arrange the league schedule so that all the matches are played during $2n-1$ days, and on each day every team plays one match.

What we are looking for is a resolvable $(2n, 2, 1)$-design. For convenience, let our ground set (of teams) be $X = \{*, 1, \ldots, 2n-1\}$, where $*$ is some symbol different from $1, \ldots, 2n-1$. Since by Theorem 7.1.6, the replication number equals

$$r = \frac{\lambda(v-1)}{k-1} = \frac{1 \cdot (2n-1)}{2-1} = 2n-1,$$

we have to show how to partite the collection $\mathcal{D}$ of all 2-element subsets of X into $2n-1$ parallel classes $\mathcal{D}_1, \ldots, \mathcal{D}_{2n-1}$; the i-th class $\mathcal{D}_i$ gives us the set of matches played at the i-th day.

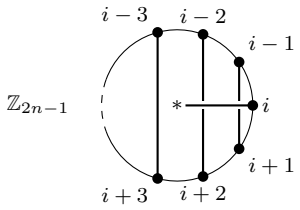


Figure 7.2: Blocks in the parallel class $\mathcal{D}_i$

Define $\{i, *\} \in \mathcal{D}_i$ for all $i \in X \setminus \{*\}$, and $\{a, b\} \in \mathcal{D}_i$, if

$$a + b \equiv 2i \pmod{2n-1}$$

for $a, b \in X \setminus \{*\}$, as illustrated in Figure 7.2. Since $2n-1$ is odd, each 2-element subset of X belongs to a unique $\mathcal{D}_i$; and the unique block in $\mathcal{D}_i$ containing an element $a \in X$ is $\{a, b\}$ where $b \equiv 2i - a \pmod{2n-1}$ if $a \neq i$, and $\{i, *\}$ if $a = i$.

Remark 7.5.2 A resolvable $(2n, 2, 1)$-design is exactly a partition of the edges of K_{2n} into $2n-1$ perfect matchings.

Example 7.5.3 (Kirkman's schoolgirl problem) Fifteen young ladies in a school walk out three abreast for seven days in succession: it is required to arrange them daily so that no two shall walk twice abreast.

This is equivalent to finding a resolvable $(15, 3, 1)$-design. One solution is given by the table below, denoting the 15 girls with the letters A to O.

Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
ABC	ADG	AEO	AIM	AFJ	AHK	ALN
DEF	BEH	BIJ	BDL	BKO	BGN	BFM
GHI	CJM	CDN	CEK	CGL	CFI	CHO
JKL	FKN	FHL	FGO	DHM	DJO	DIK
MNO	ILO	GKM	HJN	EIN	ELM	EGJ

Example 7.5.4 Let $X = [9]$. Then

$$\mathcal{D} = \{123, 456, 789, 168, 249, 357, 159, 267, 348, 147, 258, 369\}$$

is a resolvable $(9, 3, 1)$-design, as shown in Figure 7.3.

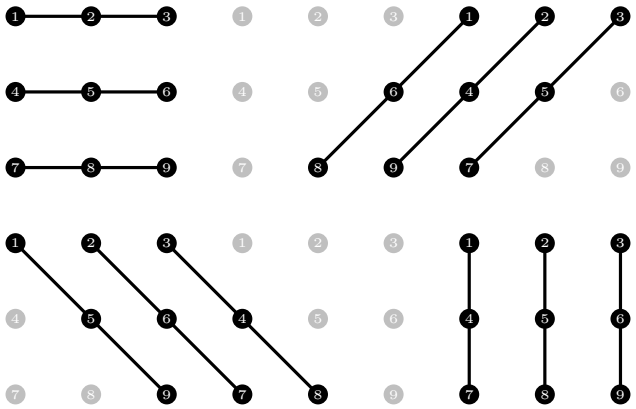


Figure 7.3: A resolvable $(9, 3, 1)$-design

Examples 7.5.3 and 7.5.4 suggest a natural class of combinatorial designs:

Definition 7.5.5 A **Steiner triple system of order** v (abbreviated STS(v)) is a $(v, 3, 1)$-design, whose blocks are called **triples**. A **Kirkman triple system of order** v (abbreviated KTS(v)) is a resolvable STS(v).

By Theorem 7.1.6, the total number of triples in an STS(v) is $v(v-1)/6$, and the replication number is $(v-1)/2$. This implies that v must be odd and that $6 \mid v(v-1)$, which is equivalent to $3 \mid v(v-1)$.

Therefore, we obtain

$$v \equiv 1 \text{ or } 3 \pmod 6. \tag{7–5}$$

The fact that the condition (7–5) on v ($v \geqslant 3$) is sufficient for the existence of an STS(v) was proved by Raj Chandra Bose and T. Skolem.

Since each block in a KTS(v) contains three points, this is possible only when $3 \mid v$. This means that we must have

$$v \equiv 3 \pmod 6. \tag{7–6}$$

The sufficiency of the condition (7–6) on v ($v \geqslant 3$) for the existence of a KTS(v) was presented at a conference in 1968 by Dwijendra Kumar Ray-Chaudhuri and Richard Michael Wilson, and published in 1971 in the proceedings of that conference. Lu Jiaxi had actually written a proof of this result in 1961, but he was unknown at the time and through misunderstandings and mistakes, his work was rejected as "not really new".

If $\mathcal{D}$ is an STS(v) over X, then $\mathcal{D} \subset \binom{X}{3}$ and $|\mathcal{D}| = v(v-1)/6$. Since $\left|\binom{X}{3}\right| = v(v-1)(v-2)/6$, we see that $|\mathcal{D}|$ divides $\left|\binom{X}{3}\right|$. This raises the question: is it possible to partition $\binom{X}{3}$ into a family of $v-2$ disjoint STS(v)'s? Such a partition of $\binom{X}{3}$ is called a **large set of disjoint Steiner triple systems**.

The existence of large sets of disjoint STS(v)'s was established through a series of classic papers by Lu Jiaxi, who proved that they exist for all $v \neq 7$ satisfying (7–5), with six possible exceptions $v = 141$, 283, 501, 789, 1501, and 2365. The spectrum was finally completed by Teirlinck in 1991, who constructed large sets for these remaining six cases.

It is natural to ask for the existence of large sets of disjoint KTS(v)'s. This problem, however, remains largely open. In 2018, Peter Keevash proved that such large sets exist for all sufficiently large values of v satisfying (7–6).

7.6 Affine Planes

An **affine plane** AG(q) of order q is a $\big(q^2, q, 1\big)$-design. By Theorem 7.1.6, each point of this plane belongs to $r = \big(q^2-1\big)/(q-1) = q+1$ lines, and we have $b = vr/k = q^2+q$ lines altogether. Put otherwise, an affine plane of order q has q^2 points and satisfies the following conditions:

- ◇ Every line has q points.
- ◇ Any two points lie on a unique line.
- ◇ Any point lies on exactly $q+1$ lines.
- ◇ There are exactly q^2+q lines.

Hence, the main difference from a projective plane is that now we can have "parallel" lines, i.e., lines which do not meet each other.

There are two basic constructions of affine planes.

Construction 1 An affine plane can be obtained from a projective plane by removing any one of its lines. Let $(X, \mathcal{L})$ be a projective plane of order q. Fix one of its lines $L_0 \in \mathcal{L}$ and consider the design $(X', \mathcal{L}')$ where $X' = X \setminus L_0$ and $\mathcal{L}' = \{L \setminus L_0 : L \in \mathcal{L}, L \neq L_0\}$. Since X' has $q^2 + q + 1 - (q + 1) = q^2$ points, each line in $\mathcal{L}'$ has $(q + 1) - 1 = q$ points, and any two distinct points of X' lie on a unique line of $\mathcal{L}'$, the obtained design $(X', \mathcal{L}')$ is an affine plane of order q. The line L_0 is called the **line at infinity**. For each line $L' \in \mathcal{L}'$ of the affine plane there is a unique point $x \in L_0$ such that $L' \cup \{x\} \in \mathcal{L}$; this point is called the **infinite point** of L'.

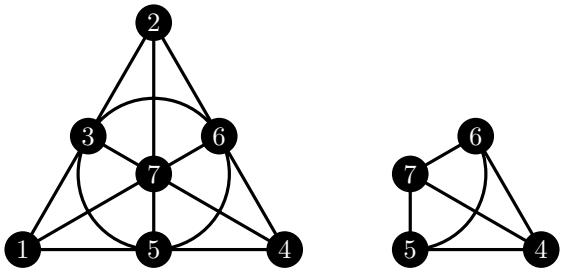


Figure 7.4: Construction 1 applied to the Fano plane: $L_0 = \{1, 2, 3\}$ is the removed line

Construction 2 Let q be a prime power, and consider the set of points $X = \mathbb{F}_q \times \mathbb{F}_q$. Let $\mathcal{D}$ be the set of all blocks of the form

$$L(a, b) \coloneqq \{(x, y) \in X : y = ax + b\}$$

and

$$L(c) \coloneqq \{(c, y) : y \in \mathbb{F}_q\},$$

where $a, b, c \in \mathbb{F}_q$. We will show that $\mathcal{D}$ is a $\left(q^2, q, 1\right)$-design. Clearly, there are q^2 points and each block contains exactly q of them. Hence, we only need to show that every pair of points $(x_1, y_1), (x_2, y_2) \in \mathbb{F}_q \times \mathbb{F}_q$ is contained in a unique block. If $x_1 = x_2$ then the unique block containing this pair is $L(x_1)$. If $x_1 \neq x_2$ then the system of equations $y_1 = ax_1 + b$, $y_2 = ax_2 + b$ has a unique solution (a, b); hence, the unique block containing that pair is $L(a, b)$, and we are done.

Remark 7.6.1 (1) Construction 2 is just Construction 1 applied to the projective plane constructed in Section 7.4.

(2) The affine plane $\mathrm{AG}(q)$ is a resolvable $\left(q^2, q, 1\right)$-design, see Problem **??**.

Chapter 8 Intersecting Families

8.1 Basic Definitions

Definition 8.1.1 An **intersecting family** of sets is a collection of sets such that every two sets intersect.

Example 8.1.2 Here are two intersecting families of size 2^{n-1} in $2^{[n]}$:

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

(2) $\mathcal{F} = \{A \subset [n] : |A| > n/2\}$, where n is odd.

Fact 8.1.3 Every intersecting family $\mathcal{F} \subset 2^{[n]}$ satisfies $|\mathcal{F}| \leqslant 2^{n-1}$.

Proof Partition $2^{[n]}$ into pairs of the form $\{A, A^c\}$, where $A \subset [n]$. There are exactly 2^{n-1} such pairs, and since no two complementary sets intersect, the family $\mathcal{F}$ can contain at most one set from each pair. Hence $|\mathcal{F}| \leq 2^{n-1}$. □

Example 8.1.4 Here are two examples of intersecting families of uniform size:

(1) $\mathcal{F} = \left\{A \in \binom{[n]}{k} : 1 \in A\right\}$, with $|\mathcal{F}| = \binom{n-1}{k-1}$.

(2) $\mathcal{F} = \binom{[n]}{k}$ when $k > n/2$, with $|\mathcal{F}| = \binom{n}{k}$.

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 $n \geqslant 2k$ since otherwise any two k-element sets intersect, and there is nothing to prove.

The following result, found by Erdös, Ko, and Rado in 1938 (but published only 23 years later), answers the question.

Theorem 8.1.5 (Erdös–Ko–Rado theorem) If $2k \leqslant n$ then every intersecting family of k-element subsets of $[n]$ has size at most $\binom{n-1}{k-1}$. Moreover, if $2k < n$, then equality holds if and only if the family is a **star**, i.e., consists of all k-element subsets containing a fixed element.

Remark 8.1.6 When $2k = n$, the extremal families can be obtained by choosing one set from each complementary pair $\{A, A^c\}$ with $A \in \binom{[n]}{k}$. This yields an intersecting family of size $\frac{1}{2}\binom{n}{k} = \binom{n-1}{k-1}$, which need not be of the form described above.

8.2 Proof 1 of Erdös–Ko–Rado Theorem

The first proof we present is due to Katona (1972) and uses a double counting argument. The idea is to study all permutations of the elements of $[n]$, estimating how often the consecutive elements of these permutations can constitute one of the sets in our family.

Proof 1 of Theorem 8.1.5 Let $\mathcal{F}$ be an intersecting family of k-element subsets of $[n]$. Take a cyclic permutation $\pi = (a_1\, a_2\, \ldots\, a_n)$ of $[n]$. A set $A \subset [n]$ is said to be *contained* in π if $A = \{a_i, a_{i+1}, \ldots, a_{i+k-1}\}$ for some i, where the indices are taken modulo n. For each π, define

$$\mathcal{F}_\pi = \{A \in \mathcal{F} : A \text{ is contained in } \pi\}.$$

Claim 1 $|\mathcal{F}_\pi| \leqslant k$ for all cyclic permutations π.

Pick any $A \in \mathcal{F}_\pi$. There are exactly $2k-2$ distinct k-element sets B that are contained in π and intersect A, excluding A itself. These $2k-2$ sets can be grouped into $k-1$ pairs of disjoint subsets. Since $\mathcal{F}$ is intersecting, at most one set from each pair can belong to $\mathcal{F}$. Hence $|\mathcal{F}_\pi| \leqslant 1 + (k-1) = k$. //

We now count in two ways the number L of pairs (π, A), where π is a cyclic permutation of $[n]$ and $A \in \mathcal{F}_\pi$. By Claim 1,

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

On the other hand, fix a set $A \in \mathcal{F}$. The number of cyclic permutations π such that $A \in \mathcal{F}_\pi$ is $k!(n-k)!$. Indeed, we can arrange the elements of A in $k!$ ways, and the elements of $[n] \setminus A$ in $(n-k)!$ ways. Thus,

$$L = \sum_{A \in \mathcal{F}} k!(n-k)! = |\mathcal{F}| \cdot k!(n-k)!.$$

Combining this with the previous estimate, we obtain

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

Now suppose $2k < n$ and the equality holds, so that for every cyclic permutation π we have $|\mathcal{F}_\pi| = k$. We shall prove that this forces $\mathcal{F}$ to be a star.

Claim 2 Let $\pi = (a_1\, a_2\, \ldots\, a_n)$ be a cyclic permutation of $[n]$. Then there exists an $r \in [n]$ such that

$$\mathcal{F}_\pi = \big\{\{a_{r+j}, a_{r+j+1}, \ldots, a_{r+j+k-1}\} : 1 \leqslant j \leqslant k\big\},$$

where indices are taken modulo n.

Since $n > 2k$, the k sets in $\mathcal{F}_\pi$ cannot cover the whole circle, otherwise some of them would be disjoint. Hence their union is a proper arc of the cycle. Choose a block $A' \in \mathcal{F}_\pi$ whose boundary meets an endpoint of this arc. Shifting A' successively by one position along the circle produces $k-1$ further consecutive k-blocks, all intersecting A'. These k blocks already account for every k-block lying

entirely within the same arc; any other k-block would extend beyond the arc and be disjoint from A', contradicting the intersecting property. //

By Claim 2, we can assume $\mathcal{F}_\pi = \{A_1, A_2, \ldots, A_k\}$, where $A_j = \{a_j, a_{j+1}, \ldots, a_{j+k-1}\}$ for $1 \leqslant j \leqslant k$. Then $A_1 \cap A_2 \cap \cdots \cap A_k = \{a_k\}$. If $\mathcal{F}$ is not a star, then there exists a set $A_0 \in \mathcal{F}$ with $a_k \notin A_0$. We shall then show that $\mathcal{F} = \binom{A_1 \cup A_k}{k}$, from which it follows that $|\mathcal{F}| = \binom{2k-1}{k} = \binom{2k-1}{k-1} < \binom{n-1}{k-1}$. This contradiction will complete the proof.

Claim 3 For any $B \in \binom{A_1 \cup A_k \setminus \{a_k\}}{k-1}$, we have $B \cup \{a_k\} \in \mathcal{F}$.

Write $B = B_1 \cup B_2$ with $B_1 \subset A_1$ and $B_2 \subset A_k$. List all elements of the five sets

$$A_1 \setminus (B_1 \cup \{a_k\}), \quad B_1, \quad \{a_k\}, \quad B_2, \quad A_k \setminus (B_2 \cup \{a_k\})$$

in this order. Append the remaining elements of $[n]$ (those not contained in $A_1 \cup A_k$) in arbitrary order, and regard the resulting linear order as a cyclic permutation π'. Then $A_1, A_k \in \mathcal{F}_{\pi'}$ and $A_1 \cap A_k = \{a_k\}$. By Claim 2, we have $B \cup \{a_k\} \in \mathcal{F}_{\pi'} \subset \mathcal{F}$. //

Claim 4 $A_0 \subset A_1 \cup A_k \setminus \{a_k\}$.

