MATH 580 Notes

These notes were taken during the Fall 2025 offering of the Combinatorics course taught by Abhishek Methuku. They contain definitions, theorems, examples, and proofs written in class, and *some additional material* not covered in lectures.

The primary reference for the course is the textbook *Combinatorial Mathematics* by Douglas B. West. The course covers material from Chapters 1–9, as well as selected topics from Chapter 10 on Ramsey Theory, and Chapter 14 on the Probabilistic Method.

These notes are also partly based on Prof. Balogh's notes from his offering of MATH 580, as found here: https://sites.google.com/view/jozsefbaloghmath/teaching/math580

The material below follows the notation and terminology of the textbook whenever applicable.

A small ritual before continuing

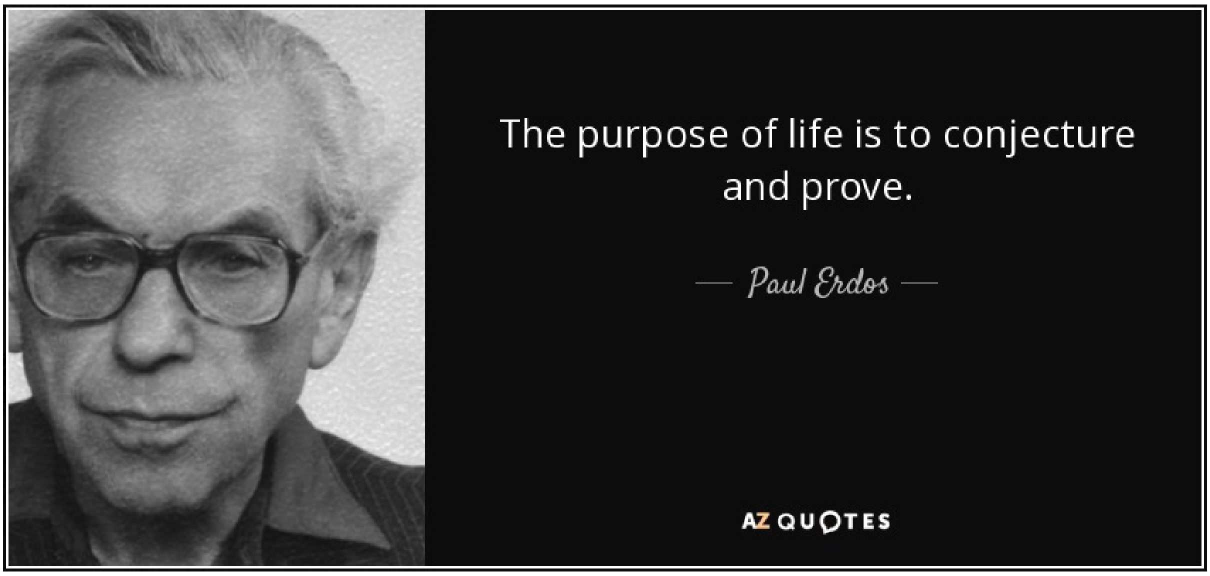
Before we proceed to learn combinatorics and graph theory,
we pause to request the mandatory traditional blessing:

May Paul Erdős — itinerant monarch, patron saint of lemmas, from whatever celestial couch he is currently borrowing, look upon these pages and not immediately close the PDF. Wherever he is currently staying, may the coffee be strong.

We ask for the following blessings:

- *taste: to prefer the sharp lemma over the bloated "generalization" nobody will use,*
- *nerve: to try the simple idea first and not hide behind machinery out of fear,*
- *precision: to keep definitions honest and hypotheses minimal (no decorative assumptions),*
- *stamina: to survive ugly computations without becoming the computation,*
- *humility: to notice when a "proof" is actually just vibes plus notation,*
- *ruthlessness: to delete a beloved argument the moment it stops pulling its weight,*
- *courage: to say "I don't know" early, before the manuscript becomes a mausoleum,*
- *and luck: that the one miraculous trick we need is the one we actually think of.*

If a lemma is ugly, may it at least be useful. If it is useless, may it at least be short. And if it is long and *useless, may it be struck from the manuscript without mercy.*



Proceed only after paying your respects.

Contents

1 Basic Combinatorics

Theorem 1.1. Let A be a finite set and suppose we can write

$$A = B_1 \,\dot\cup\, B_2 \,\dot\cup\, \cdots \,\dot\cup\, B_k$$

as a *disjoint union* (a partition) of sets $B_1, \ldots, B_k$. Then

$$|A| = |B_1| + |B_2| + \cdots + |B_k|.$$

Theorem 1.2. Let $B_1, \ldots, B_k$ be finite sets and consider their Cartesian product

$$B_1 \times B_2 \times \cdots \times B_k.$$

An element of this product is a k-tuple

$$(a_1, \ldots, a_k), \quad a_i \in B_i.$$

If there are $|B_i|$ choices for the i-th coordinate (for each i), then the total number of possible k-tuples is

$$|B_1 \times \cdots \times B_k| = |B_1| \cdot |B_2| \cdots |B_k|.$$

More generally, if $A \subseteq B_1 \times \cdots \times B_k$ and for each i the i-th coordinate can be chosen in c_i ways (possibly depending on earlier choices), then

$$|A| = c_1 c_2 \cdots c_k.$$

Goal: Find a set X that can be counted in two different ways $\Rightarrow$ we obtain an identity. Equating the two expressions for $|X|$ gives a (usually nontrivial) identity. Often, once we *guess* the identity, it could also be proved by induction, but double counting gives a more conceptual proof.

Example 1.1. We want to show

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

Consider the set

$$X := \{(a,b) : 1 \le a < b \le n\},$$

the set of all ordered pairs of distinct integers from $\{1, \ldots, n\}$ with $a < b$.

First count. Any pair $1 \le a < b \le n$ is determined by the *unordered* pair $\{a,b\}$. There are exactly $\binom{n}{2}$ unordered pairs of distinct elements of $\{1, \ldots, n\}$, and each corresponds to exactly one ordered pair (a,b) with $a < b$. Hence

$$|X| = \binom{n}{2}.$$

Second count. For each $i = 1, \dots, n-1$ define

$$B_{i+1} := \{(a, b) \in X : b = i + 1\},$$

the set of pairs whose second coordinate is $i + 1$.

Fix i. Then $b = i + 1$ and a must satisfy $1 \le a < b = i + 1$, so a can be any of $1, 2, \dots, i$. Thus $|B_{i+1}| = i$.

The sets $B_2, B_3, \dots, B_n$ form a partition of X: every pair $(a, b) \in X$ has a unique second coordinate $b \in \{2, \dots, n\}$, so it lies in exactly one B_b. Therefore, by the sum principle,

$$|X| = \sum_{i=1}^{n-1} |B_{i+1}| = \sum_{i=1}^{n-1} i.$$

Conclusion. We have counted the same set X in two ways:

$$|X| = \binom{n}{2} \quad \text{and} \quad |X| = \sum_{i=1}^{n-1} i.$$

Hence

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

Theorem 1.3 (Bijection principle)**.** If $f : A \to B$ is a bijection between finite sets, then

$$|A| = |B|.$$

Thus counting A is equivalent to counting any set B that is in bijection with A.

Example 1.2.

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

Both sides count k-element subsets of an n-element set in different ways: choosing a k-subset is equivalent to choosing its $(n-k)$-element complement.

Injections both ways imply bijection. Let A, B be finite sets. Suppose

$$f : A \to B, \qquad g : B \to A$$

are both injective (one-to-one). Then necessarily $|A| = |B|$, and hence both f and g are bijections. So, for finite sets, injections in both directions already force a bijection.