If not, we have $A_0 \not\subset A_1 \cup A_k$ and $|A_0 \cap (A_1 \cup A_k)| \leqslant k-1$. Then there exists some $B \in \binom{(A_1 \cup A_k) \setminus A_0}{k}$ such that $a_k \in B$. By Claim 3, $B \in \mathcal{F}$, contradicting the fact that $A_0 \cap B \neq \varnothing$ since $\mathcal{F}$ is intersecting. //

Claim 5 $\binom{A_1 \cup A_k}{k} \subset \mathcal{F}$.

For any $i \in A_0$, let $B_i = ((A_1 \cup A_k) \setminus A_0) \cup \{i\}$. Since $a_k \in B_i$, Claim 3 implies that $B_i \in \mathcal{F}$. Repeating the proof of Claim 3 with a_k replaced by i, we obtain that any k-element subset of $A_1 \cup A_k$ containing i belongs to $\mathcal{F}$. By Claim 4, any k-element subset of $A_1 \cup A_k$ contains some $i \in A_0$, and hence belongs to $\mathcal{F}$. This shows that $\binom{A_1 \cup A_k}{k} \subset \mathcal{F}$. //

Claim 6 $\mathcal{F} \subset \binom{A_1 \cup A_k}{k}$.

Suppose that there exists some k-element set $C \in \mathcal{F}$ with $C \not\subset A_1 \cup A_k$. Then $|(A_1 \cup A_k) \setminus C| \geqslant k$. Hence we can find a $B \in \binom{(A_1 \cup A_k) \setminus C}{k}$. By Claim 5, $B \in \mathcal{F}$, but $B \cap C = \varnothing$, contradicting the fact that $\mathcal{F}$ is intersecting. //

Claims 5 and 6 together yield $\mathcal{F} = \binom{A_1 \cup A_k}{k}$, as desired. □

8.3 Proof 2 of Erdös–Ko–Rado Theorem

Definition 8.3.1 The **eigenvalues of a graph** are defined as the eigenvalues of its adjacency matrix.

Example 8.3.2 (1) The complete graph K_n has an adjacency matrix $J - I$, where J is the all-ones matrix and I is the identity matrix. The rank of J is 1, i.e., there is one nonzero eigenvalue equal to n, with an eigenvector $(1, 1, \ldots, 1)$. All the remaining eigenvalues are 0. Therefore the eigenvalues of K_n are $\lambda_1 = n - 1$ and $\lambda_2 = \cdots = \lambda_n = -1$.

(2) If G is d-regular, then $\mathbf{1} = (1, 1, \ldots, 1)$ is an eigenvector. We get $A\mathbf{1} = d\mathbf{1}$, and hence d is an eigenvalue. It is easy to show that no eigenvalue can be larger than d.

Definition 8.3.3 The greatest integer r such that G contains an independent set of size r is the **independence number** of G, and is denoted by $\alpha(G)$.

Theorem 8.3.4 (Hoffman's bound) Let G be a d-regular graph on n vertices with eigenvalues $d = \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. Then

$$\alpha(G) \leqslant n \cdot \frac{-\lambda_n}{\lambda_1 - \lambda_n}.$$

Proof Let $V(G) = [n]$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be an orthonormal basis of eigenvectors of the adjacency matrix $A(G) = (a_{i,j})_{n\times n}$ corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$. That is,

$$A(G)\mathbf{v}_i = \lambda_i \mathbf{v}_i, \quad \|\mathbf{v}_i\| = 1, \quad \langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0 \text{ for } i \neq j.$$

Let $I \subset V(G)$ be an independent set of size $\alpha(G)$. Define $\mathbb{1}_I \in \{0,1\}^n$ to be the characteristic vector of I, i.e., the j-th coordinate of $\mathbb{1}_I$ is 1 if $j \in I$ and 0 otherwise. We can write

$$\mathbb{1}_I = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$$

for some real numbers α_i. Then

$$|I| = \langle \mathbb{1}_I, \mathbb{1}_I \rangle = \sum_{i=1}^{n} \alpha_i^2. \tag{8–1}$$

Since G is d-regular, we have $\lambda_1 = d$ and $\mathbf{v}_1 = \left(1/\sqrt{n}, \ldots, 1/\sqrt{n}\right)^{\mathsf{T}}$. Hence,

$$\alpha_1 = \langle \mathbb{1}_I, \mathbf{v}_1 \rangle = \frac{|I|}{\sqrt{n}}. \tag{8–2}$$

Because I is an independent set, no two vertices in I are adjacent. Thus,

$$\mathbb{1}_I{}^{\mathsf{T}} A(G) \mathbb{1}_I = \sum_{i,j=1}^{n} (\mathbb{1}_I)_i a_{i,j} (\mathbb{1}_I)_j = 0.$$

On the other hand, we can express this as

$$\left(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i{}^{\mathsf{T}}\right)\left(\sum_{j=1}^{n} \alpha_j A(G)\mathbf{v}_j\right) = \left(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i{}^{\mathsf{T}}\right)\left(\sum_{j=1}^{n} \alpha_j \lambda_j \mathbf{v}_j\right) = \sum_{i=1}^{n} \alpha_i^2 \lambda_i.$$

By (8–1) and (8–2),

$$0 = \sum_{i=1}^{n} \alpha_i^2 \lambda_i \geqslant \alpha_1^2 \lambda_1 + \left(\alpha_2^2 + \cdots + \alpha_n^2\right)\lambda_n = \frac{|I|^2}{n}\lambda_1 + \left(|I| - \frac{|I|^2}{n}\right)\lambda_n.$$

Rearranging gives

$$|I|(\lambda_1 - \lambda_n) + n\lambda_n \leqslant 0,$$

which implies

$$\alpha(G) = |I| \leqslant n \cdot \frac{-\lambda_n}{\lambda_1 - \lambda_n}.$$

□

Definition 8.3.5 The **Kneser graph** $K(n,k)$ is the graph whose vertices correspond to $\binom{[n]}{k}$, and where two vertices are adjacent if and only if the two corresponding sets are disjoint.

Example 8.3.6 (1) $K(n,1)$ is the complete graph K_n.

(2) $K(5,2)$ is the Petersen graph.

Theorem 8.3.7 The spectrum of the Kneser graph $K(n,k)$ consists of $k+1$ distinct eigenvalues:

$$\mu_j = (-1)^j \binom{n-k-j}{k-j}, \quad j = 0, \ldots, k.$$

Moreover μ_j occurs with multiplicity $\binom{n}{j} - \binom{n}{j-1}$ for $j > 0$ and μ_0 has multiplicity 1.

Proof (Reference) A detailed proof can be found in *Algebraic Graph Theory* by Chris Godsil and Gordon Royle, Theorem 9.4.3. □

The Erdős–Ko–Rado theorem (the first part of Theorem 8.1.5) can also be described in terms of independent sets in Kneser graphs. Indeed, an intersecting family of k-element subsets of $[n]$ corresponds to an independent set in $K(n,k)$.

Theorem 8.3.8 (Erdös–Ko–Rado theorem) For $n \geqslant 2k$, we have $\alpha(K(n,k)) \leqslant \binom{n-1}{k-1}$.

Proof Consider the eigenvalues of the Kneser graph $K(n,k)$, denoted

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_{\binom{n}{k}}.$$

From Theorem 8.3.7, we know that

$$\lambda_1 = \mu_0 = \binom{n-k}{k} \quad \text{and} \quad \lambda_{\binom{n}{k}} = \mu_1 = -\binom{n-k-1}{k-1}.$$

By Theorem 8.3.4, we have

$$\alpha(K(n,k)) \leqslant \binom{n}{k} \frac{-\lambda_{\binom{n}{k}}}{\lambda_1 - \lambda_{\binom{n}{k}}} = \binom{n}{k} \frac{\binom{n-k-1}{k-1}}{\binom{n-k}{k} + \binom{n-k-1}{k-1}} = \binom{n-1}{k-1}.$$

□

Chapter 9 Chains and Antichains

9.1 Partially Ordered Sets

Definition 9.1.1 A **partial order**, is a relation $\preccurlyeq$ on a set X that is reflexive, antisymmetric, and transitive. That is, for all $x, y, z \in X$, it must satisfy:

(1) Reflexivity: $x \preccurlyeq x$, i.e. every element is related to itself.

(2) Antisymmetry: if $x \preccurlyeq y$ and $y \preccurlyeq x$ then $x = y$, i.e. no two distinct elements precede each other.

(3) Transitivity: if $x \preccurlyeq y$ and $y \preccurlyeq z$ then $x \preccurlyeq z$.

A **partially ordered set** (**poset** for short) is an ordered pair $P = (X, \preccurlyeq)$ consisting of a set X (called the **ground set** of P) and a partial order $\preccurlyeq$ on X. We write $x \prec y$ if $x \preccurlyeq y$ but $x \neq y$, and we say that x is a **predecessor** (or **child**) of y. Elements x and y are said to be **comparable** if either $x \preccurlyeq y$ or $y \preccurlyeq x$; otherwise, they are **incomparable**.

Example 9.1.2 Standard examples of posets arising in mathematics include:

(1) The power set of a given set ordered by inclusion.

(2) The set of positive integers ordered by divisibility.

Definition 9.1.3 Let $P = (X, \preccurlyeq)$ be a poset. We say that an element x is an **immediate predecessor** of y (or that y **covers** x) if $x \prec y$ and there is no $z \in X$ such that $x \prec z \prec y$. In this case, we write $x \lessdot y$.

Fact 9.1.4 For $x, y \in (X, \preccurlyeq)$, one has $x \prec y$ if and only if there exist $z_1, \ldots, z_k \in X$ with

$$x \lessdot z_1 \lessdot \cdots \lessdot z_k \lessdot y,$$

for some $k \geqslant 0$; when $k = 0$, this just means $x \lessdot y$.

Proof It suffices to prove the "only if" direction. Define

$$M_{x,y} = \{z \in X : x \prec z \prec y\}.$$

We proceed by induction on $|M_{x,y}|$. If $|M_{x,y}| = 0$, then $x \lessdot y$ and we are done. Now suppose $n \coloneqq |M_{x,y}| > 0$ and assume the claim holds for all $x' \prec y'$ with $|M_{x',y'}| < n$. Pick some $z \in M_{x,y}$. Then $|M_{x,z}|, |M_{z,y}| < n$, so by the inductive hypothesis, there exist $s_1, \ldots, s_k$ and $t_1, \ldots, t_m$ such that

$$x \lessdot s_1 \lessdot \cdots \lessdot s_k \lessdot z \quad \text{and} \quad z \lessdot t_1 \lessdot \cdots \lessdot t_m \lessdot y.$$

Combining these two chains gives the desired result. □

Definition 9.1.5 A **Hasse diagram** is type of mathematical diagram used to represent a finite partially ordered set, in the form of a drawing of its transitive reduction. Concretely, for a partially ordered set $P = (X, \preccurlyeq)$ one represents each element of X as a vertex in the plane and draws a line segment or curve that goes *upward* from one vertex x to another vertex y whenever y covers x.

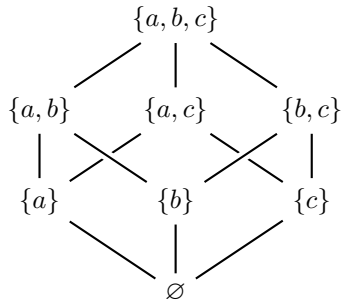


Figure 9.1: The Hasse diagram of the power set of $\{a, b, c\}$ ordered by inclusion

Fact 9.1.4 can now be rephrased as follows: in a Hasse diagram, there is a path going upward from x to y if and only if $x \prec y$.

Definition 9.1.6 Let $P = (X, \preccurlyeq)$ be a poset.

(1) A **chain** in P is a subset $C \subset X$ such that any two elements of C are comparable. The length of the longest chain in P is known as its **height**, denoted by $\omega(P)$.

(2) An **antichain** in P is a subset $A \subset X$ such that no two elements of A are comparable. The size of the largest antichain in P is known as its **width**, denoted by $\alpha(P)$.

Definition 9.1.7 Let $P = (X, \preccurlyeq)$ be a poset.

(1) An element $g \in X$ is a **maximal element** if there is no element $a \in X$ such that $g \prec a$.

(2) An element $m \in X$ is a **minimal element** if there is no element $b \in X$ such that $b \prec m$.

Fact 9.1.8 The set of all maximal (respectively, minimal) elements of a poset forms an antichain.

Theorem 9.1.9 Let $P = (X, \preccurlyeq)$ be a poset. Then $\omega(P)\alpha(P) \geqslant |X|$.

Proof We inductively construct posets $P_i = (X_i, \preccurlyeq)$ and sets $M_i \subset X_i$ such that M_i is the set of minimal elements of P_i, and

$$X_i = X \setminus \bigcup_{j=0}^{i-1} M_j, \quad M_0 = \varnothing.$$

Set $P_1 = P = (X, \preccurlyeq)$ and $X_1 = X$. Assume $P_i = (X_i, \preccurlyeq)$ and M_{i-1} have been defined. Let

$$M_i = \{\text{all minimal elements of } P_i\}, \quad X_{i+1} = X \setminus \bigcup_{j=0}^{i} M_j,$$

and let P_{i+1} be the subposet of P induced by X_{i+1}. Continue this process until $X_{\ell+1} = \varnothing$.

By Fact 9.1.8, each M_i is an antichain in P_i, and since P_i is a subposet of P, each M_i is also an antichain of P. Hence

$$|M_i| \leqslant \alpha(P) \quad \text{for all } i.$$

To complete the proof, it suffices to find a chain

$$x_1 \prec x_2 \prec \cdots \prec x_\ell$$

in P with $x_i \in M_i$ for every $i \in [\ell]$. Indeed, if such a chain exists, then

$$X = M_1 \sqcup M_2 \sqcup \cdots \sqcup M_\ell \quad \text{and} \quad |X| = \sum_{i=1}^{\ell} |M_i| \leqslant \ell\alpha(P) \leqslant \omega(P)\alpha(P). \tag{9–1}$$

We show that the desired chain exists. By the definition of M_i, for every $x \in M_i$ with $2 \leqslant i \leqslant \ell$ there exists some $y \in M_{i-1}$ such that $y \lessdot x$; otherwise x would already have been minimal in P_{i-1}, contradicting $x \notin M_{i-1}$. Thus each $x \in M_i$ covers at least one element of M_{i-1}, and by choosing one such predecessor at each step, we obtain the required chain $x_1 \prec x_2 \prec \cdots \prec x_\ell$.

This establishes (9–1), completing the proof. □

Corollary 9.1.10 Let $P = (X, \preccurlyeq)$ be a poset and let r and s be positive integers such that $|X| \geqslant rs + 1$. Then P contains either a chain of length $r + 1$ or an antichain of size $s + 1$.

9.2 Erdős–Szekeres Theorem

Theorem 9.2.1 (Erdős–Szekeres) Let $(a_1, \ldots, a_{rs+1})$ be a sequence of $rs + 1$ real numbers. Then the sequence contains either an increasing subsequence of length $r+1$ or a decreasing subsequence of length $s + 1$.

Proof 1 Let $X = [rs+1]$ and define a poset $P = (X, \preccurlyeq)$ by setting $i \preccurlyeq j$ if and only if $i \leqslant j$ and $a_i \leqslant a_j$. By Theorem 9.1.9, we have

$$\omega(P)\alpha(P) \geqslant |X| = rs + 1.$$

Hence, either $\omega(P) \geqslant r + 1$ or $\alpha(P) \geqslant s + 1$.

- ◇ If $\omega(P) \geqslant r + 1$, then there exist indices $i_1 < i_2 < \cdots < i_{r+1}$ such that $a_{i_1} \leqslant a_{i_2} \leqslant \cdots \leqslant a_{i_{r+1}}$.
- ◇ If $\alpha(P) \geqslant s + 1$, then there exist indices $j_1 < j_2 < \cdots < j_{s+1}$ such that they are pairwise incomparable in P. This means that $a_{j_1} > a_{j_2} > \cdots > a_{j_{s+1}}$. □

Exercise 9.2.2 Show that the bound in Theorem 9.2.1 is optimal by constructing a sequence of rs real numbers that contains no increasing subsequence of length $r+1$ and no decreasing subsequence of length $s+1$.