Theorem 1.4 (Pigeonhole principle)**.** If more than kn objects are placed into n boxes, then at least one box contains at least $k + 1$ objects. Equivalently: in any distribution of objects into boxes,

$$\max\{\text{occupancies of boxes}\} \;\ge\; \text{average occupancy},$$

with equality only when all boxes contain the same number of objects.

Theorem 1.5 (Polynomial principle). Let $P(x)$ and $Q(x)$ be polynomials of degree at most d over a field (e.g. $\mathbb{R}$ or $\mathbb{C}$). If

$$P(x) = Q(x)$$

for at least $d + 1$ distinct values of x, then in fact

$$P(x) \equiv Q(x)$$

as polynomials (all coefficients are equal).

Definition 1.1 (k-word). Fix a finite set A (the *alphabet*). A *k-word over A* is an ordered list of length k of elements of A, i.e. an element of A^k.

Definition 1.2 (Simple word). A *simple k-word* (or *word with distinct letters*) is a k-word in which no letter repeats. Equivalently, it is an ordered k-tuple $(a_1, \dots, a_k)$ with $a_i \in A$ and $a_i \neq a_j$ for $i \neq j$.

Definition 1.3 (k-set). A *k-set* from A is a k-element subset of A (order does not matter, no repetition).

We use the standard shorthand

$$[n] := \{1, 2, \dots, n\}.$$

Then

$$\binom{[n]}{k}$$

denotes the collection of all k-subsets of $[n]$, and

$$\left|\binom{[n]}{k}\right| = \binom{n}{k}$$

is the number of k-element subsets of an n-element set.

We can classify size–k selections from an n-element set according to whether order matters and whether repetitions are allowed:

	no repetitions	repetitions allowed
ordered	simple k-words / k-permutations	k-words
unordered	subsets of size k	multisets of size k

A (possibly repeated) word of length k from $[n]$ is just an ordered k-tuple

$$(a_1, \dots, a_k) \in [n]^k.$$

By the product principle, each coordinate has n choices independently, so

$$\#\{\text{words of length } k \text{ from } [n]\} = n^k.$$

A *simple k-word* is a word of length k with all entries distinct. To count them, choose the letters one by one:

- position 1: n choices;
- position 2: $n-1$ choices;
- ...
- position k: $n-k+1$ choices.

By the product principle,

$$\#\{\text{simple } k\text{-words from } [n]\} = n(n-1)\cdots(n-k+1) = \prod_{i=0}^{k-1}(n-i).$$

It is convenient to introduce shorthand notation.

Definition 1.4 (Falling and rising factorials). For integers $k \geq 0$ and n we define the *falling factorial*

$$n^{\underline{k}} := n(n-1)(n-2)\cdots(n-k+1) = \prod_{i=0}^{k-1}(n-i),$$

and the *rising factorial*

$$n^{\overline{k}} := n(n+1)(n+2)\cdots(n+k-1) = \prod_{i=0}^{k-1}(n+i).$$

When $k=0$ we use the empty product convention:

$$n^{\underline{0}} = n^{\overline{0}} = 1.$$

With this notation,

$$\#\{\text{simple } k\text{-words from } [n]\} = n^{\underline{k}}.$$

Note that

$$n! = n^{\underline{n}}, \qquad n^{\overline{k}} = (n+k-1)^{\underline{k}}.$$

Each simple k-word from $[n]$ corresponds to a k-element subset of $[n]$ together with an ordering of its elements. Conversely, given any k-subset, there are exactly $k!$ ways to order its elements. Thus we can obtain the number of k-subsets of $[n]$ by either:

- *forgetting order*: many different simple words represent the same subset, or
- by the *division principle*: divide by $k!$.

Hence

$$\binom{n}{k} = \frac{\#\{\text{simple } k\text{-words from } [n]\}}{k!} = \frac{n^{\underline{k}}}{k!}.$$

Definition 1.5 (Binomial coefficient). For integers $n \geq 0$ and $0 \leq k \leq n$, the *binomial coefficient*

$$\binom{n}{k}$$

denotes the number of ways to choose k elements from an n-element set. Equivalently, it is the number of k-subsets of $[n] = \{1, \ldots, n\}$.

2 Binomial coefficients

Definition 2.1 (Binomial coefficient)**.** For integers $n \geq 0$ and $0 \leq k \leq n$, the *binomial coefficient*

$$\binom{n}{k}$$

denotes the number of ways to choose k elements from an n-element set. Equivalently, it is the number of k-subsets of $[n] = \{1, \dots, n\}$.

2.1 Binomial coefficients

Theorem 2.1 (Binomial theorem)**.** For every integer $n \geq 0$ and for all real (or complex) numbers x, y,

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

Proof. Expand $(x + y)^n$ as

$$(x + y)(x + y) \cdots (x + y)$$

(n factors). To get a monomial $x^k y^{n-k}$ in the expansion, we must choose x from exactly k of the factors and y from the remaining $n - k$ factors.

The number of ways to choose which k factors contribute x is $\binom{n}{k}$ (choose the set of positions where we pick x). Thus the coefficient of $x^k y^{n-k}$ is $\binom{n}{k}$, which yields the stated identity. □

2.2 Multisets, Stars and Bars

Definition 2.2 (Multiset)**.** A k-element multiset from $[n]$ is a multiset whose underlying set is a subset of $[n]$ and whose total multiplicity (counting repetitions) is k. Equivalently, it is a sequence of multiplicities

$$(x_1, \dots, x_n) \in \mathbb{Z}_{\geq 0}^n$$

such that

$$x_1 + \cdots + x_n = k,$$

where x_i is the multiplicity of i in the multiset.

Thus we have a bijection:

$$\{k\text{-element multisets from } [n]\} \longleftrightarrow \{(x_1, \dots, x_n) \in \mathbb{Z}_{\geq 0}^n : x_1 + \cdots + x_n = k\}.$$

Theorem 2.2 (Stars and Bars)**.** The number of k-element multisets from $[n]$ equals

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

Equivalently, the number of n-tuples $(x_1, \dots, x_n)$ of nonnegative integers with $\sum_{i=1}^{n} x_i = k$ is $\binom{n+k-1}{k}$.

Proof. Consider a solution $(x_1, \dots, x_n)$ with $x_i \geq 0$ and $x_1 + \cdots + x_n = k$. Write a string of k dots ("stars") and $n-1$ bars:

$$\underbrace{\bullet\ \bullet\ \dots\ \bullet}_{x_1} \mid \underbrace{\bullet\ \dots\ \bullet}_{x_2} \mid \dots \mid \underbrace{\bullet\ \dots\ \bullet}_{x_n}.$$

The number of stars before the first bar is x_1, between the first and second bar is x_2, etc., and after the last bar is x_n.

Conversely, given any string of k stars and $n-1$ bars, reading the numbers of stars in each segment between consecutive bars recovers a unique n-tuple $(x_1, \dots, x_n)$ with sum k. Thus we have a bijection between such n-tuples and strings of length $k+n-1$ with k stars and $n-1$ bars.

The number of such strings is

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

since we just choose the positions of the k stars (or of the $n-1$ bars). This is the desired count. □

Definition 2.3 (Composition). A *composition* of a positive integer k into n parts is an ordered n-tuple $(y_1, \dots, y_n)$ of positive integers such that

$$y_1 + \cdots + y_n = k.$$

Corollary 2.3. The number of compositions of k with n parts is

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

Proof. Write $y_i = x_i + 1$ where $x_i \geq 0$. Then

$$y_1 + \cdots + y_n = k \quad \Longleftrightarrow \quad x_1 + \cdots + x_n = k - n.$$

So compositions of k with n positive parts correspond bijectively to solutions of $x_1 + \cdots + x_n = k - n$ with $x_i \geq 0$. By the previous theorem, the number of such solutions is

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

□

We can view the following as correspondences:

- k-words over $S \longleftrightarrow$ functions $f\colon [k] \to S$;
- subsets of $S \longleftrightarrow$ indicator functions $f\colon S \to \{0,1\}$;
- multisets from $S \longleftrightarrow$ multiplicity functions $f\colon S \to \mathbb{N}$ with a fixed total sum;

2.3 Binomial Identities

A *double counting* proof establishes an identity by counting the *same finite set* of objects in two different ways.

More precisely, suppose we want to prove an equality

$$\text{LHS} = \text{RHS}.$$

We look for a concrete finite set Ω such that:

- the left-hand side LHS counts $|\Omega|$ by one natural method (e.g. by choosing parameters, splitting into cases, or summing over a statistic), and
- the right-hand side RHS counts $|\Omega|$ by a different method.

Since both expressions count the same quantity $|\Omega|$, they must be equal.

Goal in practice: When we want to prove an identity (especially involving binomial coefficients) using *double counting*, our job is to *invent* a set Ω so that each side becomes an honest count of Ω under a different viewpoint.

Theorem 2.4 (Pascal)**.** For $1 \le k \le n$,

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

Proof. Interpret $\binom{n}{k}$ as the number of k-subsets of $[n]$.

Partition all k-subsets of $[n]$ into two classes:

- those that do *not* contain n;
- those that *do* contain n.

The first class is in bijection with k-subsets of $[n-1]$ (we just ignore n), so there are $\binom{n-1}{k}$ of them. The second class is in bijection with $(k-1)$-subsets of $[n-1]$: if a k-subset contains n, the remaining $k-1$ elements lie in $[n-1]$.

Hence the total number of k-subsets is

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

which proves the formula. □

Theorem 2.5. For every $n \ge 0$,

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

Proof. Right-hand side: 2^n is the number of subsets of $[n]$, since each of the n elements may be either in or out independently.

Left-hand side: group all subsets of $[n]$ by their size. For each k, there are $\binom{n}{k}$ subsets of size k. By the sum principle,

$$\#\{\text{subsets of } [n]\} = \sum_{k=0}^{n} \binom{n}{k}.$$

Equating both counts gives the identity. □

Fix integers $n \geq r \geq 0$.

Theorem 2.6 (Hockey-stick identity).

$$\sum_{k=r}^{n} \binom{k}{r} = \binom{n+1}{r+1}.$$

Proof. Right-hand side: $\binom{n+1}{r+1}$ counts $(r+1)$-subsets of $[n+1] = \{1, \dots, n+1\}$.
Left-hand side: partition all $(r+1)$-subsets of $[n+1]$ by their largest element. For $k = r, \dots, n$, consider the group of subsets whose largest element is exactly $k+1$. Such a subset must contain $k+1$ and choose the remaining r elements from $\{1, \dots, k\}$, so there are $\binom{k}{r}$ of them.
Thus

$$\#\{(r+1)\text{-subsets of } [n+1]\} = \sum_{k=r}^{n} \binom{k}{r}.$$

Equating the two expressions for this count gives the identity. □

Theorem 2.7. For each fixed $d \geq 0$, the functions $f_0, f_1, \dots f_d$ where $f_i(k) = \binom{k}{i}$ form a basis of the real vector space $\mathcal{P}_d$ of polynomials in k of degree at most d.

Proof. First recall that $\mathcal{P}_d$ has dimension $d+1$, since $\{1, k, k^2, \dots, k^d\}$ are linearly independent.
Evaluate the binomial polynomials at the integer points $0, 1, \dots, d$. We use that

$$\binom{m}{j} = \begin{cases} 0, & m < j, \\ 1, & m = j. \end{cases}$$

Form the $(d+1) \times (d+1)$ matrix $A = (a_{mj})$ with $a_{mj} = \binom{m}{j}$, where $m, j = 0, 1, \dots, d$. Then A is lower–triangular with all diagonal entries $a_{jj} = \binom{j}{j} = 1$, so $\det A = 1 \neq 0$. Hence the functions $\binom{k}{0}, \dots, \binom{k}{d}$ are linearly independent.
Since we have $d+1$ linearly independent vectors in a vector space of dimension $d+1$, they form a basis of $\mathcal{P}_d$. □

Example 2.1. We use the above theorem to prove that

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

Remark 2.1. This can be proved easily by induction if we have a conjecture for what the expression would be.

Because $\{\binom{k}{0}, \binom{k}{1}, \binom{k}{2}\}$ is a basis of $\mathcal{P}_2$, we can write

$$k^2 = a_0 \binom{k}{0} + a_1 \binom{k}{1} + a_2 \binom{k}{2}.$$

Comparing coefficients (or solving at $k = 0, 1, 2$) gives $a_0 = 0$, $a_1 = 1$, $a_2 = 2$, so

$$k^2 = 2\binom{k}{2} + \binom{k}{1}.$$

Summing this identity over $k = 1, \dots, n$ and using $\sum_k \binom{k}{1} = \binom{n+1}{2}$ and $\sum_k \binom{k}{2} = \binom{n+1}{3}$ via hockey-stick identity yields a combinatorial proof of the well-known formula

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

This can be generalized to find expressions of the form $\sum_{k=1}^{n} P(k)$ for some polynomial $P(k)$

Theorem 2.8 (Vandermonde, 1772). For nonnegative integers m, n, r,

$$\sum_k \binom{m}{k}\binom{n}{r-k} = \binom{m+n}{r},$$

where the sum is over all integers k (terms with impossible parameters are interpreted as 0). In particular, for $m = n$ and $r = n$,

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

Proof. Right-hand side: $\binom{m+n}{r}$ is the number of r-subsets of the set $[m+n] = \{1, \dots, m+n\}$. Interpret $[m+n]$ as the disjoint union of two blocks

$$A = \{1, \dots, m\}, \qquad B = \{m+1, \dots, m+n\}.$$

Any r-subset S of $A \cup B$ has some number k of elements from A and $r-k$ from B, where $0 \le k \le r$.

For a fixed k, the number of such subsets S with $|S \cap A| = k$ is

$$\binom{m}{k}\binom{n}{r-k}$$

(choose k elements from A and $r-k$ from B). Summing over all k gives the left-hand side, which must equal the total number of r-subsets, namely $\binom{m+n}{r}$. □

Definition 2.4 (Extended binomial coefficient). Let $u \in \mathbb{R}$ (or $\mathbb{C}$) and $k \in \mathbb{N}$. Define

$$\binom{u}{k} := \frac{1}{k!} u(u-1)\cdots(u-k+1)$$

where the product is empty and equal to 1 when $k = 0$. If k is not a nonnegative integer, we set $\binom{u}{k} = 0$.

Example 2.2. For $u \in \mathbb{R}$ and $k \geq 0$,

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

Example 2.3. Take $u = 3$ and $k = 5$.

$$\binom{-3}{5} = \frac{(-3)(-4)(-5)(-6)(-7)}{5!}.$$

There are 5 negative factors in the numerator, so the numerator has sign $(-1)^5 = -1$, and

$$(-3)(-4)(-5)(-6)(-7) = (-1)^5 (3 \cdot 4 \cdot 5 \cdot 6 \cdot 7).$$

Thus

$$\binom{-3}{5} = (-1)^5 \frac{3 \cdot 4 \cdot 5 \cdot 6 \cdot 7}{5!} = (-1)^5 \binom{7}{5} = (-1)^5 \binom{3+5-1}{5},$$

Theorem 2.9 (Newton's generalized binomial theorem)**.** Let $u \in \mathbb{R}$ and $x \in \mathbb{R}$ with $|x| < 1$. Then

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

For integer $u = n \geq 0$ this reduces to the ordinary binomial theorem, since $\binom{n}{k} = 0$ for $k > n$ and the sum is finite.