Proof Consider the sequence

$$\underbrace{s, s-1, \ldots, 1}_{\text{block } 1}, \quad \underbrace{2s, 2s-1, \ldots, s+1}_{\text{block } 2}, \quad \ldots, \quad \underbrace{rs, rs-1, \ldots, (r-1)s+1}_{\text{block } r}.$$

Note that any subsequence of length $r+1$ must contain at two elements from the same block, which are in decreasing order. Thus, there is no increasing subsequence of length $r+1$. Similarly, any subsequence of length $s+1$ must contain two elements from different blocks, which are in increasing order. Thus, there is no decreasing subsequence of length $s+1$. □

9.3 Miscellaneous Applications of Pigeonhole Principle

Theorem 9.3.1 Let G be a graph with n vertices where $n \geqslant 2$. Then there are at least two vertices in G with the same degree.

Proof Assume for the sake of contradiction that all vertices in G have distinct degrees. Then the degrees of the vertices must be $0, 1, \ldots, n-1$. However, if there is a vertex with degree $n-1$, then there cannot be a vertex with degree 0. This is a contradiction. □

Exercise 9.3.2 Given $n \geqslant 2$, find a graph with n vertices that has exactly two vertices with the same degree.

Solution We construct such a graph by induction on n. For $n = 2$, the construction is trivial. Assume we have such a graph on n vertices. There are two cases to consider:

- ◇ If no vertex has degree 0, we can add an isolated vertex. All existing degrees remain unchanged, so the only duplicated degree is the original pair.
- ◇ If some vertex has degree 0, then no vertex has degree $n-1$. We can add a new vertex connected to all existing vertices. The new vertex has degree n, and every old degree increases by 1, so none becomes n. Thus the only duplicated degree is again the original pair. □

Definition 9.3.3 Let G be a finite graph. The **chromatic number** of G, denoted by $\chi(G)$, is the minimum number of colors in a coloring of the vertices of G with the property that no two adjacent vertices have the same color. In other words, $\chi(G)$ is the smallest number of independent sets into which the vertex set of G can be partitioned.

Theorem 9.3.4 In any graph G with n vertices, we have $\alpha(G)\chi(G) \geqslant n$. Here, $\alpha(G)$ denotes the independence number of G (see Definition 2.1.14).

Proof Consider the vertices of G partitioned into $\chi(G)$ color classes. By the pigeonhole principle, one of the classes must contain at least $n/\chi(G)$ vertices, and these vertices are pairwise non-adjacent. Thus $\alpha(G) \geqslant n/\chi(G)$, as desired. □

Remark 9.3.5 There is a standard translation between Theorem 9.1.9 and Theorem 9.3.4. Given a poset P, its **comparability graph** G_P has edges between comparable elements. In G_P:

- ◇ Cliques correspond to chains, thus $\chi(G_P) \geqslant \omega(P)$.
- ◇ Independent sets correspond to antichains, thus $\alpha(G_P) = \alpha(P)$.

Therefore, By Theorem 9.1.9,

$$\alpha(G_P)\chi(G_P) \geqslant \alpha(P)\omega(P) \geqslant n.$$

Theorem 9.3.6 For any subset $S \subset [2n]$ with $|S| \geqslant n+1$, there exist two distinct elements $a, b \in S$ such that $a \mid b$.

Proof For each odd number $2k-1$ with $k \in [n]$, define

$$S_{2k-1} = \big\{2^i(2k-1) \in S : i \geqslant 0\big\}.$$

These sets form a partition of S into at most n subsets. Since $|S| \geqslant n+1$, at least one of these sets contains at least two elements, say $2^i(2k-1)$ and $2^j(2k-1)$ with $i < j$. Then $2^i(2k-1) \mid 2^j(2k-1)$. □

Theorem 9.3.7 (Dirichlet's approximation theorem) For any $x \in \mathbb{R}$ and any positive integer n, there is a rational number p/q such that $1 \leqslant q \leqslant n$ and

$$\left|x - \frac{p}{q}\right| < \frac{1}{nq} \leqslant \frac{1}{q^2}.$$

Proof Let $\{x\} \coloneqq x - \lfloor x \rfloor$ denote the fractional part of the real number x. Consider the $n+1$ numbers

$$\{ax\}, \quad a = 1, 2, \ldots, n+1.$$

We put these numbers into the n pigeonholes

$$[0, 1/n), \quad [1/n, 2/n), \ldots, [1-1/n, 1).$$

By the pigeonhole principle, some interval contains more than one of the numbers, say $\{ax\}$ and $\{bx\}$ with $a > b$, which therefore differ by less than $1/n$. Letting $q = a - b$, we see that there exists an integer $p = \lfloor ax \rfloor - \lfloor bx \rfloor$ such that

$$|qx - p| = |(ax - \lfloor ax \rfloor) - (bx - \lfloor bx \rfloor)| = |\{ax\} - \{bx\}| < 1/n,$$

from which the result follows on division by q. Moreover, q is the difference between two integers in $[n+1]$, so $1 \leqslant q \leqslant n$. □

We end this section with alternative proofs of the Erdős–Szekeres theorem and Mantel's theorem using the pigeonhole principle.

Proof 2 of Theorem 9.2.1 For each index $i \in [rs+1]$, let f_i be the maximum length of an increasing subsequence starting at a_i. We may assume that $f_i \in [r]$ for all i, since otherwise we are done. By the pigeonhole principle, there exist $s+1$ indices $i_1 < i_2 < \cdots < i_{s+1}$ such that $f_{i_1} = f_{i_2} = \cdots = f_{i_{s+1}} = t$.

We now claim that the corresponding terms form a decreasing sequence:

$$a_{i_1} \geqslant a_{i_2} \geqslant \cdots \geqslant a_{i_{s+1}}. \tag{9–2}$$

Indeed, suppose for contradiction that $a_{i_j} < a_{i_{j+1}}$ for some j. Since $f_{i_{j+1}} = t$, there is an increasing

subsequence

$$a_{i_{j+1}} < b_2 < \cdots < b_t.$$

But then

$$a_{i_j} < a_{i_{j+1}} < b_2 < \cdots < b_t$$

is an increasing subsequence of length $t+1$ starting at a_{i_j}, contradicting $f_{i_j} = t$. Thus (9–2) holds. □

Proof 4 of Theorem 4.4.1 To avoid ceilings and floorings, we will prove the theorem for graphs on an even number $2n$ of vertices. We want to prove that every such graph with at least n^2+1 edges must contain a triangle. We argue by induction on n. If $n = 1$, then G cannot have n^2+1 edges; hence the statement is true. Assuming the result for n, we now consider a graph G on $2(n+1)$ vertices with at least $(n+1)^2+1$ edges. Let x and y be adjacent vertices in G, and let H be the induced subgraph on the remaining $2n$ vertices. If H contains at least n^2+1 edges then we are done by the induction hypothesis. Suppose that H has at most n^2 edges, and therefore at least

$$(n+1)^2 + 1 - n^2 - 1 = 2n + 1$$

edges of G emanate from x and y to vertices in H. By the pigeonhole principle, among these $2n+1$ edges there must be an edge from x and an edge from y to the same vertex z in H. Hence G contains the triangle $\{x, y, z\}$. □

9.4 Decomposition in Chains and Antichains

A decomposition of a poset is its partition into mutually disjoint chains or antichains. Given a poset P, our goal is to decompose it into as few chains (or antichains) as possible. One direction is easy: if a poset P has a chain (antichain) of size r then it cannot be partitioned into fewer than r antichains (chains). The reason here is simple: any two points of the same chain must lie in different members of a partition into antichains.

Is this optimal? If P has no chain (or antichain) of size greater than r, is it then possible to partition P into r antichains (or chains, respectively)? One direction is straightforward (see Problem **??**). The dual result looks similar, but its proof is more involved. This result, uniformly known as Dilworth's decomposition theorem (Dilworth 1950), has played an important role in motivating research into posets. There are several elegant proofs; the one we present is due to F. Galvin.

Theorem 9.4.1 (Dilworth's decomposition theorem) Suppose that the largest antichain in a poset P has size r. Then P can be partitioned into r chains.

Proof We use induction on the cardinality of P. Let a be a maximal element of P, and let r be the size of a largest antichain in $P' := P \setminus \{a\}$. Then P' is the union of r disjoint chains $C_1, \ldots, C_r$. We have to show that P either contains an $(r+1)$-element antichain or else is the union of r disjoint chains. Now, every r-element antichain in P' consists of one element from each C_i. Let a_i be the maximal element in C_i which belongs to some r-element antichain in P'. It is easy to see that $A = \{a_1, \ldots, a_r\}$ is an antichain in P'. If $A \cup \{a\}$ is an antichain in P, we are done: we have found an antichain of size $r+1$. Otherwise, we have $a \succcurlyeq a_i$ for some i. Then $K = \{a\} \cup \{x \in C_i : x \preccurlyeq a_i\}$ is a chain in P, and there are no r-element antichains in $P \setminus K$ (since a_i was the maximal element of C_i participating in such an antichain), whence $P \setminus K$ is the union of $r-1$ chains. □

To recognize the power of this theorem, let us show that it contains Hall's Marriage Theorem 6.1.2 as a special case.

Proof 2 of Theorem 6.1.2 Suppose that Hall's condition (6–1) holds. We construct a poset P as follows. The points of P are the elements of $X := S_1 \cup \cdots \cup S_m$ and symbols $y_1, \ldots, y_m$, with $x \prec y_i$ if $x \in S_i$, and no other comparabilities. It is clear that X is an antichain in P. We claim that there is no larger antichain. To show this, let A be an antichain, and set $I := \{i : y_i \in A\}$. Then A contains no point of $\bigcup_{i \in I} S_i$, for if $x \in S_i$ then x is comparable with y_i, and hence, A cannot contain both of these points. Hence, Hall's condition (6–1) implies that

$$|A| \leqslant |I| + |X| - \left|\bigcup_{i \in I} S_i\right| \leqslant |X|,$$

as desired. Now, Theorem 9.4.1 implies that P can be partitioned into $|X|$ chains. Since the antichain X is maximal, each of the chains in the partition must contain a point of X. Let the chain through y_i be $\{x_i, y_i\}$. Then $(x_1, \ldots, x_m)$ is a desired system of distinct representatives: for $x_i \in S_i$ (since $x_i \prec y_i$) and $x_i \neq x_j$ (since the chains are disjoint). □

Chapter 10 Ramsey Theory

10.1 Ramsey's Theorem for Graphs

How many people can be invited to a party such that among any three guests, there are two who know each other and two who do not? It turns out that at most five people can attend such a party.

To see this, consider the following simple game. Mark six points on a sheet of paper, with no three points collinear. Two players take turns; one uses a red pencil and the other a blue pencil. On each turn, a player draws a line in their own color between two points that have not yet been joined (line crossings are allowed). The goal of each player is to form a monochromatic triangle. If you try playing this game with a friend, you will find that it always ends with one player winning; a draw is impossible.

We can model this scenario using a graph G on six vertices, where each vertex represents a participant and an edge exists between two vertices if and only if those participants know each other. To prove that a draw is impossible, we must show that any such graph contains either a triangle K_3 or an independent set I_3. Consider vertex 1; it must either be adjacent to at least three vertices or non-adjacent to at least three vertices. By symmetry, we may assume vertex 1 is adjacent to vertices 2, 3, and 4. If any pair among $\{2,3,4\}$ is adjacent, those two vertices together with vertex 1 form a K_3. If no two of these vertices are adjacent, then $\{2,3,4\}$ itself forms an I_3. In either case, the property is satisfied.

Definition 10.1.1 A **k-edge-coloring** of a graph is an assignment of one of k colors to each edge of the graph.

Definition 10.1.2 For $s,t \geqslant 2$, the **Ramsey number** $R(s,t)$ is the smallest integer n such that any 2-edge-coloring of K_n contains a red K_s or a blue K_t.

Fact 10.1.3 (1) $R(s,t) = R(t,s)$.

(2) $R(2,t) = t$.

(3) $R(s,2) = s$.

Example 10.1.4 The previous example shows that any 2-edge-coloring of K_6 contains a monochromatic triangle, so $R(3,3) \leqslant 6$. To see that $R(3,3) > 5$, consider the following 2-edge-coloring of K_5 shown in Figure 10.1, which contains no monochromatic K_3. Hence $R(3,3) = 6$.

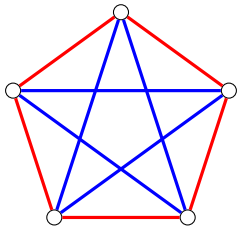


Figure 10.1: A 2-edge-coloring of K_5 with no monochromatic K_3

Theorem 10.1.5 $R(s,t) \leqslant R(s,t-1) + R(s-1,t)$.

Proof Let $G = (V, E)$ be a graph on $n = R(s,t-1) + R(s-1,t)$ vertices. Consider any 2-edge-coloring of G using the colors red and blue. Take an arbitrary vertex $x \in V$, and split $V \setminus \{x\}$ into two subsets:

$$S = \{y \in V \setminus \{x\} : xy \text{ is blue}\}, \quad T = \{y \in V \setminus \{x\} : xy \text{ is red}\}.$$

Since

$$R(s,t-1) + R(s-1,t) = n = |S| + |T| + 1,$$

we have either $|S| \geqslant R(s,t-1)$ or $|T| \geqslant R(s-1,t)$.

If $|S| \geqslant R(s,t-1)$, then the induced subgraph $G[S]$ of G contains either a red K_s or a blue K_{t-1}. In the former case, we are done. In the latter case, adding vertex x to this blue K_{t-1} gives a blue K_t in G. The case $|T| \geqslant R(s-1,t)$ is similar. □

Theorem 10.1.6 (Ramsey's theorem) $R(s,t) \leqslant \binom{s+t-2}{s-1} = \binom{s+t-2}{t-1} < \infty$.

Proof We proceed by induction on $s+t$. The base case $s = t = 2$ is clear by Fact 10.1.3. Assuming the result holds for all pairs (s', t') with $s' + t' < s + t$, we have by Theorem 10.1.5 and Fact 1.1.2 that

$$\begin{aligned} R(s,t) &\leqslant R(s,t-1) + R(s-1,t) \leqslant \binom{s+(t-1)-2}{(t-1)-1} + \binom{(s-1)+t-2}{t-1} \\ &= \binom{s+t-3}{t-2} + \binom{s+t-3}{t-1} = \binom{s+t-2}{t-1}. \end{aligned}$$

□

Theorem 10.1.7 If the Ramsey numbers $R(s,t-1)$ and $R(s-1,t)$ are both even, then Theorem 10.1.5 can be strengthened to

$$R(s,t) \leqslant R(s,t-1) + R(s-1,t) - 1.$$

Proof Let $n = R(s,t-1) + R(s-1,t) - 1$. Given the hypothesis that both Ramsey numbers are even, it follows that n is odd. Consider any 2-edge-coloring of K_n using the colors red and blue. For any vertex x, we define the sets $S_x = \{y : xy \text{ is red}\}$ and $T_x = \{y : xy \text{ is blue}\}$. As in the proof of Theorem 10.1.5, if $|S_x| \geqslant R(s,t-1)$ or $|T_x| \geqslant R(s-1,t)$, we are done. Thus, we may assume that for every vertex x, $|S_x| \leqslant R(s,t-1) - 1$ and $|T_x| \leqslant R(s-1,t) - 1$. Then

$$R(s,t-1) + R(s-1,t) - 1 = n = |S_x| + |T_x| + 1 \leqslant [R(s,t-1) - 1] + [R(s-1,t) - 1] + 1,$$

which forces $|S_x| = R(s,t-1) - 1$ and $|T_x| = R(s-1,t) - 1$ for every vertex x. Now, consider the graph G formed exclusively by the red edges. Every vertex of G has degree $R(s,t-1) - 1$, which is odd. However, this contradicts Theorem 2.2.6 since G has an odd number of vertices. □

Corollary 10.1.8 $R(3,4) = 9$.

Proof Since both $R(2,4) = 4$ (Fact 10.1.3 (2)) and $R(3,3) = 6$ (Example 10.1.4) are even, Theorem 10.1.7 applies and yields

$$R(3,4) \leqslant R(2,4) + R(3,3) - 1 = 4 + 6 - 1 = 9.$$ □

To see that $R(3,4) > 8$, consider the following 2-edge-coloring of K_8 shown in Figure 10.2, which contains no red K_3 or blue K_4. Hence $R(3,4) = 9$.