Proof. Consider the function

$$f(x) := (1+x)^u$$

on $(-1, 1)$. We first compute its derivatives. By the chain rule,

$$f'(x) = u(1+x)^{u-1}.$$

Differentiating repeatedly and using induction on k gives

$$f^{(k)}(x) = u(u-1)\cdots(u-k+1)\,(1+x)^{u-k} \quad \text{for all } k \geq 0.$$

(For $k = 0$ this is the definition of f; assuming the formula for k, differentiating once more yields the formula for $k+1$.)

In particular, evaluating at $x = 0$ we obtain

$$f^{(k)}(0) = u(u-1)\cdots(u-k+1) = k! \binom{u}{k}.$$

Now recall the Taylor expansion of a C^∞ function around 0: if f is analytic on $(-1, 1)$, then for $|x| < 1$,

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

and this series converges to $f(x)$. Our function $f(x) = (1+x)^u$ is analytic on $(-1, 1)$. Substituting $f^{(k)}(0) = k!\binom{u}{k}$ into the Taylor series gives

$$(1+x)^u = \sum_{k=0}^{\infty} \frac{k!\binom{u}{k}}{k!} x^k = \sum_{k=0}^{\infty} \binom{u}{k} x^k,$$

for all $|x| < 1$, as claimed. □

3 Combinatorial arguments

3.1 Delannoy numbers

Definition 3.1. For integers $m, n \geq 0$, the *Delannoy number* $d_{m,n}$ is the number of lattice paths from $(0,0)$ to (m,n) using only the three types of steps

$$(1,0), \quad (0,1), \quad (1,1).$$

Example 3.1. Compute $d_{2,2}$.

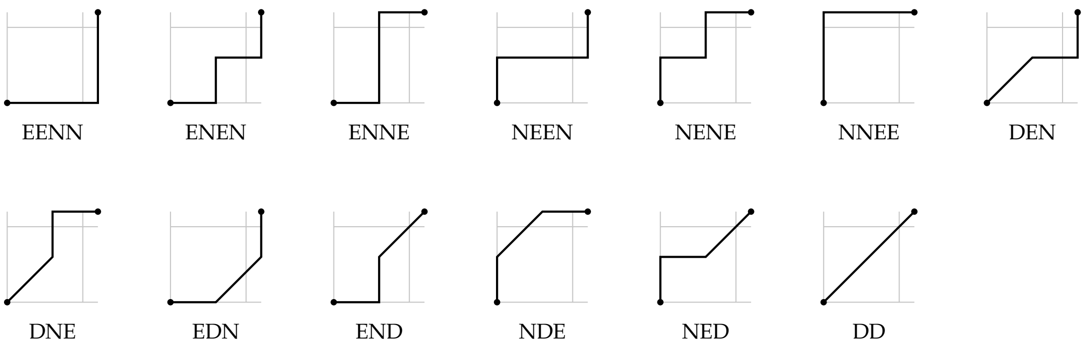
- Paths using only $(1,0)$ and $(0,1)$: we need 2 horizontal and 2 vertical steps, in any order:

$$\binom{4}{2} = 6.$$

- Paths using two diagonal steps $(1,1)$: the only possibility is $(1,1),(1,1)$, so this contributes 1.
- Paths using exactly one diagonal step: then we need 1 horizontal step and 1 vertical step in addition; there are $3!$ permutations of the multiset $\{(1,1),(1,0),(0,1)\}$, so this contributes 6.

Thus

$$d_{2,2} = 6 + 1 + 6 = 13.$$



Theorem 3.1. For all $m, n \geq 0$,

$$d_{m,n} = \sum_j \binom{m}{j}\binom{n+m-j}{m} = \sum_k \binom{m}{k}\binom{n+k}{m},$$

where the sums range over integers for which the binomial coefficients are defined, and $k = m - j$.

Proof. Partition all paths from $(0,0)$ to (m,n) according to the number j of diagonal steps $(1,1)$ they use.

Fix j. Then:

- The path uses $m - j$ horizontal steps $(1,0)$ and $n - j$ vertical steps $(0,1)$, since the total x–increment is m and the total y–increment is n.
- The total number of steps is

$$(m-j) + (n-j) + j = m + n - j.$$

Among these steps, exactly m of them increase the x–coordinate: the $(m-j)$ horizontal steps and the j diagonal steps. Choosing the positions of these m steps among the $m+n-j$ total steps gives

$$\binom{m+n-j}{m}$$

possibilities.

- Among those m x–increasing steps, we must decide which j are diagonal and which $m-j$ are horizontal. This can be done in

$$\binom{m}{j}$$

ways.

Therefore, for fixed j, the number of paths with exactly j diagonal steps is

$$\binom{m}{j}\binom{m+n-j}{m}.$$

Summing over all admissible j gives

$$d_{m,n} = \sum_{j}\binom{m}{j}\binom{m+n-j}{m}.$$

Finally, substituting $k = m - j$ yields the equivalent form

$$d_{m,n} = \sum_{k}\binom{m}{k}\binom{n+k}{m}.$$

□

3.2 Lattice balls in $\mathbb{Z}^n$

Definition 3.2 (Lattice ball). Fix integers $n \geq 1$ and $m \geq 0$. The *lattice ball* of radius m in $\mathbb{Z}^n$ is

$$B_m^{(n)} := \left\{ x = (x_1, \ldots, x_n) \in \mathbb{Z}^n : |x_1| + \cdots + |x_n| \leq m \right\}.$$

Equivalently, $B_m^{(n)}$ is the set of lattice points that can be reached from $0 = (0, \ldots, 0)$ in at most m steps of the form $\pm e_i$ (where e_i are the standard basis vectors).

Example 3.2. In $\mathbb{Z}^2$ with radius $m = 2$, the ball consists of the 13 points

$$(0,0),\ (\pm 1, 0),\ (0, \pm 1),\ (\pm 2, 0),\ (0, \pm 2),\ (\pm 1, \pm 1).$$

Theorem 3.2. For integers $n \geq 1$ and $m \geq 0$, the size of the lattice ball is

$$|B_m^{(n)}| = \sum_{k=0}^{\min\{n,m\}} \binom{n}{k}\binom{m}{k} 2^k.$$

Proof. Counting lattice points in $B_m^{(n)}$ is the same as counting integer solutions of

$$|x_1| + \cdots + |x_n| \le m.$$

We group solutions according to the number k of nonzero coordinates.

Step 1: choose which coordinates are nonzero. If a solution has exactly k nonzero coordinates, there are

$$\binom{n}{k}$$

ways to choose their positions.

Step 2: choose signs. Once the positions of the k nonzero coordinates are fixed, each of those coordinates can be positive or negative independently, giving

$$2^k$$

choices of signs.

Thus it remains to count, for fixed k, the number of solutions with *positive* values in those k chosen coordinates.

Step 3: positive solutions. Let $y_1, \ldots, y_k$ be the absolute values of the k nonzero coordinates; these are integers with $y_i > 0$ and

$$y_1 + \cdots + y_k \le m.$$

The number of such k-tuples is the same as the number of nonnegative integer solutions to a single equality, via a "slack variable" trick.

First shift $y_i = x_i + 1$, where $x_i \ge 0$. Then

$$y_1 + \cdots + y_k \le m \iff (x_1 + 1) + \cdots + (x_k + 1) \le m \iff x_1 + \cdots + x_k \le m - k.$$

Introduce one extra nonnegative variable x_{k+1} and write

$$x_1 + \cdots + x_k + x_{k+1} = m - k, \qquad x_i \ge 0.$$

By stars and bars, the number of nonnegative integer solutions is

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

Hence, for fixed k, there are $\binom{m}{k}$ ways to choose the absolute values of the k nonzero coordinates.