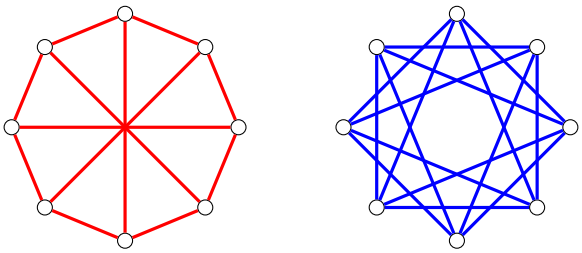


Figure 10.2: A 2-edge-coloring of K_8 with no red K_3 or blue K_4

Definition 10.1.9 For $c \geqslant 3$ and $s_1, \ldots, s_c \geqslant 2$, the **multicolor Ramsey number** $R_c(s_1, \ldots, s_c)$ is the smallest integer n such that any c-edge-coloring of K_n contains a monochromatic K_{s_i} in color i for some $i \in [c]$.

Theorem 10.1.10 If $c \geqslant 3$, then $R_c(s_1, \ldots, s_c) \leqslant R_{c-1}(s_1, \ldots, s_{c-2}, R(s_{c-1}, s_c))$.

Proof Consider a complete graph of $R_{c-1}(s_1, \ldots, s_{c-2}, R(s_{c-1}, s_c))$ vertices and color its edges with c colors. Now "go color-blind" and pretend that c–1 and c are the same color. Thus the graph is now $(c-1)$-colored. Due to the definition of $R_{c-1}(s_1, \ldots, s_{c-2}, R(s_{c-1}, s_c))$, such a graph contains either a monochromatic K_{s_i} in color i for some $i \in [c-2]$ or a monochromatic $K_{R(s_{c-1}, s_c)}$ in the combined color of $c-1$ and c. In the former case we are finished. In the latter case, we recover our sight again and see from the definition of $R(s_{c-1}, s_c)$ that this $K_{R(s_{c-1}, s_c)}$ contains either a monochromatic $K_{s_{c-1}}$ in color $c-1$ or a monochromatic K_{s_c} in color c. In either case the proof is complete. □

Remark 10.1.11 The right-hand side of the inequality expresses a Ramsey number for c colors in terms of Ramsey numbers for fewer colors. Therefore, any $R_c(s_1, \ldots, s_c)$ is finite for any number of colors.

Theorem 10.1.12 (Schur's theorem) For every integer $c \geqslant 2$, there exists a positive integer $N = N(c)$ such that for any coloring $\varphi\colon [N] \to [c]$, there exist $x, y, z \in [N]$ with $x + y = z$ and $\varphi(x) = \varphi(y) = \varphi(z)$.

Proof 1 Let $N = R_c(3, \ldots, 3)$. Given a coloring $\varphi\colon [N] \to [c]$, define a c-edge-coloring of K_N as follows: for any $i, j \in [N]$, assign the edge $\{i, j\}$ the color $\varphi(|i-j|)$. By the definition of $R_c(3, \ldots, 3)$, this coloring contains a monochromatic triangle, say with vertices i, j, k. Without loss of generality, assume $i < j < k$. Then

$$\varphi(k-j) = \varphi(j-i) = \varphi(k-i).$$

Let $x = k-j$, $y = j-i$, and $z = k-i$, all of which lie in $[N]$. We then have $x + y = z$ and $\varphi(x) = \varphi(y) = \varphi(z)$, as desired. □

Proof 2 Let $N = \lceil ec! \rceil$. Suppose, for the sake of contradiction, that there do not exist positive integers x, y with $x + y \leqslant N$ such that $\varphi(x) = \varphi(y) = \varphi(x + y)$.

Let c_0 be a color that appears most frequently among the integers in $[N]$, and let $x_0 < x_1 < \cdots < x_{n_1-1}$ be the integers of color c_0. By the pigeonhole principle, we have $N \leqslant cn_1$.

Consider the set $A_0 = \{x_i - x_0 : 1 \leqslant i < n_1\}$. By our assumption, no element of A_0 can receive color c_0; hence A_0 is colored using only $c-1$ colors. Let c_1 be a color that appears most frequently on A_0, and let $y_0 < y_1 < \cdots < y_{n_2-1}$ be the elements of A_0 colored c_1. Then $n_1 - 1 \leqslant (c-1)n_2$.

Next, consider the set $A_1 = \{y_i - y_0 : 1 \leqslant i < n_2\}$. Again by our assumption, no element of A_1 can receive either of the colors c_0 or c_1. Thus, A_1 is colored by at most $c-2$ colors. Let c_2 be a color appearing most frequently in A_1, and let $z_0 < z_1 < \cdots < z_{n_3-1}$ be the elements of A_1 colored c_2. Then $n_2 - 1 \leqslant (c-2)n_3$.

Continuing this procedure, we eventually obtain a sequence of integers $n_1, n_2, \ldots, n_k$ such that $n_k = 1$. Since there are only c colors, this process must terminate for some $k \leqslant c$. Collecting the inequalities, we have

$$N \leqslant cn_1 \quad \text{and} \quad n_i \leqslant (c-i)n_{i+1} + 1 \quad \text{for } i \in [k-1],$$

with $n_k = 1$. Combining these inequalities yields

$$N \leqslant \sum_{i=0}^{c-1} c(c-1)(c-2)\cdots(c-i) = \sum_{i=0}^{c-1} \frac{c!}{i!} < c! \sum_{i=0}^{\infty} \frac{1}{i!} = ec!.$$

This contradicts our choice of N, completing the proof. □

Schur (1916) used Theorem 10.1.12 to show that Fermat's Last Theorem is false for the finite field $\mathbb{F}_p$ for any sufficiently large prime p.

Theorem 10.1.13 For every integer $n \geqslant 1$, there exists p_0 such that for all primes $p \geqslant p_0$, the equation

$$x^n + y^n \equiv z^n \pmod{p}$$

has a nontrivial solution.

Proof The multiplicative group $\mathbb{F}_p^* = \{1, 2, \ldots, p-1\}$ is known to be cyclic and hence it has a generator g. Each element of $\mathbb{F}_p^*$ can be expressed as $x = g^{nj+i}$ where $0 \leqslant i < n$. We define a coloring $\varphi\colon \mathbb{F}_p^* \to \{0, 1, \ldots, n-1\}$ by $\varphi(x) = i$ if $x = g^{nj+i}$. By Theorem 10.1.12, for sufficiently large p, there exist $x', y', z' \in \mathbb{F}_p^*$ such that $x' + y' = z'$ and $\varphi(x') = \varphi(y') = \varphi(z') = i$. Therefore, $x' = g^{nj_x+i}$, $y' = g^{nj_y+i}$, $z' = g^{nj_z+i}$ and

$$g^{nj_x+i} + g^{nj_y+i} \equiv g^{nj_z+i} \pmod{p}.$$

Setting $x = g^{j_x}$, $y = g^{j_y}$, and $z = g^{j_z}$ gives

$$x^n + y^n \equiv z^n \pmod{p}.$$

□

10.2 Ramsey's Theorem for Sets

We now consider colorings of k-element subsets of $[n]$ for $k > 2$.

Theorem 10.2.1 (Ramsey's theorem) For any natural numbers $1 \leqslant k \leqslant s$ and $r \geqslant 2$ there exists a natural number $n = R_r(k; s)$ such that whenever k-subsets of $[n]$ are colored in r colors, there is an s-subset of $[n]$ whose all k-subsets receive the same color.

Proof We first observe that it is enough to consider the case of $r = 2$ colors.

Claim $R_{r+1}(k; s) \leqslant R_r(k; R_2(k; s))$.

Let $N = R_r(k; R_2(k; s))$ and consider an arbitrary coloring of the k-subsets of an N-element set X using $r + 1$ colors $0, 1, \ldots, r$. Now "go color-blind" and pretend that colors 0 and 1 are the same color. Thus, the k-subsets of X are now colored using only r colors. By the definition of $R_r(k; R_2(k; s))$, there exists $Y \subset X$ with $|Y| = R_2(k; s)$ so that all k-subsets of Y are colored with the same color in this new coloring. If this color is the combined color of 0 and 1, then by the definition of $R_2(k; s)$, there exists an s-subset of Y whose all k-subsets are colored either 0 or 1. Otherwise, all k-subsets of Y are colored with the same color from $\{2, \ldots, r\}$, and we are done. //

By this claim, it suffices to show that $R_2(k; s)$ exists.

In order to argue by induction, we define a more "granulated" version of the Ramsey number $R_2(k; s)$. Namely, let $R(k; s, t)$ be the smallest integer n such that whenever the k-subsets of an n-element set are colored in two colors (say 0 and 1), there exists either an s-subset whose all k-subsets are colored 0 or a t-subset whose all k-subsets are colored 1. Clearly, $R_2(k; s) = R(k; s, s)$. We will prove a more general statement that $R(k; s, t) \leqslant n$, where

$$n \coloneqq R(k - 1; R(k; s - 1, t), R(k; s, t - 1)) + 1.$$

We proceed by induction on k and on s, t. Observe that, by the pigeonhole principle, $R(1; s, t) = s+t-1$ for all s and t and, moreover, $R(k; x, k) = R(k; k, x) = x$ for all k and $x \geqslant k$. By induction, we may assume that the numbers $R(k; s - 1, t)$ and $R(k; s, t - 1)$ exist, and take an arbitrary n-element set X, where n is defined as above.

Let φ be a coloring of the k-subsets of X using colors 0 and 1. Fix a point $x \in X$, and let $X' \coloneqq X \backslash \{x\}$. We define a new coloring φ' of the $(k - 1)$-subsets A of X' by setting $\varphi'(A) = \varphi(A \cup \{x\})$. By the choice of n and by symmetry, we can assume that there exists a subset $Y \subset X'$ with $|Y| = R(k; s - 1, t)$ such that all $(k - 1)$-subsets of Y are colored 0 by φ'.

Now consider the coloring φ restricted to the k-subsets of Y. By the definition of $R(k; s-1, t)$, there exists either an $(s - 1)$-subset of Y whose all k-subsets are colored 0 by φ or a t-subset of Y whose all k-subsets are colored 1 by φ. In the latter case, we are done. In the former case, adding the point x to this $(s - 1)$-subset gives an s-subset of X whose all k-subsets are colored 0 by φ. □

One of the earliest and most popular applications of Theorem 10.2.1 is due to Erdős and Szekeres.

Theorem 10.2.2 (Erdős–Szekeres theorem) Let $m \geqslant 3$ be a positive integer. Then there exists a positive integer n such that any set of n points in the Euclidean plane, no three of which are collinear, contains m points that are the vertices of a convex m-gon.

Proof Choose $n = R_2(3; m)$, the number from Theorem 10.2.1, and let A be any set of n points in the plane, no three of which are collinear. For $a, b, c \in A$, let $|abc|$ denote the number of points of A which lie in the interior of the triangle formed by a, b, c. We color the 3-subsets of A as follows: if $|abc|$ is even, color $\{a, b, c\}$ by 0; otherwise, color it by 1. By the choice of n and by Theorem 10.2.1, there exists an m-subset $B \subset A$ such that all 3-subsets of B receive the same color. We claim that B is the vertex set of a convex m-gon. Otherwise, there would exist four points $a, b, c, d \in B$ such that one of them, say d, lies in the interior of the triangle formed by the other three. Since no three points of B are collinear, we have

$$|abc| = |abd| + |acd| + |bcd| + 1,$$

which is a contradiction since the parity of $|abc|$ differs from that of $|abd|$, $|acd|$, and $|bcd|$. □

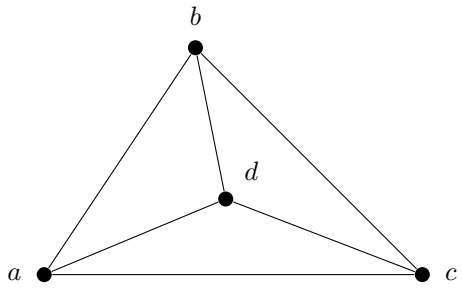


Figure 10.3: Point d lies in none of the lines determined by points a, b, and c

Chapter 11 Probabilistic Method

11.1 Probabilistic Preliminaries

Definition 11.1.1 A **finite probability space** consists of a finite set Ω and a function $\mathbb{P}\colon 2^{\Omega} \to [0,1]$, such that

- $\mathbb{P}(\varnothing) = 0$.
- $\mathbb{P}(\Omega) = 1$.
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ for disjoint sets $A, B \subseteq \Omega$.

Subsets $A \subset \Omega$ are called **events**. A **random variable** is a function $X\colon \Omega \to \mathbb{R}$. The **expectation** of a random variable X is defined as

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

Fact 11.1.2 (Union bound) $\mathbb{P}(A \cup B) \leqslant \mathbb{P}(A) + \mathbb{P}(B)$ for all events $A, B \subseteq \Omega$.

Fact 11.1.3 (Linearity of expectation) If $X_1, \ldots, X_n$ are random variables on the same probability space and $a_1, \ldots, a_n \in \mathbb{R}$, then

$$\mathbb{E}[a_1 X_1 + \cdots + a_n X_n] = a_1 \mathbb{E}[X_1] + \cdots + a_n \mathbb{E}[X_n].$$

Roughly speaking, the probabilistic method works as follows: trying to prove that an object with certain properties exists, one defines an appropriate probability space of objects and shows that a randomly chosen element of this space has the desired properties with a positive probability.

11.2 Applications of Union Bound

Theorem 11.2.1 Let n and s satisfy $\binom{n}{s}2^{1-\binom{s}{2}} < 1$. Then $R(s,s) > n$.

Proof We wish to show that there exists a 2-edge-coloring of K_n containing no monochromatic K_s.

Let Ω be the set of all 2-edge-colorings of K_n, and choose $c \in \Omega$ uniformly at random. Let A be the event that c contains a monochromatic copy of K_s. For $S \in \binom{[n]}{s}$, let A_S be the event that the vertices in S span a monochromatic K_s under c. Then

$$A = \bigcup_{S \in \binom{[n]}{s}} A_S \quad \text{and} \quad \mathbb{P}(A_S) = 2^{1-\binom{s}{2}}.$$

By Fact 11.1.2,

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

Hence with positive probability the coloring c contains no monochromatic K_s. Thus, such a coloring exists. □

Corollary 11.2.2 $R(s,s) > \frac{1}{e\sqrt{2}} s 2^{s/2}$.

Proof Let $n = \frac{1}{e\sqrt{2}} s 2^{s/2}$. By Fact 1.4.8 and Theorem 1.4.1,

$$\binom{n}{s}2^{1-\binom{s}{2}} < \frac{n^s}{s!} \cdot 2^{1-\binom{s}{2}} \leqslant \frac{n^s}{e(s/e)^s}2^{1-\binom{s}{2}} = 1.$$

Hence by Theorem 11.2.1, $R(s,s) > n$. □

Definition 11.2.3 (Erdős–Rényi model) The random graph $G(n,p)$ is obtained by taking the vertex set $\{1,\dots,n\}$ and including each of the $\binom{n}{2}$ possible edges independently with probability p.

Remark 11.2.4 Here $G(n,p)$ is a graph-valued random variable. By "the graph $G(n,p)$" we mean a single realization of this random variable.

Definition 11.2.5 Let A be a graph property. We say that the random graph $G(n,1/2)$ **almost surely satisfies** A if

$$\lim_{n\to\infty} \mathbb{P}(G(n,1/2) \text{ satisfies } A) = 1,$$

and **almost surely does not satisfy** A if

$$\lim_{n\to\infty} \mathbb{P}(G(n,1/2) \text{ satisfies } A) = 0.$$

Theorem 11.2.6 The random graph $G(n,1/2)$ is almost surely not bipartite.

Proof Let A be the event that $G(n,1/2)$ is bipartite. For any subset $S \subseteq [n]$, let A_S denote the event that all edges of $G(n,1/2)$ run between S and its complement $[n] \setminus S$. Then $G(n,1/2)$ is bipartite if and only if there exists such a partition, and hence

$$A = \bigcup_{S \subset [n]} A_S.$$

For a fixed S, the number of graphs satisfying A_S is $2^{|S|(n-|S|)}$. Therefore,

$$\mathbb{P}(A_S) = \frac{2^{|S|(n-|S|)}}{2^{\binom{n}{2}}} \leqslant \frac{2^{n^2/4}}{2^{n(n-1)/2}} = 2^{-\frac{n^2}{4}+\frac{n}{2}},$$

where we used the bound $|S|(n-|S|) \leqslant n^2/4$. Applying Fact 11.1.2, we obtain

$$0 \leqslant \mathbb{P}(A) \leqslant \sum_{S \subset [n]} \mathbb{P}(A_S) \leqslant 2^n \cdot 2^{-\frac{n^2}{4}+\frac{n}{2}} = 2^{-\frac{n^2}{4}+\frac{3n}{2}}.$$

Thus $\lim_{n\to\infty} \mathbb{P}(A) = 0$, and the result follows. □

11.3 Applications of Linearity of Expectation

Definition 11.3.1 A subset B of an additive group is called **sum-free** if $x + y \notin B$ for all $x, y \in B$.

Example 11.3.2 (1) The set of all odd integers in $[n]$ is a sum-free subset of $[n]$ of size $\lceil n/2 \rceil$.

(2) The set $\{\lfloor n/2 \rfloor + 1, \lfloor n/2 \rfloor + 2, \ldots, n\}$ is a sum-free subset of $[n]$ of size $\lceil n/2 \rceil$.

Exercise 11.3.3 Show that the maximum size of a sum-free subset of $[n]$ is $\lceil n/2 \rceil$.

Proof Let B be a sum-free subset of $[n]$ and m be its largest element. Consider the pairs

$$(1, m-1), (2, m-2), (3, m-3), \ldots, (\lfloor m/2 \rfloor, \lceil m/2 \rceil).$$

Since B is sum-free, it can contain at most one element from each of these pairs, and it cannot contain $m/2$ if m is even. Thus,

$$|B| \leqslant 1 + (\lfloor m/2 \rfloor - 1) + \mathbb{1}_{\{m \text{ is odd}\}} = \lceil m/2 \rceil.$$

Since $m \leqslant n$, we have $|B| \leqslant \lceil n/2 \rceil$. Combined with Example 11.3.2, this completes the proof. □

Theorem 11.3.4 Let $A \subset \mathbb{Z} \setminus \{0\}$ be finite. Then there exists a sum-free subset B of A with $|B| > |A|/3$.

Proof Let $p = 3k + 2$ be a prime number such that $p > 2 \max_{a \in A} |a|$. Such a prime exists by Dirichlet's prime number theorem. Write $S = \{k+1, k+2, \ldots, k+(k+1)\}$, and observe that S is a sum-free subset of $\mathbb{Z}_p$. Indeed, the sum $(k+1)+(k+1) = 2k+2 > k+(k+1)$ is too large, whereas the sum $(2k+1)+(2k+1) = 4k+2 \equiv k \pmod p < k+1$ is too small.

We choose a subset of A as follows. Pick a random element $t \in \mathbb{Z}_p \setminus \{0\}$, and let

$$A_t = \{a \in A : \text{the remainder of } at \text{ modulo } p \text{ lies in } S\}.$$

Note that A_t is sum-free, because for any $a, b \in A_t$, the remainders of at and bt modulo p lie in S, whereas the remainder of $(a+b)t$ modulo p cannot lie in S, by the sum-free property of S in $\mathbb{Z}_p$.

It remains to show that there exists t such that $|A_t| > |A|/3$. To do this, observe that for any fixed $a \in A$, as t ranges over all numbers in $[p-1]$, the remainder of at modulo p takes each value in $\mathbb{Z}_p \setminus \{0\}$ exactly once. Therefore, for any fixed $a \in A$,

$$\mathbb{P}(\text{the remainder of } at \text{ modulo } p \text{ lies in } S) = \frac{|S|}{|\mathbb{Z}_p \setminus \{0\}|} = \frac{|S|}{p-1} = \frac{k+1}{3k+1} > \frac{1}{3}.$$

By Fact 11.1.3, we have

$$\mathbb{E}[|A_t|] = \sum_{a\in A} \mathbb{P}(a \in A_t) = \sum_{a\in A} \mathbb{P}(\text{the remainder of } at \text{ modulo } p \text{ lies in } S) > \frac{|A|}{3}.$$

By the pigeonhole property of expectation, there is a value of t for which $|A_t| > |A|/3$. □

Definition 11.3.5 A **dominating set** of vertices in a graph $G = (V, E)$ is a set $S \subset V$ such that every vertex of G belongs to S or has a neighbor in S.

Theorem 11.3.6 If $G = (V, E)$ is an n-vertex graph with minimum degree $\delta > 1$, then G has a dominating set with at most

$$n \cdot \frac{1 + \ln(\delta + 1)}{\delta + 1}$$

vertices.

Proof Construct a random vertex subset $A \subset V$ by including each vertex independently with probability $p \in [0, 1]$ to be chosen later. Given A, let B be the set of vertices not in A and not adjacent to any vertex in A; adding B to A yields a dominating set. Since each vertex appears in A with probability p, we have $\mathbb{E}[|A|] = np$.

To estimate $\mathbb{E}[|B|]$, observe that for any vertex $v \in V$, the probability that $v \in B$ is the probability that neither v nor any of its at least δ neighbors appear in A, which is bounded above by $(1-p)^{\delta+1}$. Thus,

$$\mathbb{E}[|B|] = \sum_{v\in V} (1-p)^{\delta+1} = n(1-p)^{\delta+1} \leqslant ne^{-p(\delta+1)}.$$

Therefore,

$$\mathbb{E}[|A \cup B|] \leqslant \mathbb{E}[|A|] + \mathbb{E}[|B|] \leqslant np + ne^{-p(\delta+1)}.$$

Choosing $p = \dfrac{\ln(\delta+1)}{\delta+1}$ minimizes the right-hand side, yielding

$$\mathbb{E}[|A \cup B|] \leqslant n \cdot \frac{1 + \ln(\delta + 1)}{\delta + 1}.$$

Hence there exists a dominating set of size at most $n \cdot \dfrac{1 + \ln(\delta + 1)}{\delta + 1}$. □

Theorem 11.3.7 Let G be a graph on n vertices and let d_i denote the degree of the i-th vertex. Then

$$\alpha(G) \geqslant \sum_{i=1}^{n} \frac{1}{d_i + 1}.$$

Proof Let $V = [n]$ and let $\pi \colon V \to V$ be a random permutation taking its values uniformly and independently with probability $1/n!$. Let A_i be the event that $\pi(j) > \pi(i)$ for all d_i neighbors j of vertex i. There are $\binom{n}{d_i+1}$ ways to choose a $(d_i + 1)$-element set $S \subset V$ of possible π-images of i and all its d_i neighbors. After that there are $(|S| - 1)! = d_i!$ ways to arrange the π-images of the neighbors of i within S after fixing the smallest element of S to be $\pi(i)$, and $(n - |S|)! = (n - d_i - 1)!$ ways to arrange the remaining π-images. Thus,

$$\mathbb{P}(A_i) = \frac{\binom{n}{d_i+1} d_i!(n - d_i - 1)!}{n!} = \frac{1}{d_i + 1}.$$

Let U be the set of all vertices i such that A_i occurs. By Fact 11.1.3,

$$\mathbb{E}[|U|] = \sum_{i=1}^{n} \mathbb{P}(A_i) = \sum_{i=1}^{n} \frac{1}{d_i + 1}.$$

Thus, for some specific ordering,

$$|U| \geqslant \sum_{i=1}^{n} \frac{1}{d_i + 1}.$$

Now let $\{i, j\}$ be an edge of G. Then either $\pi(i) < \pi(j)$ or $\pi(j) < \pi(i)$. In the first case $j \notin U$, and in the second case $i \notin U$. That is, U is an independent set of G, and the result follows. □

Corollary 11.3.8 If G is a graph with n vertices and m edges, then $\alpha(G) \geqslant \frac{n^2}{2m+n}$.

Proof With Jensen's inequality and Lemma 2.2.5, Theorem 11.3.7 yields

$$\alpha(G) \geqslant \sum_{v \in V(G)} \frac{1}{d_G(v) + 1} \geqslant \frac{n}{\frac{1}{n} \sum\limits_{v \in V(G)} d_G(i) + 1} = \frac{n^2}{2m + n}. \tag{11–1}$$

□

Corollary 11.3.9 If G is a graph on n vertices with average degree $d \geqslant 1$, then $\alpha(G) \geqslant n/(d+1)$.

Proof By (11–1) and Lemma 2.2.5,

$$\alpha(G) \geqslant \frac{n}{2m/n + 1} = \frac{n}{d+1}.$$

□

Below, we present two proofs of a rough form of Turán's theorem 4.5.4, namely Corollary 4.5.5.

Proof 1 Let G be a graph with n vertices and m edges so that

$$m > \left(1 - \frac{1}{r}\right)\frac{n^2}{2}.$$

Then the complement $\overline{G}$ of G has fewer than

$$\binom{n}{2} - \left(1 - \frac{1}{r}\right)\frac{n^2}{2}$$

edges, and hence (11–1) yields

$$\alpha\big(\overline{G}\big) > \frac{n^2}{2\left[\binom{n}{2} - \left(1 - \frac{1}{r}\right)\frac{n^2}{2}\right] + n} = r.$$

By Remark 2.1.13, $\alpha\big(\overline{G}\big) \geqslant r + 1$ implies that G contains a K_{r+1}. Therefore,

$$\operatorname{ex}(n, K_{r+1}) \leqslant \left(1 - \frac{1}{r}\right)\frac{n^2}{2}.$$

□

Proof 2 Let $G = ([n], E)$ be a K_{r+1}-free graph. Consider functions $p\colon [n] \to [0, 1]$ satisfying

$$p(1) + \cdots + p(n) = 1.$$

We seek to maximize

$$f(p) \coloneqq \sum_{\{i,j\}\in E} p(i)p(j)$$

over all such functions p. Let p be a maximizer of f, chosen so that, among all maximizers, the number of vertices k with $p(k) > 0$ is minimal.

Claim The subset $S \coloneqq \{i \in [n] : p(i) > 0\}$ is a clique of G.

Suppose not. Then there exist non-adjacent vertices i_0 and j_0 with $p(i_0), p(j_0) > 0$. Define

$$S_{i_0} = \sum_{k:\{k,i_0\}\in E} p(k) \quad \text{and} \quad S_{j_0} = \sum_{k:\{k,j_0\}\in E} p(k).$$

Without loss of generality, assume that $S_{i_0} \geqslant S_{j_0}$. Define a new function $p^*\colon [n] \to [0,1]$ by

$$p^*(k) = \begin{cases} p(i_0) + p(j_0), & \text{if } k = i_0, \\ 0, & \text{if } k = j_0, \\ p(k), & \text{otherwise.} \end{cases}$$

Then

$$\begin{aligned} f(p^*) &= \sum_{\{i,j\}\in E} p^*(i)p^*(j) \\ &= \sum_{\{i,j\}\in E} p(i)p(j) + \sum_{k:\{k,i_0\}\in E} p(k)[p^*(i_0) - p(i_0)] + \sum_{k:\{k,j_0\}\in E} p(k)[p^*(j_0) - p(j_0)] \\ &= f(p) + S_{i_0}p(j_0) - S_{j_0}p(j_0) \\ &\geqslant f(p). \end{aligned}$$

Since p is a maximizer of f, we must have $f(p^*) = f(p)$. However, p^* has fewer positive values than p, contradicting our choice of p. This proves the claim. //

Since G is K_{r+1}-free, $|S| \leqslant r$. Thus,

$$\begin{aligned} \max_p f(p) &= \sum_{i,j\in S} p(i)p(j) \\ &= \frac{1}{2}\left\{ \left(\sum_{k\in S} p(k) \right)^2 - \sum_{k\in S} p(k)^2 \right\} \\ &= \frac{1}{2}\left\{ 1 - \sum_{k\in S} p(k)^2 \right\} \\ &\leqslant \frac{1}{2}\left\{ 1 - |S| \left(\frac{1}{|S|} \sum_{k\in S} p(k) \right)^2 \right\} \\ &= \frac{1}{2}\left(1 - \frac{1}{|S|} \right) \\ &\leqslant \frac{1}{2}\left(1 - \frac{1}{r} \right). \end{aligned}$$

On the other hand, choosing $p(k) = 1/n$ for all $k \in [n]$ yields

$$\max_p f(p) \geqslant \frac{|E|}{n^2}.$$

Combining the two inequalities gives

$$|E| \leqslant \left(1 - \frac{1}{r}\right)\frac{n^2}{2}.$$

□

11.4 2-Colorable Families

Definition 11.4.1 Let $\mathcal{F}$ be a family of subsets of some finite set. We say that $\mathcal{F}$ is 2**-colorable** if there exists a coloring of the elements of the underlying set with two colors such that no member of $\mathcal{F}$ is monochromatic.

Example 11.4.2 The complete graph K_3 is not 2-colorable, since any 2-coloring of its vertices yields a monochromatic edge.

Theorem 11.4.3 Every k-uniform family with fewer than 2^{k-1} members is 2-colorable.

Proof Let $\mathcal{F}$ be an arbitrary k-uniform family of subsets of some finite set X. Consider a random 2-coloring obtained by coloring each point independently and uniformly at random with one of the two colors. Then for any fixed $A \in \mathcal{F}$, the probability that A is monochromatic is $2 \cdot 2^{-k} = 2^{1-k}$. By Fact 11.1.3, the expected number of monochromatic members of $\mathcal{F}$ is

$$\sum_{A \in \mathcal{F}} \mathbb{P}(A \text{ is monochromatic}) = |\mathcal{F}| \cdot 2^{1-k} < 1.$$

Hence there exists a 2-coloring of X with no monochromatic member of $\mathcal{F}$. □

11.5 Tournaments

Definition 11.5.1 A **tournament** is a directed graph with exactly one edge between each two vertices, in one of the two possible directions. Equivalently, a tournament is an orientation of an undirected complete graph.

Remark 11.5.2 The name tournament comes from interpreting the graph as the outcome of a round-robin tournament, a game where each player is paired against every other exactly once. In a tournament, the vertices represent the players, and the edges between players point from the winner to the loser.

Say that a tournament has the property P_k if for every set of k players there is one who beats them all, i.e., if for any k-element subset S of vertices there exists a vertex $y \notin S$ such that the edges between y and each vertex in S are directed from y to that vertex.

Theorem 11.5.3 If $\binom{n}{k}\left(1 - 2^{-k}\right)^{n-k} < 1$, then there exists a tournament of n players with property P_k.

Proof Consider a random tournament of n players, i.e., the outcome of every game is determined by a fair coin flip. For a set S of k players, let A_S be the event that no player $y \notin S$ beats all players in S. Each

$y \notin S$ has a probability of $1-2^{-k}$ of not beating all players in S, so $\mathbb{P}(A_S) = \left(1-2^{-k}\right)^{n-k}$ and

$$\mathbb{P}\left(\bigcup A_S\right) \leqslant \binom{n}{k}\left(1-2^{-k}\right)^{n-k} < 1. \tag{11–2}$$

Hence with positive probability the tournament has property P_k, and the result follows. □

Remark 11.5.4 If $k \geqslant 2$, then by Fact 1.4.8 we obtain

$$\binom{n}{k}\left(1-2^{-k}\right)^{n-k} \leqslant \frac{n^k}{k!}e^{-(n-k)/2^k} \leqslant n^k e^{-n/2^k}.$$

Thus, when $n \geqslant k^2 2^{k+1}$, the inequality (11–2) holds, and there exists a tournament of n players with property P_k.

11.6 Deletion Method

As described in previous sections, the basic probabilistic method works as follows: trying to prove that an object with certain properties exists, one defines an appropriate probability space of objects and then shows that the desired properties hold in this space with positive probability. In this section, we consider situations where the "random" object does not have all the desired properties but may have a few "blemishes". With a small alteration, we remove the blemishes, giving the desired structure.

Theorem 11.6.1 If $G = (V, E)$ is a graph on n vertices with average degree $d \geqslant 1$, then $\alpha(G) \geqslant n/2d$.