Step 4: combine the choices. For a fixed k, we have

$$\binom{n}{k} \cdot 2^k \cdot \binom{m}{k}$$

solutions with exactly k nonzero coordinates. Summing over all admissible k (i.e. $0 \le k \le \min\{n, m\}$) gives

$$|B_m^{(n)}| = \sum_{k=0}^{\min\{n,m\}} \binom{n}{k}\binom{m}{k} 2^k,$$

as claimed. □

3.3 Delannoy identity

Theorem 3.3 (Delannoy identity)**.** For all $m, n \in \mathbb{N}$ we have

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

Proof. Recall our two combinatorial interpretations:

- Let A be the set of Delannoy paths from $(0,0)$ to (m,n). $|A| = d_{m,n} = \sum_k \binom{m}{k}\binom{n+k}{m}$.
- Let B be the lattice ball of radius m in $\mathbb{Z}^n$, so $|B| = \sum_k \binom{n}{k}\binom{m}{k} 2^k$.

We will demonstrate a bijection $A \leftrightarrow B$ between Delannoy paths and lattice balls.

Map $\phi : A \to B$ (Delannoy path → lattice ball point). Take a path $P \in A$. For each $i \in \{1, \dots, n\}$ look at the portion of P between the horizontal lines $y = i - 1$ and $y = i$.

Because all steps have vertical component 0 or 1, the path crosses from $y = i - 1$ to $y = i$ exactly once. Before that crossing it may use some horizontal steps H at height $y = i - 1$; the crossing itself is precisely one step, either V or D. Thus the segment from height $i - 1$ to height i has the form

$$H\,H\,\dots\,H\,U,$$

where U is either V or D.

Let h_i be the number of those H's at height $y = i - 1$. Define

$$b_i := \begin{cases} h_i, & \text{if } S_i = V, \\ -(h_i + 1), & \text{if } S_i = D. \end{cases}$$

Intuitively $|b_i|$ is the number of steps that increase the x–coordinate between $y = i - 1$ and $y = i$ (the H's and possibly one D), and the sign records whether the step that actually raises the y–coordinate is vertical (+) or diagonal (−).

Doing this for each $i = 1, \dots, n$ gives a vector $b = (b_1, \dots, b_n) \in \mathbb{Z}^n$.

The total increase in the x–coordinate while the path rises from $y = 0$ to $y = n$ is therefore

$$\sum_{i=1}^{n} |b_i|.$$

After the path reaches height $y = n$, it may take additional horizontal steps H at height n; suppose there are r such steps. Then the endpoint (m, n) satisfies

$$m = \left(\sum_{i=1}^{n} |b_i|\right) + r, \qquad r \geq 0,$$

so $\sum_i |b_i| \leq m$. Hence $b \in B$, and we set $\phi(P) = b$.

Map $\psi : B \to A$ (Lattice ball point → Delannoy path). Now take $b = (b_1, \dots, b_n) \in B$ with

$$\sum_{i=1}^{n} |b_i| \leq m.$$

We construct a path $\psi(b)$ from $(0,0)$ to (m,n).

Start at $(0,0)$ and for $i=1,\dots,n$ repeat the following step.

Assume we currently are at $(x, i-1)$.

- If $b_i \geq 0$: take b_i horizontal steps H at height $y = i-1$, then one vertical step V to reach $(x+b_i, i)$.
- If $b_i < 0$: let $t := |b_i| = -b_i > 0$. Take $t-1$ horizontal steps H at height $y = i-1$, then one diagonal step D to reach $(x+t, i)$.

In either case, between heights $i-1$ and i we use exactly $|b_i|$ steps that increase the x–coordinate (all the H's plus possibly one D), so after finishing the nth layer we are at

$$\Big(\sum_{i=1}^{n} |b_i|,\ n\Big).$$

Finally, add

$$m - \sum_{i=1}^{n} |b_i| \ \geq 0$$

additional horizontal steps H at height $y = n$. The resulting path ends at (m,n) and uses only H, V, D steps, so $\psi(b) \in A$. □

3.4 Cayley's Formula

Definition 3.3 (Graph). Let V be a finite set. A (simple) graph on V is a pair $G = (V, E)$ with

$$E \subseteq \binom{V}{2}$$

(the edges are unordered pairs of distinct vertices).

On the fixed vertex set $V = [n] = \{1, 2, \ldots, n\}$, each of the $\binom{n}{2}$ possible edges is either present or absent, independently. Hence

$$\#\{\text{graphs on } [n]\} = 2^{\binom{n}{2}}.$$

Definition 3.4 (Tree). A *tree* is a connected, acyclic graph.

How many trees on $[n]$?

- Any n-vertex tree has exactly $n-1$ edges. Therefore the number of labelled trees on $[n]$ is at most the number of $(n-1)$-edge subsets of the edge set of K_n:

  $$\#\{\text{trees on } [n]\} \le \binom{\binom{n}{2}}{n-1}.$$

  Using the crude bounds

  $$\binom{n}{2} \le \frac{n^2}{2}, \qquad \binom{a}{b} \le \left(\frac{ea}{b}\right)^b \quad (a \ge b),$$

  we obtain

  $$\#\{\text{trees on } [n]\} \le \binom{\frac{n^2}{2}}{n} \le \left(\frac{e\,(n^2/2)}{n}\right)^n = \left(\frac{e}{2}\right)^n n^n.$$

  So up to the multiplicative factor $\left(\frac{e}{2}\right)^n$, the upper bound behaves like n^n.

- Every path on $[n]$ is a tree. A labelled path on $[n]$ is determined by an ordering of the vertices up to reversal, so the number of labelled paths is

  $$\frac{n!}{2}.$$

  Thus

  $$\#\{\text{trees on } [n]\} \ge \frac{n!}{2} \approx \left(\frac{n}{e}\right)^n$$

  (by Stirling's formula).

So the true number of labelled n-vertex trees lies between about $(n/e)^n$ and $\left(\frac{e}{2}\right)^n n^n$. Cayley's formula (1889) gives the exact value:

$$\#\{\text{labelled trees on } [n]\} = n^{n-2}.$$

Definition 3.5 (Functional Digraph). Given a function $f : [n] \to [n]$, define its *functional digraph* D_f by

$$V(D_f) = [n], \qquad E(D_f) = \{x \to f(x) : x \in [n]\}.$$

Every vertex has out-degree exactly 1.

Each weakly connected component of D_f contains exactly one directed cycle, and every other vertex in that component lies on a (directed) tree whose edges are oriented towards that cycle.

Theorem 3.4 (Cayley). The number of labelled trees on the vertex set $[n]$ is n^{n-2}.

We will follow Joyal's "functional digraph" proof.

Strategy (The Roadmap): We are establishing a bijection between the set of functions $f : [n] \to [n]$ subject to the constraints $f(1) = 1$ and $f(n) = n$ (a set of size n^{n-2}) and the set of all *labelled trees* on $[n]$ (which we want to prove has size n^{n-2}).

The transformation proceeds in three phases:

1. **Digraph Interpretation:** We view f as a directed graph. Since $f(1) = 1$ and $f(n) = n$, the vertices 1 and n are "fixed points" (loops). Any other cycles in the graph are floating components.
2. **Cycle Sorting:** We identify all cycles in the graph (including the loops at 1 and n). We define a canonical ordering for these cycles based on their smallest elements.
3. **Building the Spine:** We "cut" these cycles and stitch them together in that specific sorted order to form a unique simple path from 1 to n. This path becomes the "spine" of the tree, and all non-cycle vertices hang off this spine as subtrees, resulting in a valid tree structure.

Proof. There are n^n functions $f : [n] \to [n]$. We want n^{n-2} trees, so we will fix two values. Consider all functions $f : [n] \to [n]$ with $f(1) = 1,\ f(n) = n.$. We will construct a bijection to all labelled trees on $[n]$. Let f be such a function and consider its functional digraph D_f.

Step 1: Each component of D_f has a unique directed cycle. List these cycles: $C_1, \dots, C_r$

Step 2: For each cycle C_i, choose a cyclic order

$$C_i = (v_{i,1} \to v_{i,2} \to \cdots \to v_{i,\ell_i} \to v_{i,1})$$

and let m_i be the smallest label on that cycle. Rotate the notation so that m_i is written *last*:

$$C_i = (c_{i,1} \to c_{i,2} \to \cdots \to c_{i,\ell_i - 1} \to m_i \to c_{i,1}).$$

Thus the edge leaving m_i is $m_i \to c_{i,1}$. Because $f(1) = 1$ and $f(n) = n$, the vertices 1 and n are 1-cycles, so there are cycles $\{1\}$ and $\{n\}$.

Step 3: Now *order* the cycles so that $m_1 < m_2 < \cdots < m_r$, i.e. the smallest label is increasing. In this order we necessarily have $m_1 = 1$ and $m_r = n$.

Step 4: Forget the orientations of the edges and turn D_f into an undirected graph by ignoring arrow directions. This has n vertices and n edges. We now modify the edges lying on the cycles, keeping all tree-edges (the edges not belonging to any cycle) as they are. For each i with $1 \le i \le r - 1$:

- delete the edge m_i—$c_{i,1}$ from C_i;
- add a new edge m_i—$c_{i+1,1}$ connecting cycle C_i to the next cycle C_{i+1}.

For the last cycle C_r we simply *delete* the edge $m_r c_{r,1}$ (which is the loop from n to n).

In the above procedure, for each $i = 1, \dots, r - 1$ we remove one edge and add one edge, and for $i = r$ we remove one edge and add none. Thus the total number of edges decreases by exactly 1, so the graph has $n - 1$ edges. Each component of D_f originally contained one cycle; by inserting edges m_i—$c_{i+1,1}$ we link the cycles (and hence their attached trees) into a single connected component.

How to reverse reverse the construction:

- Given a tree T on $[n]$, there is a unique simple path from 1 to n. Write it as

$$P: \quad v_0 = 1, v_1, \dots, v_\ell = n.$$

  Call P the *spine* of T. Every vertex $x \notin V(P)$ lies in a unique subtree attached to some vertex of P. This path is where all the cycle minimum labels will lie.

- For $0 < i < \ell$ call v_i a *local minimum on P* if $v_i < v_{i-1}$ and $v_i < v_{i+1}$ We also declare the endpoints $v_0 = 1$ and $v_\ell = n$ to be local minima. List all local minima along P in order: $m_1, m_2, \dots, m_r$, where $m_1 = 1$ and $m_r = n$.

In the forward construction we start from D_f, whose components each consist of one directed cycle with rooted trees feeding into it. When we perform the "cycle surgery" (cut the edge leaving the minimum of each cycle and connect these minima in increasing order), only edges belonging to cycles are affected. All vertices that did *not* lie on cycles keep the same unique neighbour that lies closer to the cycle and therefore cannot lie on the path between 1 and n, because both 1 and n are cycle vertices. Hence the spine P is exactly the union of all cycle-vertices.

On each original cycle C, the chosen minimum m is the unique vertex whose label is smaller than the labels of its two neighbours on the cycle. After cutting at m and reconnecting the cycles in increasing order of their minima, the local picture around m on the spine is unchanged: its two neighbours on P are still vertices belonging either to its own cycle or to a later cycle, and all of those vertices have label $> m$. Conversely, no other vertex on P can have both neighbours larger, since within each cycle there is only one such vertex (the minimum), and the spine traverses the vertices of each cycle in a contiguous block. Thus the local minima on P are precisely the cycle minima used in the forward map.

For $1 \le j \le r-1$, define the j-th *block* of P by

$$B_j := (v_{i(j)}, v_{i(j)+1}, \dots, v_{i(j+1)}),$$

and set $B_r := \{m_r\}$. So $B_1, \dots, B_r$ are pairwise disjoint and their union is $V(P)$. If T came from f, then:

- the vertices in B_j are exactly those of the j-th directed cycle of D_f;
- m_j is the minimum of that cycle.

We now define a directed cycle on each block.

- For $1 \le j \le r-1$: write B_j as $B_j = (x_0, x_1, \dots, x_t) \quad$ with $x_0 = m_j,\ x_t = m_{j+1}$. On this set we put the edges $x_1 \to x_2 \to \cdots \to x_t \to x_0 \to x_1$, obtaining a directed cycle C_j whose vertices are the elements of B_j and whose minimum is $x_0 = m_j$.
- For the last block $B_r = \{m_r\}$ we put the trivial cycle $m_r \to m_r$, i.e. we set $f(m_r) = m_r$.

Note that each vertex on the spine now has out-degree 1 coming from its cycle.

Consider a vertex $u \notin V(P)$. In T there is a unique simple path from u to the spine P; let $p(u)$ be the neighbour of u on this path (the "parent" of u with respect to P). We now define a directed edge

$$u \to p(u),$$

for every such vertex u. This produces directed trees with all edges oriented *towards* the spine. Now interpret this as a functional digraph. □

3.5 Multinomial coefficients

Proposition 3.5 (Multinomial coefficient). Let $k_1, \dots, k_r$ be nonnegative integers with

$$k_1 + \cdots + k_r = m.$$

The number of words of length m over an alphabet $\{1, \dots, r\}$ in which the letter i appears exactly k_i times is

$$\frac{m!}{k_1! \cdots k_r!}.$$

Proof. Consider a multiset with k_i copies of the symbol i for each $i = 1, \dots, r$, so in total $m = k_1 + \cdots + k_r$ symbols.

Every word of length m in which symbol i appears exactly k_i times is just a permutation of this multiset, and conversely every permutation of the multiset gives such a word.

If all m symbols were distinct, there would be $m!$ permutations. But permuting the k_i identical copies of symbol i does not change the word, so we divide by $k_i!$ for each i. Hence the number of distinct words is

$$\frac{m!}{k_1!k_2! \cdots k_r!}.$$

□

Corollary 3.6. Let T be a labelled tree on $[n] = \{1, \dots, n\}$ and let d_i be the degree of vertex i in T. Assume $d_1, \dots, d_n$ are positive integers with

$$\sum_{i=1}^{n} d_i = 2(n-1)$$

(so they are a possible degree sequence for a tree on $[n]$). Then the number of labelled trees on $[n]$ with $\deg_T(i) = d_i$ for every i is

$$\frac{(n-2)!}{\prod_{i=1}^{n}(d_i - 1)!}.$$

Proof. For any tree T on $[n]$ we have $\sum_{i=1}^{n} d_i = 2|E(T)| = 2(n-1)$ by the handshake lemma.

Recall the bijection between labelled trees on $[n]$ and words of length $n-2$ over the alphabet $[n]$. Under this bijection, the number of occurrences of the letter i in the word equals $d_i - 1$, where d_i is the degree of vertex i in the tree.