Proof Construct a random subset $S \subset V$ by including each vertex independently with probability p. Let X denote the number of vertices in S, and let Y denote the number of edges with both endpoints in S. Then $\mathbb{E}[X] = np$ and by Lemma 2.2.5 $\mathbb{E}[Y] = |E|p^2 = ndp^2/2$. Hence,

$$\mathbb{E}[X-Y] = np - \frac{ndp^2}{2}.$$

We choose $p = 1/d$ to maximize the right-hand side, yielding

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

Thus, there exists a subset S of vertices such that $X - Y \geqslant n/2d$. By removing one vertex from each edge with both endpoints in S, we obtain an independent set of size at least $X - Y \geqslant n/2d$. □

We already saw in Theorem 11.2.1 that $R(k,k) > n$ provided that $\binom{n}{k}2^{1-\binom{k}{2}} < 1$. Small alterations allow us to slightly improve this result.

Theorem 11.6.2 For any positive 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. For any set S of k vertices, let X_S be the indicator random variable for the event that the induced subgraph of K_n on S is monochromatic. Let $X = \sum_{S \in \binom{[n]}{k}} X_S$.

By Fact 11.1.3,

$$\mathbb{E}[X] = \sum_{S \in \binom{[n]}{k}} \mathbb{E}[X_S] = \sum_{S \in \binom{[n]}{k}} 2 \cdot \left(\frac{1}{2}\right)^{\binom{k}{2}} = \binom{n}{k}2^{1-\binom{k}{2}}.$$

Thus, there exists a coloring with at most $\binom{n}{k}2^{1-\binom{k}{2}}$ monochromatic copies of K_k. By removing one vertex from each such copy, we obtain a 2-edge-coloring of a complete graph on at least $n - \binom{n}{k}2^{1-\binom{k}{2}}$ vertices with no monochromatic K_k. □

Corollary 11.6.3 As $k \to \infty$, we have $R(k,k) > \frac{1}{e}[1+o(1)]k2^{k/2}$.

Proof Taking $n \sim \frac{1}{e}[1-o(1)]k2^{k/2}$ in Theorem 11.6.2 gives the desired result. □

11.7 Markov's Inequality

Theorem 11.7.1 (Markov's Inequality) If X is a nonnegative random variable and $a > 0$, then

$$\mathbb{P}(X \geqslant a) \leqslant \frac{\mathbb{E}[X]}{a}.$$

Proof $\mathbb{E}[X] = \sum_i i \cdot \mathbb{P}(X=i) \geqslant \sum_{i \geqslant a} a \cdot \mathbb{P}(X=i) = a \cdot \mathbb{P}(X \geqslant a)$. □

Corollary 11.7.2 Let $(X_n)_{n \geqslant 1}$ be a sequence of nonnegative integer-valued random variables defined on probability spaces $(\Omega_n, \mathbb{P}_n)$. If $\lim_{n\to\infty} \mathbb{E}[X_n] = 0$, then $\lim_{n\to\infty} \mathbb{P}(X_n = 0) = 1$. In other words, X_n almost surely equals 0 as $n \to \infty$.

Lemma 11.7.3 For a random graph $G(n,p)$ where $p \in (0,1)$, we have

$$\lim_{n\to\infty} \mathbb{P}\left(\alpha(G(n,p)) \leqslant \left\lceil \frac{2\ln n}{p} \right\rceil\right) = 1.$$

Here p can depend on n.

Proof Let $t = \left\lceil \frac{2\ln n}{p} \right\rceil$. We define X_n to be the number of independent sets of size $t+1$ in $G(n,p)$. For each subset of vertices $S \in \binom{[n]}{t+1}$, let X_S be the indicator random variable for the event that S is an independent set. By Fact 11.1.3 and Fact 1.4.8,

$$\begin{aligned}
\mathbb{E}[X_n] &= \sum_{S \in \binom{[n]}{t+1}} \mathbb{E}[X_S] = \sum_{S \in \binom{[n]}{t+1}} \mathbb{P}(S \text{ is an independent set}) \\
&= \sum_{S \in \binom{[n]}{t+1}} (1-p)^{\binom{t+1}{2}} = \binom{n}{t+1}(1-p)^{\binom{t+1}{2}} \\
&\leqslant \frac{n^{t+1}}{(t+1)!} e^{-p\binom{t+1}{2}} = \frac{\left(ne^{-pt/2}\right)^{t+1}}{(t+1)!} \\
&\leqslant \frac{1}{(t+1)!} \xrightarrow{n\to\infty} 0,
\end{aligned}$$

where in the last inequality we used the fact that

$$ne^{-pt/2} = n\exp\left(-\frac{p}{2}\left\lceil \frac{2\ln n}{p} \right\rceil\right) \leqslant n\exp(-\ln n) = 1.$$

By Corollary 11.7.2, we have

$$\lim_{n\to\infty} \mathbb{P}(\alpha(G(n,p)) \leqslant t) = \lim_{n\to\infty} \mathbb{P}(X_n = 0) = \lim_{n\to\infty} \mathbb{E}[X_n] = 0.$$

□

Example 11.7.4 Let $\chi(G)$ denote the chromatic number of a graph G (see Definition 9.3.3).

(1) $\chi(K_n) = n$.

(2) $\chi(C_{2n+1}) = 3$.

(3) $\chi(G) \leqslant 2$ if and only if G is bipartite.

Definition 11.7.5 The **girth** $g(G)$ of a graph G is the length of the shortest cycle in G.

Theorem 11.7.6 (Erdős' theorem) For any positive integer k, there exists a finite graph G^* such that $\chi(G^*) \geqslant k$ and $g(G^*) \geqslant k$.

Proof Consider a random graph $G = G(n,p)$, where the edge probability p will be specified later. Let $t = \left\lceil \frac{2\ln n}{p} \right\rceil$. By Lemma 11.7.3, we have $\mathbb{P}(\alpha(G) \leqslant t) \to 1$ as $n \to \infty$.

Let X_n denote the total number of cycles in G with length strictly less than k. Then

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

where $\frac{n(n-1)\cdots(n-i+1)}{2i}$ represents the number of possible cycles of length i in a complete graph K_n. We can bound this expectation as follows:

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

By Theorem 11.7.1,

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

Now, let $p = n^{-\frac{k-1}{k}}$, which implies $np = n^{1/k}$. Substituting this into the inequality, we get

$$\mathbb{P}(X_n > n/2) \leqslant \frac{2(n-1)}{n\left(n^{1/k} - 1\right)} \xrightarrow{n\to\infty} 0.$$

For sufficiently large n, there exists a graph G on n vertices satisfying

$$\frac{n^{1/k}}{6\ln n} \geqslant k, \quad X_n \leqslant n/2 \quad \text{and} \quad \alpha(G) \leqslant t = \left\lceil \frac{2\ln n}{p} \right\rceil = \left\lceil 2n^{\frac{k-1}{k}} \ln n \right\rceil \leqslant 3n^{\frac{k-1}{k}} \ln n.$$

By removing one vertex from each cycle of length less than k, we obtain a graph G^* with at least $n/2$ vertices, girth at least k. Moreover, by Theorem 9.3.4,

$$\chi(G^*) \geqslant \frac{n/2}{\alpha(G^*)} \geqslant \frac{n/2}{\alpha(G)} \geqslant \frac{n/2}{3n^{\frac{k-1}{k}} \ln n} = \frac{n^{1/k}}{6\ln n} \geqslant k.$$

□

Chapter 12 Algebraic Method

12.1 Odd–Even Town

A town of n people forms clubs in such a way that every club has an odd number of people while every two clubs have an even (possibly 0) number of people in common. What is the maximum possible number of clubs that can be formed?

Example 12.1.1 Let A_i be the set of people in the i-th club.

(1) If $A_i = \{i\}$ for $1 \leqslant i \leqslant n$, then there are n clubs.

(2) If n is even, then we can form n clubs by letting $A_i = [n] \setminus \{i\}$ for $1 \leqslant i \leqslant n$.

(3) If n is even, then we can form n clubs by letting $A_1 = [n] \setminus \{1\}$, $A_2 = [n] \setminus \{2\}$, and $A_i = \{1, 2, i\}$ for $3 \leqslant i \leqslant n$.

Theorem 12.1.2 (Odd–Even town) Let $\mathcal{F} \subset 2^{[n]}$ be a family of subsets such that for all distinct $A, B \in \mathcal{F}$, $|A|$ is odd and $|A \cap B|$ is even. Then $|\mathcal{F}| \leqslant n$.

Proof Let $\mathbb{1}_A \in \mathbb{F}_2^n$ be the characteristic vector of $A \subset [n]$. Then for all distinct $A, B \in \mathcal{F}$, we have

$$\langle \mathbb{1}_A, \mathbb{1}_A \rangle = 1 \quad \text{and} \quad \langle \mathbb{1}_A, \mathbb{1}_B \rangle = 0.$$

We claim that the vectors in $\{\mathbb{1}_A\}_{A \in \mathcal{F}}$ are linearly independent. Indeed, suppose that

$$\sum_{A \in \mathcal{F}} \lambda_A \mathbb{1}_A = 0$$

for some $\lambda_A \in \mathbb{F}_2$. Then for any $B \in \mathcal{F}$, we have

$$0 = \langle 0, \mathbb{1}_B \rangle = \left\langle \sum_{A \in \mathcal{F}} \lambda_A \mathbb{1}_A, \mathbb{1}_B \right\rangle = \sum_{A \in \mathcal{F}} \lambda_A \langle \mathbb{1}_A, \mathbb{1}_B \rangle = \lambda_B.$$

Hence, all $\lambda_A = 0$, and the vectors are linearly independent. Since they lie in $\mathbb{F}_2^n$, we have $|\mathcal{F}| \leqslant n$. □

Remark 12.1.3 We can also prove the theorem using rank arguments. Let M be the matrix whose columns are the characteristic vectors of the sets in $\mathcal{F}$. There are $|\mathcal{F}|$ columns and n rows. Then the matrix $M^\mathsf{T}M$ is an identity matrix, so $|\mathcal{F}| = \operatorname{rank}(M^\mathsf{T}M) \leqslant \operatorname{rank}(M) \leqslant n$.

12.2 Even–Odd Town

A town of n people forms clubs in such a way that every club has an even number of people while every two clubs have an odd number of people in common. What is the maximum possible number of clubs that can be formed?

Theorem 12.2.1 (Even–Odd town) Let $\mathcal{F} \subset 2^{[n]}$ be a family of subsets such that for all distinct $A, B \in \mathcal{F}$, $|A|$ is even and $|A \cap B|$ is odd. Then $|\mathcal{F}| \leqslant n$.

Proof Adding a new element to the town and including it in every club, we transform the even–odd town into an odd–even town with $n+1$ people. By Theorem 12.1.2, $|\mathcal{F}| \leqslant n+1$. Hence, it suffices to show that $|\mathcal{F}| \neq n+1$. Suppose for contradiction that $\mathcal{F} = \{A_1, \ldots, A_{n+1}\}$. For $1 \leqslant i \leqslant n+1$, let $\mathbb{1}_{A_i} \in \mathbb{F}_2^n$ be the characteristic vector of A_i. Since $|\mathcal{F}| = n+1 > \dim \mathbb{F}_2^n$, the vectors $\{\mathbb{1}_{A_i}\}_{i=1}^{n+1}$ are linearly dependent. Thus, there exist $\lambda_1, \ldots, \lambda_{n+1} \in \mathbb{F}_2$, not all zero, such that

$$\sum_{i=1}^{n+1} \lambda_i \mathbb{1}_{A_i} = 0.$$

Since for all distinct i, j, we have

$$\langle \mathbb{1}_A, \mathbb{1}_A \rangle = 0 \quad \text{and} \quad \langle \mathbb{1}_{A_i}, \mathbb{1}_{A_j} \rangle = 1,$$

we get for any $1 \leqslant k \leqslant n+1$,

$$0 = \langle 0, \mathbb{1}_{A_k} \rangle = \left\langle \sum_{i=1}^{n+1} \lambda_i \mathbb{1}_{A_i}, \mathbb{1}_{A_k} \right\rangle = \sum_{i=1}^{n+1} \lambda_i \langle \mathbb{1}_{A_i}, \mathbb{1}_{A_k} \rangle = -\lambda_k + \sum_{i=1}^{n+1} \lambda_i.$$

This implies that all λ_k are equal. Since not all λ_k are zero, we have $\lambda_k = 1$ for all k. Then the identity

$$1 = \lambda_1 = \sum_{i=1}^{n+1} \lambda_i = n+1 \quad \text{in } \mathbb{F}_2$$

shows that n is even. Now, consider the family $\mathcal{F}^c := \{A^c\}_{A \in \mathcal{F}}$, which is also an even–odd town because for all distinct $A, B \in \mathcal{F}$, we have

- $|A^c| = n - |A|$ is even, and
- $|A^c \cap B^c| = n - |A \cup B| = n - |A| - |B| + |A \cap B|$ is odd.

By the same argument as above, we know that

$$\sum_{i=1}^{n+1} \mathbb{1}_{A_i^c} = 0.$$

However, we then have

$$0 = 0 + 0 = \sum_{i=1}^{n+1} \left(\mathbb{1}_{A_i} + \mathbb{1}_{A_i^c}\right) = (n+1)\mathbf{1},$$

where $\mathbf{1} \in \mathbb{F}_2^n$ is the all-ones vector. This is a contradiction since $n+1$ is odd. □

Remark 12.2.2 We can use the rank argument to improve the bound in Theorem 12.2.1. Suppose $|\mathcal{F}| = m$ and let M be the matrix whose columns are the characteristic vectors $\mathbb{1}_{A_1}, \ldots, \mathbb{1}_{A_m}$. Then the sum of the n rows of M is 0, so $\operatorname{rank}(M) \leqslant n-1$. Moreover, $M^\mathsf{T} M$ is the $m \times m$ matrix with 0's on the diagonal and 1's elsewhere. By the matrix determinant lemma (see the proof of Theorem 7.1.9), we have

$$\det\left(M^\mathsf{T} M\right) = (-1)^{m-1}(m-1) = \begin{cases} 1, & \text{if } m \text{ is even,} \\ 0, & \text{if } m \text{ is odd.} \end{cases}$$

- ◇ If m is even, then
$$|\mathcal{F}| = m = \operatorname{rank}\left(M^\mathsf{T} M\right) \leqslant \operatorname{rank}(M) \leqslant n-1.$$
- ◇ If m is odd, then the above calculation shows that the first $m-1$ rows of $M^\mathsf{T} M$ are linearly independent, so $\operatorname{rank}\left(M^\mathsf{T} M\right) = m-1$. Hence,
$$|\mathcal{F}| = m = \operatorname{rank}\left(M^\mathsf{T} M\right) + 1 \leqslant \operatorname{rank}(M) + 1 \leqslant (n-1) + 1 = n.$$
It is not hard to see that the equality can be achieved by letting $A_i = [n] \setminus \{i\}$ for $1 \leqslant i \leqslant n$.

Theorem 12.2.3 (Even–Even Town) Let $\mathcal{F} \subset 2^{[n]}$ be a family of subsets such that for all $A, B \in \mathcal{F}$, $|A \cap B|$ is even. Then $|\mathcal{F}| \leqslant 2^{\lfloor n/2 \rfloor}$.

Proof Suppose $\mathcal{F} = \{A_1, \ldots, A_m\}$ and let $\mathbb{1}_{A_i} \in \mathbb{F}_2^n$ be the characteristic vector of A_i. Let $S \subset \mathbb{F}_2^n$ be the subspace spanned by $\{\mathbb{1}_{A_1}, \ldots, \mathbb{1}_{A_m}\}$, and let $S^\perp$ be its orthogonal complement. Since $\langle \mathbb{1}_{A_i}, \mathbb{1}_{A_j} \rangle = 0$ for all $1 \leqslant i, j \leqslant m$, we have $S \subset S^\perp$. In particular, $\dim S \leqslant \dim S^\perp$. Then by the rank–nullity theorem

$$|\mathcal{F}| = m \leqslant 2^{\dim S} \leqslant 2^{\lfloor n/2 \rfloor}.$$

□

Remark 12.2.4 The equality can be achieved by having the n people partner up (possibly leaving one unmatched person whom we cavalierly throw out in the construction) and making the clubs all $2^{\lfloor n/2 \rfloor}$ subsets of the $\lfloor n/2 \rfloor$ pairs.

12.3 Fisher's Inequality

Theorem 12.3.1 (Fisher's inequality) Let $A_1, \ldots, A_m$ be distinct subsets of $[n]$ such that $|A_i \cap A_j| = k$ for some fixed $k \in [n]$ and all $1 \leqslant i < j \leqslant m$. Then $m \leqslant n$.

Proof Let $\mathbb{1}_{A_i} \in \mathbb{R}^n$ be the characteristic vector of A_i for each i. It suffices to show that these m vectors are linearly independent. Assume for contradiction that there exist $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$, not all zero, such that

$$\sum_{i=1}^{m} \lambda_i \mathbb{1}_{A_i} = 0.$$

Since

$$\langle \mathbb{1}_{A_i}, \mathbb{1}_{A_j} \rangle = \begin{cases} |A_i|, & \text{if } i = j, \\ k, & \text{if } i \neq j, \end{cases}$$

we have

$$\begin{aligned} 0 = \left\langle \sum_{i=1}^{m} \lambda_i \mathbb{1}_{A_i}, \sum_{j=1}^{m} \lambda_j \mathbb{1}_{A_j} \right\rangle &= \sum_{i=1}^{m} \lambda_i^2 \langle \mathbb{1}_{A_i}, \mathbb{1}_{A_i} \rangle + \sum_{1 \leqslant i \neq j \leqslant m} \lambda_i \lambda_j \langle \mathbb{1}_{A_i}, \mathbb{1}_{A_j} \rangle \\ &= \sum_{i=1}^{m} \lambda_i^2 |A_i| + k \sum_{1 \leqslant i \neq j \leqslant m} \lambda_i \lambda_j = \sum_{i=1}^{m} \lambda_i^2 (|A_i| - k) + k \left(\sum_{i=1}^{m} \lambda_i \right)^2. \end{aligned}$$

Clearly, $|A_i| \geqslant k$ for all i and $|A_i| = k$ for at most one i, since otherwise the intersection condition would force two of the sets to be equal. But then the right-hand side is positive (because the last sum can vanish only if at least two of the coefficients λ_i are nonzero), a contradiction. □

Remark 12.3.2 We can also prove the theorem using rank arguments. Let M be the matrix whose columns are the characteristic vectors $\mathbb{1}_{A_1}, \ldots, \mathbb{1}_{A_m}$. Then $M^{\mathsf{T}}M$ is the $m \times m$ matrix with $|A_1|, \ldots, |A_m|$ on the diagonal and k elsewhere. For any nonzero vector $x \in \mathbb{R}^m$, we have

$$\begin{aligned} x^{\mathsf{T}}(M^{\mathsf{T}}M)x &= \begin{pmatrix} x_1 & \cdots & x_m \end{pmatrix} \left\{ k \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix} + \begin{pmatrix} |A_1| - k & & \\ & \ddots & \\ & & |A_m| - k \end{pmatrix} \right\} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \\ &= k \left(\sum_{i=1}^{m} x_i \right)^2 + \sum_{i=1}^{m} (|A_i| - k) x_i^2 > 0. \end{aligned}$$

This shows that $M^{\mathsf{T}}M$ is positive definite, so it has full rank. Therefore,

$$m = \operatorname{rank}(M^{\mathsf{T}}M) \leqslant \operatorname{rank}(M) \leqslant n.$$

Theorem 12.3.3 (de Bruijn–Erdős theorem) Suppose P is a set of n points in $\mathbb{R}^2$. Then either the points are collinear, or they determine at least n distinct lines.

Proof Let L be the set of all lines determined by the points $P = \{x_1, x_2, \ldots, x_n\}$. Our goal is to show that either $|L| = 1$ or $|L| \geqslant |P| = n$. For each point $x_i \in P$, define $L_i \subset L$ as the set of lines passing through x_i:

$$L_i = \{\ell \in L : x_i \in \ell\}.$$

We observe the following properties:

- $|L_i \cap L_j| = 1$ for all $i \neq j$.
- $L_i = L_j$ for some $i \neq j$ if and only if all n points lie on the same line.

Therefore, if the points are not collinear, then $L_i \neq L_j$ for all $i \neq j$. In this case, let $\mathcal{F} = \{L_i\}_{i=1}^{n}$. Since $|\mathcal{F}| = n$ and every pair of sets in $\mathcal{F}$ shares exactly one common element, Theorem 12.3.1 applies and gives $n \leqslant |L|$. □

Remark 12.3.4 Similar ideas can be used to prove the inequalities in Theorems 7.1.9 and 7.2.2.

We end this section with a constructive proof of a rough bound for diagonal Ramsey numbers.

Lemma 12.3.5 Let $G = (V, E)$ be a graph with $V = \binom{[k]}{3}$ and edges defined as follows: for any distinct $A, B \in V$, $AB \in E$ if and only if $|A \cap B| = 1$. Then G contains no clique or independent set of size $k+1$.

Proof Let $\{A_1, \ldots, A_m\}$ be a maximum clique in G. Then $|A_i \cap A_j| = 1$ for all $1 \leqslant i < j \leqslant m$. By Theorem 12.3.1, we have $m \leqslant k$.

Now, let $\{B_1, \ldots, B_n\}$ be a maximum independent set in G. Then $|B_i| = 3$ is odd and $|B_i \cap B_j| = 0$ or 2 is even for all $1 \leqslant i < j \leqslant n$. By Theorem 12.1.2, we have $n \leqslant k$. □

Corollary 12.3.6 $R(k+1, k+1) > \binom{k}{3}$.

12.4 Erdős Distance Problem

Theorem 12.4.1 (Unit distance problem) Given n points in $\mathbb{R}^2$, the number of unordered pairs of points at distance exactly 1 is at most $O(n^{3/2})$.

Proof Let P be a set of n points in $\mathbb{R}^2$. Define a graph G with vertex set P by joining two vertices $a, b \in P$ by an edge if and only if the distance between a and b is exactly 1. Then the number of pairs of points at distance 1 is exactly $e(G)$.

We claim that G is $K_{2,3}$-free. Indeed, suppose $a, b \in P$ are two distinct vertices. Any common neighbor of a and b must lie on the intersection of two circles of radius 1 centered at a and b. Since two distinct circles in the plane intersect in at most two points, a and b can have at most two common neighbors. Consequently, G contains no $K_{2,3}$ as a subgraph. By Corollary 4.2.4,

$$e(G) \leqslant \mathrm{ex}(n, K_{2,3}) \leqslant Cn^{2-1/2} = O(n^{3/2}).$$ □

Theorem 12.4.2 (1-distance problem) There are at most $n+1$ points in $\mathbb{R}^n$ such that the distance between any two distinct points is equal to 1.

Proof We first note that equality is attained by the vertices of an n-simplex with side length 1. Next, consider any collection of such points, say $m+1$ points in $\mathbb{R}^n$. We claim that $m \leqslant n$. By translating the configuration if necessary, we may assume that one of the points is the origin 0. Let the remaining points be denoted by $v_1, \ldots, v_m \in \mathbb{R}^n$. Then

- $\langle v_i, v_i \rangle = \|v_i - 0\|^2 = 1$ for all $1 \leqslant i \leqslant m$, and
- $\langle v_i, v_j \rangle = \frac{1}{2}\big(\|v_i\|^2 + \|v_j\|^2 - \|v_i - v_j\|^2\big) = \frac{1}{2}(1 + 1 - 1) = \frac{1}{2}$ for all $1 \leqslant i < j \leqslant m$.

Let M be the matrix whose columns are the vectors $v_1, \ldots, v_m$. Then $M^\mathsf{T} M$ is the $m \times m$ matrix with 1's on the diagonal and $\frac{1}{2}$ elsewhere. By the matrix determinant lemma (see the proof of Theorem 7.1.9), we have $\det(M^\mathsf{T} M) = (m+1)/2^m \neq 0$. Thus,

$$m = \mathrm{rank}(M^\mathsf{T} M) \leqslant \mathrm{rank}(M) \leqslant n.$$ □

When applying linear algebraic methods to combinatorics, it is often more effective to associate sets with multivariate polynomials $f(x_1, \ldots, x_n)$ rather than their standard characteristic vectors. By doing so, one can establish the size of a set system by proving that its corresponding polynomials are linearly independent within a functional vector space. Many of these applications rely on the following simple yet powerful lemma, which bridges the gap between evaluation properties and linear independence:

Lemma 12.4.3 (Independence criterion) Let Ω be an arbitrary set and $\mathbb{F}$ be any field. For $i = 1, \ldots, m$, let $f_i \colon \Omega \to \mathbb{F}$ be functions and $v_i \in \Omega$ be elements such that:

- $f_i(v_i) \neq 0$ for all $1 \leqslant i \leqslant m$, and
- $f_i(v_j) = 0$ for all $1 \leqslant j < i \leqslant m$.

Then $f_1, \ldots, f_m$ are linearly independent in the vector space $\mathbb{F}^{\Omega}$.

Proof Suppose to the contrary that there exist $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$, not all zero, such that

$$\lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_m f_m = 0.$$

Let i be the largest index such that $\lambda_i \neq 0$. Evaluating the linear relation at the point v_i, we obtain

$$\sum_{j=1}^{i} \lambda_j f_j(v_i) = 0.$$

However, by the assumptions, we have $f_j(v_i) = 0$ for all $j < i$ and $f_i(v_i) \neq 0$. Thus, we get $\lambda_i f_i(v_i) = 0$, which contradicts the choice of i. □

Definition 12.4.4 A **2-distance set** is a set of points for which the distances between any two distinct points are always one of exactly two specific values.

Example 12.4.5 (1) The unique (up to similarity) largest 2-distance set in the plane is the regular pentagon (Figure 10.1).

(2) Let $S \subset \mathbb{R}^n$ be the set of all 0-1 vectors that contain exactly two entries equal to 1. Then $|S| = \binom{n}{2}$, and the distance between any two distinct points in S is either $\sqrt{2}$ or 2.

Theorem 12.4.6 (Larman–Rogers–Seidel theorem) Every 2-distance set in $\mathbb{R}^n$ has size at most

$$\binom{n}{2} + 3n + 2 = \frac{(n+1)(n+4)}{2}.$$

Proof Let $A = \{a^{(1)}, \ldots, a^{(m)}\}$ be a 2-distance set of distinct points in $\mathbb{R}^n$. Let the two distances be d_1 and d_2. To each point $a_i \in A$, we associate the following polynomial in n real variables $x \in \mathbb{R}^n$:

$$f_i(x) \coloneqq \left(\|x - a^{(i)}\|^2 - d_1^2\right)\left(\|x - a^{(i)}\|^2 - d_2^2\right).$$

Then $f_i\left(a^{(i)}\right) = (d_1 d_2)^2 \neq 0$, but $f_i\left(a^{(j)}\right) = 0$ for all $j \neq i$. By Lemma 12.4.3, these polynomials are linearly independent in the vector space of real-valued functions on $\mathbb{R}^n$.

To find an upper bound for m, we identify a vector space of limited dimension that contains all such polynomials. Every $f_i(x)$ can be expanded into a linear combination of the following basis polynomials:

$$\left(\sum_{i=1}^{n} x_i^2\right)^2, \quad \left(\sum_{i=1}^{n} x_i^2\right) x_j, \quad x_i x_j, \quad x_i, \quad 1, \quad \text{for } i, j = 1, \ldots, n.$$

Their total number is

$$1 + n + \left[\binom{n}{2} + n\right] + n + 1 = \binom{n}{2} + 3n + 2.$$

Since the m polynomials are linearly independent, their count cannot exceed the dimension of the space they reside in. Thus, $m \leqslant \binom{n}{2} + 3n + 2$, completing the proof. □

Remark 12.4.7 The above proof can be extended to the case of s-distance sets (see Problem **??**).

We can rewrite the upper bound in Theorem 12.4.6 as

$$m \leqslant \binom{n}{2} + 3n + 2 = \binom{n+2}{2} + n + 1.$$

A significant improvement was achieved by Blokhuis in 1981, who showed that the second term $n+1$ here is redundant. His trick was to show that the polynomials $f_1(x), \ldots, f_m(x)$ together with the polynomials $x_1, \ldots, x_n, 1$ are linearly independent.

Theorem 12.4.8 (Blokhuis' theorem) Every 2-distance set in $\mathbb{R}^n$ has size at most $\binom{n+2}{2}$.

Proof With the same notation as in the proof of Theorem 12.4.6, we claim that the polynomials

$$f_1(x), \ldots, f_m(x), x_1, \ldots, x_n, 1 \tag{12–1}$$

are linearly independent. Suppose that

$$\sum_{i=1}^{m} \alpha_i \frac{f_i(x)}{d_1^2 d_2^2} + \sum_{j=1}^{n} \beta_j x_j + \gamma = 0$$

for some $\alpha_i, \beta_j, \gamma \in \mathbb{R}$. Since $f_i\big(a^{(k)}\big) = d_1^2 d_2^2 \delta_{i,k}$, evaluating the above identity at $x = a^{(k)}$ gives

$$\alpha_k + \sum_{j=1}^{n} \beta_j a_j^{(k)} + \gamma = 0. \tag{12–2}$$

Let $e^{(j)}$ be the j-th standard basis vector in $\mathbb{R}^n$. Evaluating the identity at $x = \mu e^{(j)}$ yields

$$\frac{1}{d_1^2 d_2^2} \sum_{i=1}^{m} \alpha_i \Big(\mu^2 - 2\mu a_j^{(i)} + \big\|a^{(i)}\big\|^2 - d_1^2\Big)\Big(\mu^2 - 2\mu a_j^{(i)} + \big\|a^{(i)}\big\|^2 - d_2^2\Big) + \beta_j + \gamma = 0.$$

Comparing the coefficients of μ^4 and μ^3 on both sides, we obtain

$$\sum_{i=1}^{m} \alpha_i = 0 \quad \text{and} \quad \sum_{i=1}^{m} \alpha_i a_j^{(i)} = 0.$$

Multiplying (12–2) by α_k and summing over k gives

$$\begin{aligned} 0 &= \sum_{k=1}^{m} \alpha_k^2 + \sum_{k=1}^{m} \sum_{j=1}^{n} \alpha_k \beta_j a_j^{(k)} + \gamma \sum_{k=1}^{m} \alpha_k \\ &= \sum_{k=1}^{m} \alpha_k^2 + \sum_{j=1}^{n} \beta_j \sum_{k=1}^{m} \alpha_k a_j^{(k)} + \gamma \sum_{k=1}^{m} \alpha_k \\ &= \sum_{k=1}^{m} \alpha_k^2. \end{aligned}$$

This implies that all $\alpha_k = 0$. Since $x_1, \ldots, x_n, 1$ are linearly independent, we also have all $\beta_j = 0$ and

$\gamma = 0$. Thus, the claim holds.

Finally, as in the proof of Theorem 12.4.6, the $m+n+1$ polynomials in (12–1) are linearly independent, their count cannot exceed the dimension of the space they reside in. Therefore,

$$m + n + 1 \leqslant \binom{n+2}{2} + n + 1.$$

Hence, $m \leqslant \binom{n+2}{2}$, completing the proof. □

12.5 L-Intersecting Families

Definition 12.5.1 Let $\mathcal{F} \subset 2^{[n]}$ and $L \subset \{0, 1, \ldots, n\}$. We say that $\mathcal{F}$ is an **L-intersecting family** if for all distinct $A, B \in \mathcal{F}$, $|A \cap B| \in L$.

Suppose we know only the size of L. How large can an L-intersecting family be?

- ◇ Fisher's inequality (Theorem 12.3.1) tells us that $|\mathcal{F}| \leqslant n$ when $|L| = 1$.
- ◇ The Erdős–Ko–Rado theorem (Theorem 8.1.5) tells us that if $\mathcal{F}$ is k-uniform and $L = [k-1]$ (where $2k \leqslant n$), then $|\mathcal{F}| \leqslant \binom{n-1}{k-1}$.
- ◇ In the odd–even town and even–odd town problem, we have $L = \{\text{even}\}$ or $L = \{\text{odd}\}$.

Theorem 12.5.2 (Frankl–Wilson theorem) If $\mathcal{F} \subset 2^{[n]}$ is an L-intersecting family, then $|\mathcal{F}| \leqslant \sum_{i=0}^{|L|} \binom{n}{i}$.

Remark 12.5.3 (1) This result is best possible: for $L = \{0, 1, \ldots, s-1\}$ one can take the family of all subsets of $[n]$ of size at most s.

(2) In the case of uniform families, the celebrated result of Ray-Chaudhuri and Wilson (1975) gives the upper bound $|\mathcal{F}| \leqslant \binom{n}{|L|}$ (see Problem **??**). This result is best possible as well, as shown by taking $L = \{0, 1, \ldots, s-1\}$ and $\mathcal{F} = \binom{[n]}{s}$.