Hence, if we want all trees with degree sequence $(d_1, \dots, d_n)$, we must count all words of length $n-2$ in which letter i appears exactly $d_i - 1$ times. Since

$$\sum_{i=1}^{n}(d_i - 1) = \left(\sum_{i=1}^{n} d_i\right) - n = 2(n-1) - n = n-2,$$

the multinomial coefficient applies with

$$m = n-2, \qquad k_i = d_i - 1.$$

By the proposition, the number of such words is

$$\frac{(n-2)!}{\prod_{i=1}^{n}(d_i-1)!}.$$

□

Definition 3.6 (Multinomial coefficient). For nonnegative integers $k_1, \dots, k_t$ with $k_1 + \cdots + k_t = k$, the *multinomial coefficient* is

$$\binom{k}{k_1, \dots, k_t} := \frac{k!}{k_1! \cdots k_t!}.$$

Theorem 3.7 (Multinomial theorem). Let $x_1, \dots, x_n$ be variables and $k \in \mathbb{N}$. Then

$$\Big(\sum_{i=1}^{n} x_i\Big)^k = \sum_{\substack{k_1,\dots,k_n \geq 0 \\ k_1+\cdots+k_n=k}} \binom{k}{k_1, \dots, k_n} \prod_{i=1}^{n} x_i^{k_i},$$

Proof. Expand the product

$$(x_1 + \cdots + x_n) \cdots (x_1 + \cdots + x_n)$$

(k factors) by distributivity. Each term in the expansion is a monomial $x_1^{k_1} \cdots x_n^{k_n}$ with $k_1 + \cdots + k_n = k$, obtained by choosing x_i from exactly k_i of the k factors. The number of ways to make such a choice is the multinomial coefficient $\binom{k}{k_1,\dots,k_n}$, giving the stated formula. □

Theorem 3.8 (Fermat's Little Theorem). Let p be prime and $n \in \mathbb{Z}$. Then if $p \nmid n$, then $n^p \equiv n \pmod{p}$.

Proof. Apply the multinomial theorem with $x_1 = \cdots = x_n = 1$ and $k = p$:

$$n^p = \sum_{\substack{k_1,\dots,k_n \geq 0 \\ k_1+\cdots+k_n=p}} \binom{p}{k_1, \dots, k_n} 1^{k_1} \cdots 1^{k_n} = \sum_{\substack{k_1,\dots,k_n \geq 0 \\ k_1+\cdots+k_n=p}} \binom{p}{k_1, \dots, k_n}.$$

Claim. If $(k_1, \dots, k_n) \neq (p, 0, \dots, 0)$ and not a permutation of it, then

$$\binom{p}{k_1, \dots, k_n} \equiv 0 \pmod{p}.$$

Indeed, for such a $(k_1, \dots, k_n)$ we have $0 \leq k_i \leq p-1$ for every i, so none of the factorials $k_i!$ is divisible by p, while $p!$ is divisible by p. Hence p divides the numerator but not the denominator of

$$\binom{p}{k_1, \dots, k_n} = \frac{p!}{k_1! \cdots k_n!},$$

so the multinomial coefficient is $0 \pmod{p}$.

The only multi-indices that can give a nonzero term modulo p are therefore those with one $k_i = p$ and all others 0. For each $i \in \{1, \ldots, n\}$ we have

$$\binom{p}{0, \ldots, 0, \overset{i}{p}, 0, \ldots, 0} = \frac{p!}{p!\,0! \cdots 0!} = 1,$$

so modulo p the sum reduces to

$$n^p \equiv \underbrace{1 + \cdots + 1}_{n \text{ times}} = n \pmod{p},$$

□

3.6 Ballot Theorem

Theorem 3.9 (Ballot theorem)**.** Let candidates A and B receive a and b votes respectively, with $a \geq b$. Assume that the $(a+b)$ votes are revealed in a uniformly random order. Then the number of vote sequences in which, after every initial segment, the number of votes for A is *not less* than the number of votes for B (i.e. A never trails) is

$$\binom{a+b}{a} - \binom{a+b}{a+1}.$$

Strategy: Directly counting "good" paths (those that stay below the diagonal) is hard because the constraint is global—it applies at every step. Instead, we use the **complementary counting** strategy:

1. Count *all* paths from $(0,0)$ to (a,b).
2. Count the "bad" paths (those that touch or cross the forbidden line $y = x+1$).
3. Subtract bad from all.

The genius of André's Reflection Principle is a bijection for step 2: if a path hits the forbidden line, we reflect the remainder of the path across that line. This creates a one-to-one correspondence between "bad paths to (a,b)" and "all paths to a specific reflected endpoint," which are easy to count.

Proof. Represent a vote for A by a right step $(1,0)$ and a vote for B by an up step $(0,1)$. Then each ordering of the a votes for A and b votes for B corresponds to a lattice path from $(0,0)$ to (a,b) using only steps $(1,0)$ and $(0,1)$.

- The total number of such paths is
$$\binom{a+b}{a},$$
since we must choose which a of the $a+b$ steps are the horizontal ones (the remaining b are vertical).
- A *ballot* path (a good sequence) is one that never goes *above* the diagonal $y = x$; equivalently, for every prefix of the sequence, we have $\#A \geq \#B$.
- A *non-ballot* path is one for which at some point $\#B > \#A$, i.e. the path goes strictly above the diagonal.

Fix a non-ballot path from $(0,0)$ to (a,b) with $a \geq b$. Let k be the smallest integer such that at the point $(x,y) = (k,k+1)$. This is exactly the first time the path has B strictly leads ahead of A, so such k exists for every non-ballot path).

Up to this time the vote counts are

$$\text{A-votes} = k, \qquad \text{B-votes} = k+1.$$

After this time, along the original path, we must still reach (a,b), so the remaining steps contribute

$$\text{A-votes} = a-k, \qquad \text{B-votes} = b-k-1.$$

We now *reflect the path after* $(k,k+1)$ *across the diagonal* $y = x+1$, which just interchanges the roles of horizontal and vertical steps in the suffix.

In terms of vote counts this gives a new path with:

	up to $(k, k+1)$		after $(k, k+1)$		endpoint
original path	$A : k$	$B : k+1$	$A : a-k$	$B : b-k-1$	(a, b)
new path	$A : k$	$B : k+1$	$A : b-k-1$	$B : a-k$	$(b-1, a+1)$

Thus the image of our non-ballot path ends at

$$\big(k + (b-k-1),\ (k+1) + (a-k)\big) = (b-1, a+1).$$

So this construction defines a map

$$\Phi : \{\text{non-ballot paths from } (0,0) \text{ to } (a,b)\} \longrightarrow \{\text{all paths from } (0,0) \text{ to } (b-1, a+1)\}.$$

Claim. Φ is a bijection.

Injective: given the image path, the first time it hits the line $y = x+1$ is again the point $(k, k+1)$; reflecting the suffix back across this line recovers the original path. Thus we can invert Φ.

Surjective: every path from $(0,0)$ to $(b-1, a+1)$ must at some point reach a point with $y = x+1$ (since $b-1 < a+1$ we end strictly below the diagonal). Let $(k, k+1)$ be the *first* such point; reflecting the suffix across $y = x+1$ produces a path from $(0,0)$ to (a,b) whose first visit to $y = x+1$ is exactly $(k, k+1)$, and hence B leads there and the path is non-ballot. This is the inverse of Φ.

Therefore

$$\#\{\text{non-ballot paths from } (0,0) \text{ to } (a,b)\} \;=\; \#\{\text{paths from } (0,0) \text{ to } (b-1, a+1)\} \;=\; \binom{a+b}{a+1}.$$

Substituting into

$$\#\{\text{ballot paths}\} = \binom{a+b}{a} - \#\{\text{non-ballot paths}\}$$

gives the Ballot theorem:

$$\#\{\text{ballot paths}\} = \binom{a+b}{a} - \binom{a+b}{a+1}.$$

□

Lemma 3.10. Let $m \geq 1$. Consider lattice paths in $\mathbb{Z}^2$ that start at $(0,0)$ and use only steps $(1,0)$ (right) and $(0,1)$ (up), and have total length $2m$. Then the following three families of paths all have the same cardinality, namely $\binom{2m}{m}$:

1. paths that end at (m, m);
2. paths that never go strictly above the diagonal $y = x$;
3. paths that never return to the diagonal $y = x$ after time 0.