Proof of Theorem 12.5.2 Let $\mathcal{F} = \{A_1, \ldots, A_m\}$ where $|A_1| \leqslant \cdots \leqslant |A_m|$. Let $L = \{l_1, \ldots, l_s\}$ be the set of all possible intersection sizes between distinct sets in $\mathcal{F}$. For each $i \in [m]$, define the polynomial f_i in n variables $x = (x_1, \ldots, x_n)$ by

$$f_i(x) = \prod_{k : l_k < |A_i|} (\langle \mathbb{1}_{A_i}, x \rangle - l_k).$$

Observe that $f_i(v_j) = 0$ for all $j < i$ and $f_i(v_i) \neq 0$ for all i. By Lemma 12.4.3, the polynomials $f_1, \ldots, f_m$ are linearly independent. To find an upper bound for m, we identify a vector space of limited dimension that contains all such polynomials. The domain being $\{0,1\}^n$ implies that $x_i^2 = x_i$ for each variable x_i. Thus, pure monomials of degree at most s form a basis (where a pure monomial is a product of distinct variables), and we have only $\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{s}$ of them. □

Using essentially the same argument, we can prove the following "modular" version of the Frankl–Wilson theorem, which can be applied to the odd–even town problem. Write $r \in L \pmod p$ if $r \equiv l \pmod p$ for at least one $l \in L$.

Theorem 12.5.4 (Deza–Frankl–Singhi) Let p be a prime and $L \subset \{0, 1, \ldots, p-1\}$. Suppose $\mathcal{F} \subset 2^{[n]}$ is a family of subsets such that

- $|A| \notin L \pmod p$ for all $A \in \mathcal{F}$, and
- $|A \cap B| \in L \pmod p$ for all distinct $A, B \in \mathcal{F}$.

Then $|\mathcal{F}| \leqslant \sum_{i=0}^{|L|} \binom{n}{i}$.

Proof Let $\mathcal{F} = \{A_1, \ldots, A_m\}$ and define the polynomial $f_i \colon \mathbb{F}_p^n \to \mathbb{F}_p$ by

$$f_i(x) = \prod_{\ell \in L} (\langle \mathbb{1}_{A_i}, x \rangle - \ell).$$

Then $f_i(v_j) = 0$ for all $j < i$ and $f_i(v_i) \neq 0$ for all i. By Lemma 12.4.3, the polynomials $f_1, \ldots, f_m$ are linearly independent. Repeating the same dimension counting argument as in the proof of Theorem 12.5.2 gives the desired bound. □

Theorem 12.5.5 (Frankl–Wilson theorem) For any prime p, there exists a graph G on $n = \binom{p^3}{p^2-1}$ vertices such that both the maximum size of a clique and the maximum size of an independent set in G are at most $\sum_{i=0}^{p-1} \binom{p^3}{i}$.

Proof Let $G = (V, E)$ be the graph with vertex set $V = \binom{[p^3]}{p^2-1}$, where two distinct vertices $A, B \in V$ are adjacent if and only if $|A \cap B| \not\equiv p - 1 \pmod p$.

Let $\{A_1, \ldots, A_m\}$ be a maximum clique in G. By the definition of adjacency in G,

- $|A_i \cap A_j| \not\equiv p - 1 \pmod p$ for all distinct i and j, and
- $|A_i| = p^2 - 1 \equiv p - 1 \pmod p$ for all i.

By Theorem 12.5.4 with $L = \{0, 1, \ldots, p-2\}$, we have $m \leqslant \sum_{i=0}^{p-1} \binom{p^3}{i}$.

Now, let $\{B_1, \ldots, B_t\}$ be a maximum independent set in G. Then $|B_i \cap B_j| \equiv p - 1 \pmod p$ for all distinct i and j. Hence, $\{B_1, \ldots, B_t\}$ is an L^*-intersecting family with

$$L^* = \{p - 1, 2p - 1, \ldots, (p-1)p - 1\}.$$

By Theorem 12.5.2, we have $t \leqslant \sum_{i=0}^{|L^*|} \binom{p^3}{i} = \sum_{i=0}^{p-1} \binom{p^3}{i}$. □

Corollary 12.5.6 $R(k+1, k+1) \geqslant k^{\Omega(\log k / \log\log k)}$. Here, $f(k) = \Omega(g(k))$ means that there exists a constant $C > 0$ such that $f(k) \geqslant Cg(k)$ for all sufficiently large k.

Proof With the same notation as in Theorem 12.5.5, we have

$$R(k+1, k+1) > n, \quad \text{where } k = \sum_{i=0}^{p-1} \binom{p^3}{i} \text{ and } n = \binom{p^3}{p^2 - 1}.$$

To simplify the relationship between n and k, we apply the following asymptotic approximations:

$$k \approx \binom{p^3}{p} \approx \left(\frac{p^3}{p}\right)^p = p^{2p}, \quad n \approx \left(\frac{p^3}{p^2-1}\right)^{p^2-1} \approx p^{p^2}.$$

Thus,

$$\log k \approx \log\left(p^{2p}\right) \approx p\log p, \quad \log\log k \approx \log(p\log p) \approx \log p,$$

which gives

$$\frac{\log k}{\log\log k} \approx \frac{p\log p}{\log p} = p.$$

Therefore,

$$n \approx \left(p^{2p}\right)^{p/2} \approx k^p = k^{\Omega(\log k/\log\log k)}.$$ □

Borsuk conjectured that it is possible to cut a d-dimensional shape of generalized diameter 1 into $d+1$ pieces each with diameter smaller than the original. It is true for $d = 2, 3$ and when the boundary is "smooth". However, the minimum number of pieces required has been shown to increase as $\sim 1.1^{\sqrt{d}}$ (Theorem 12.5.8). Since $1.1^{\sqrt{d}} > d+1$ at $d = 9162$, the conjecture becomes false at high dimensions.

Lemma 12.5.7 For any prime p, there exists a set $\mathcal{F}$ consisting of $\frac{1}{2}\binom{4p}{2p}$ vectors in $\{-1,1\}^{4p}$ such that every subset of $\mathcal{F}$ of size $2\binom{4p}{p-1}$ contains at least one pair of orthogonal vectors.

Proof Let $Q = \left\{I \in \binom{[4p]}{2p} : 1 \in I\right\}$. By symmetry, $|Q| = \frac{1}{2}\binom{4p}{2p}$. For each $I \in Q$, we define a vector $v^I \in \{-1,1\}^{4p}$ by its coordinates:

$$v_i^I = \begin{cases} 1, & \text{if } i \in I, \\ -1, & \text{if } i \notin I. \end{cases}$$

Let $\mathcal{F} = \left\{v^I : I \in Q\right\}$. Note that $|\mathcal{F}| = |Q| = \frac{1}{2}\binom{4p}{2p}$.

Claim 1 Two vectors $v^I, v^J \in \mathcal{F}$ are orthogonal if and only if $|I \cap J| \equiv 0 \pmod p$.

Since

$$\left\langle v^I, v^J \right\rangle = |I \cap J| - |I^c \cap J| - |I \cap J^c| + |I^c \cap J^c|$$

and

$$\begin{aligned} 4p &= |I \cap J| + |I \cap J^c| + |I^c \cap J| + |I^c \cap J^c| \\ &= |I \cap J| + (2p - |I \cap J|) + (2p - |I \cap J|) + |I^c \cap J^c| \\ &= 4p - |I \cap J| + |I^c \cap J^c|, \end{aligned}$$

we have

$$\left\langle v^I, v^J \right\rangle = |I \cap J| - 2(2p - |I \cap J|) + |I \cap J| = 4|I \cap J| - 4p.$$

Hence, $\left\langle v^I, v^J \right\rangle = 0$ if and only if $|I \cap J| = p$, which is equivalent to $|I \cap J| \equiv 0 \pmod p$ since $1 \leqslant |I \cap J| < 2p$ for distinct $I, J \in Q$. //

Claim 2 Any subset $\mathcal{F}' \subset \mathcal{F}$ that contains no orthogonal pairs satisfies

$$|\mathcal{F}'| \leqslant \sum_{k=0}^{p-1} \binom{4p}{k} < 2\binom{4p}{p-1}.$$

Let $Q' = \left\{I \in Q : v^I \in \mathcal{F}'\right\}$. Based on Claim 1, $Q' \subset \binom{[4p]}{2p}$ satisfies:

- $|A| = 2p \equiv 0 \pmod p$ for all $A \in Q'$.
- $|A \cap B| \not\equiv 0 \pmod p$ for all distinct $A, B \in Q'$.

By Theorem 12.5.4 with $L = [p-1]$, we obtain the first inequality. To see the second inequality, note that

$$\frac{\binom{4p}{k}}{\binom{4p}{k+1}} = \frac{k+1}{4p-k}$$

is increasing for $0 \leqslant k \leqslant p-1$. Thus, by substituting $k = p-1$, we have

$$\frac{\binom{4p}{k}}{\binom{4p}{k+1}} \leqslant \frac{(p-1)+1}{4p-(p-1)} = \frac{p}{3p+1} < \frac{1}{3}$$

for all $0 \leqslant k \leqslant p-2$. Therefore,

$$\sum_{k=0}^{p-1} \binom{4p}{k} < \binom{4p}{p-1} \sum_{k=0}^{p-1} \frac{1}{3^k} < 2\binom{4p}{p-1}. \qquad /\!/$$

The conclusion of the lemma now follows from Claim 2. □

Theorem 12.5.8 (Kahn–Kalai theorem) For a sufficiently large dimension d, there exists a bounded set $X \subset \mathbb{R}^d$ so that any partition of X into $1.1^{\sqrt{d}}$ subsets contains a subset with the same diameter as X.

Proof Let $\mathcal{F}$ be the family of vectors in $\{-1,1\}^n$ (where $n = 4p$) defined in Lemma 12.5.7. Define the set $X \subset \mathbb{R}^{n^2}$ by

$$X = \{vv^\mathsf{T} : v \in \mathcal{F}\}.$$

Let the dimension be $d = n^2 = (4p)^2 = 16p^2$. For any element $w = vv^\mathsf{T} \in X$, the squared norm is

$$\|w\|^2 = \sum_{1\leqslant i,j\leqslant n} w_{i,j}^2 = \sum_{1\leqslant i,j\leqslant n} v_i^2 v_j^2 = \sum_{i=1}^n v_i^2 \sum_{j=1}^n v_j^2 = n^2.$$

Thus, $\|w\| = n$ for all $w \in X$.

For any two elements $w = vv^\mathsf{T}$ and $w' = v'v'^\mathsf{T}$ in X, we have

$$\langle w, w'\rangle = \sum_{1\leqslant i,j\leqslant n} w_{i,j} w'_{i,j} = \sum_{1\leqslant i,j\leqslant n} v_i v_j v'_i v'_j = \sum_{i=1}^n v_i v'_i \sum_{j=1}^n v_j v'_j = \langle v, v'\rangle^2.$$

This implies that $\langle w, w'\rangle = 0$ if and only if $\langle v, v'\rangle = 0$. Thus,

$$\|w - w'\|^2 = \|w\|^2 + \|w'\|^2 - 2\langle w, w'\rangle = 2n^2 - 2\langle v, v'\rangle^2 \leqslant 2n^2.$$

This shows that $\operatorname{diam} X = \sqrt{2}n$, which occurs precisely when $\langle v, v'\rangle = 0$. Note that $|X| = |\mathcal{F}| = \frac{1}{2}\binom{4p}{2p}$.

By Lemma 12.5.7, any subset of $\mathcal{F}$ with size $2\binom{4p}{p-1}$ contains an orthogonal pair of vectors v, v'. Consequently, any $2\binom{4p}{p-1}$-subset of X must contain a pair $w = vv^\mathsf{T}$ and $w' = vv'^\mathsf{T}$ such that $\langle v, v'\rangle = 0$, yielding the maximum distance $\|w - w'\| = \operatorname{diam} X$.

To ensure that no subset in a partition contains the full diameter, each subset must have a size strictly less than $2\binom{4p}{p-1}$. Therefore, the number of subsets required is at least

$$\frac{|X|}{2\binom{4p}{p-1}} = \frac{\frac{1}{2}\binom{4p}{2p}}{2\binom{4p}{p-1}} = \frac{1}{4}\frac{(3p+1)\cdots(2p+1)}{(2p)\cdots p} > \frac{1}{4}\left(\frac{3}{2}\right)^{p+1} \geqslant C\left(\frac{3}{2}\right)^{\frac{\sqrt{d}}{4}} \geqslant 1.1^{\sqrt{d}},$$

where $d = 16p^2$ is the dimension of the space. This completes the proof. □

12.6 Bollobás' Theorem

Theorem 12.6.1 (Bollobás' theorem) Let $A_1, \dots, A_m$ and $B_1, \dots, B_m$ be two sequences of sets such that $A_i \cap B_j = \varnothing$ if and only if $i = j$. Then

$$\sum_{i=1}^{m} \binom{a_i + b_i}{a_i}^{-1} \leqslant 1,$$

where $a_i = |A_i|$ and $b_i = |B_i|$.

Proof Let $X = \bigcup_{i=1}^{m} (A_i \cup B_i)$. We argue by induction on $n = |X|$. For $n = 1$ the claim is obvious, so assume it holds for $n - 1$ and prove it for n. For every point $x \in X$, consider the family of pairs

$$\mathcal{F}_x \coloneqq \{(A_i, B_i \setminus \{x\}) : x \notin A_i\}.$$

Since each of these families $\mathcal{F}_x$ has less than n points, we can apply the induction hypothesis to obtain

$$\sum_{i \in [m],\, x \notin A_i} \binom{|A_i| + |B_i \setminus \{x\}|}{|A_i|}^{-1} \leqslant 1, \quad \forall x \in X.$$

Summing over all $x \in X$ gives

$$\sum_{x \in X} \sum_{i \in [m],\, x \notin A_i} \binom{|A_i| + |B_i \setminus \{x\}|}{|A_i|}^{-1} \leqslant n.$$

The left-hand side counts $n - a_i - b_i$ times each term $\binom{a_i+b_i}{a_i}^{-1}$, corresponding to points $x \notin A_i \cup B_i$, and b_i times each term $\binom{a_i+b_i-1}{a_i}^{-1}$, corresponding to points $x \in B_i$. Thus, interchanging the order of summation yields

$$\sum_{i=1}^{m} \left\{ (n - a_i - b_i) \binom{a_i + b_i}{a_i}^{-1} + b_i \binom{a_i + b_i - 1}{a_i}^{-1} \right\} \leqslant n. \tag{12–3}$$

Since $\binom{k-1}{\ell} = \frac{k-\ell}{k}\binom{k}{\ell}$, the i-th term of this sum simplifies to

$$(n - a_i - b_i) \binom{a_i + b_i}{a_i}^{-1} + b_i \cdot \frac{a_i + b_i}{b_i} \binom{a_i + b_i}{a_i}^{-1} = n \binom{a_i + b_i}{a_i}^{-1}.$$

Dividing both sides of (12–3) by n completes the proof. □

Remark 12.6.2 (1) The condition $A_i \cap B_j \neq \varnothing$ for all $i \neq j$ cannot be weakened to $i < j$. If this requirement is relaxed, the following counterexamples show that the conclusion of Theorem 12.6.1 may fail:

- ◇ For $m = 2$: Let $A_1 = B_2 = \{1\}$ and $A_2 = B_1 = \varnothing$. Then $\sum_{i=1}^{2} \binom{a_i + b_i}{a_i}^{-1} = 1 + 1 > 1$.

◇ For $m = 3$: Let $A_1 = B_2 = \{1\}$, $A_2 = \{2\}$, $A_3 = B_1 = \{3\}$, and $B_3 = \{1, 2\}$. Then

$$\sum_{i=1}^{3} \binom{a_i + b_i}{a_i}^{-1} = \frac{1}{2} + \frac{1}{2} + \frac{1}{3} = \frac{4}{3} > 1.$$

(2) Tuza (1984) observed that Bollobás' theorem implies both Sperner's theorem 3.3.3 and the LYM inequality 3.3.8. Let $\{A_1, \ldots, A_m\}$ be a Sperner family over a set X. Take the complements $B_i = X \setminus A_i$ and let $a_i = |A_i|$ for all i. Then $b_i = n - a_i$ and by Theorem 12.6.1,

$$\sum_{i=1}^{m} \binom{n}{|A_i|}^{-1} = \sum_{i=1}^{m} \binom{a_i + b_i}{a_i}^{-1} \leqslant 1.$$