Proof. **(i)** A path of length $2m$ ends at (m, m) iff it has exactly m right-steps and m up-steps. Choosing the positions of the m right-steps gives

$$\#\{\text{paths ending at } (m,m)\} = \binom{2m}{m}.$$

(ii) Fix integers $a \geq b \geq 0$ and let $\ell = a + b$. By the ballot / reflection argument, the number of paths from $(0,0)$ to (a,b) that never go above $y = x$ is

$$\binom{\ell}{a} - \binom{\ell}{a+1}.$$

(Among all $\binom{\ell}{a}$ paths to (a,b), exactly $\binom{\ell}{a+1}$ go above the diagonal; reflect those at the first step above $y = x$ to obtain a bijection with paths to $(a+1, b-1)$.)

Now take $\ell = 2m$ and sum over all admissible endpoints (a,b) with $a \geq b$ and $a + b = 2m$, i.e. over $a = m, m+1, \ldots, 2m$:

$$\sum_{a=m}^{2m} \left[\binom{2m}{a} - \binom{2m}{a+1} \right] = \binom{2m}{m} - \binom{2m}{2m+1} = \binom{2m}{m},$$

since the sum telescopes. This is exactly the number of paths of length $2m$ from $(0,0)$ that never go above $y = x$.

(iii) We count paths of length $2m$ from $(0,0)$ that never return to $y = x$ after time 0.

Such a path must leave the diagonal immediately, so its first step is either $(1,0)$ or $(0,1)$. By symmetry, the numbers of paths with first step $(1,0)$ and with first step $(0,1)$ are equal. Hence

$$\#\{\text{paths of length } 2m \text{ never returning to } y = x\} = 2 \cdot N,$$

where N is the number of such paths whose first step is $(1,0)$.

After the first step $(1,0)$ the path is at $(1,0)$ and has $2m-1$ steps remaining. The condition "never return to $y = x$" is equivalent to "never cross the line $y = x$", i.e. staying strictly below $y = x$. Shifting the coordinate system by $(-1,0)$, this is the same as a path of length $2m-1$ starting at $(0,0)$ that never goes above the line $y = x - 1$, which (after another shift) is equivalent to a path that never goes above the diagonal.

Thus, by part (ii) with $2m-1$ in place of $2m$, we have

$$N = \binom{2m-1}{m},$$

so

$$\#\{\text{paths of length } 2m \text{ never returning to } y = x\} = 2\binom{2m-1}{m} = \binom{2m}{m},$$

using the identity $2\binom{2m-1}{m} = \binom{2m}{m}$.

Combining (i)–(iii) shows that all three families have size $\binom{2m}{m}$. □

Theorem 3.11. For every integer $n \geq 0$,

$$\sum_{k=0}^{n} \binom{2k}{k}\binom{2n-2k}{n-k} = 4^n.$$

Proof. Interpret the right-hand side combinatorially. A lattice path of length $2n$ with steps $(1,0)$ and $(0,1)$ is determined by the choice of each step, so there are $2^{2n} = 4^n$ such paths starting at $(0,0)$.

Group these paths according to the *last* time they are on the diagonal $y = x$. For a given $k \in \{0, \ldots, n\}$, consider those paths whose last visit to the diagonal is at the point (k,k).

- The prefix from $(0,0)$ to (k,k) is an arbitrary path of length $2k$ ending at (k,k), so there are $\binom{2k}{k}$ choices.
- The suffix of length $2n-2k$ starts at (k,k) and never returns to $y=x$. Translating (k,k) to the origin, the number of such suffixes equals, by the lemma with $m=n-k$ and property (iii),

$$\binom{2(n-k)}{n-k} = \binom{2n-2k}{n-k}.$$

Thus the number of paths whose last visit to the diagonal is at (k,k) is $\binom{2k}{k}\binom{2n-2k}{n-k}$. Summing over all possible k gives

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

paths in total, which must equal the total number 4^n of length-$2n$ paths. This proves the identity. □

Definition 3.7 (Ballot path and ballot sequence)**.** A *ballot path of length* $2n$ is a lattice path from $(0,0)$ to (n,n) using steps $(1,0)$ (east) and $(0,1)$ (north) that never goes strictly above the diagonal $x=y$.

Equivalently, a *ballot sequence* of length $2n$ is a word in $\{0,1\}$ containing n zeros and n ones such that in every initial segment the number of 1's is at least the number of 0's.

3.7 Catalan numbers

Definition 3.8 (Catalan numbers)**.** The *nth Catalan number* C_n is the number of ballot paths (or ballot sequences) of length $2n$.

Theorem 3.12 (Closed form for Catalan numbers)**.** For every $n \geq 0$,

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

Definition 3.9.

- A *rooted tree* is a tree together with a distinguished vertex called the *root*.
- In a rooted tree, the parent of a vertex $v \neq$ root is the previous vertex on the unique path from the root to v.
- A *leaf* is a vertex of degree 1 (except in the trivial tree).
- An *ordered rooted tree* is a rooted tree in which, for every vertex, the children are linearly ordered (from "left" to "right"). The vertices themselves are not labeled.
- A *rooted ordered binary tree* is an ordered rooted tree in which each vertex has either 0 or 2 children. (If there are two children, one is designated "left" and one "right".)

Theorem 3.13. 1. The number of rooted ordered binary trees with $n+1$ leaves is C_n.

2. The number of triangulations of a convex $(n+2)$-gon is C_n.

Proof. **(i) Binary trees and ballot sequences.** We describe a bijection between rooted ordered binary trees with $n+1$ leaves and ballot sequences of length $2n$.

Given such a tree T, perform the following depth-first (preorder) traversal, starting at the root:

- When a vertex is first visited, mark it visited, write a 1 if it has children (i.e. it is an internal vertex) or 0 if it is a leaf.
- Then recursively visit the left subtree, then the right subtree, returning upwards along edges as usual. When we return to a visited vertex we do not write anything new.

Let the sequence obtained be $b_1 b_2 \ldots b_{2n} \in \{0,1\}^{2n}$.

A rooted ordered binary tree with $n+1$ leaves has n internal vertices, so we write n ones and n zeros. Thus the sequence has length $2n$ with n 1's and n 0's.

We claim that this is a ballot sequence. Consider any initial segment of the traversal. Whenever we write a 1 we "create" two new children; whenever we write a 0 we finish a leaf and effectively close off one child. A short induction on the steps of the traversal shows that after any initial segment, the number of 1's is at least the number of 0's: otherwise we would have closed more leaves than the number of child positions created, and there would not be any vertex to continue the traversal from. Hence the sequence never goes below the line "#1's = #0's", so it is ballot.

Conversely, given a ballot sequence of length $2n$ with n ones and n zeros, one can reconstruct a unique rooted ordered binary tree by the reverse procedure: scan the sequence from left to right, starting with a root whose two child positions are "open". Whenever a 1 appears, we replace one open child position by an internal vertex with two new open child positions; when a 0 appears we close one open child position by making it a leaf. The ballot condition guarantees that we never run out of open child positions, and the total number of zeros ensures we finish with no open positions left. This reconstructs a unique tree with $n+1$ leaves.

Hence we have a bijection, and the number of such trees is C_n.

(ii) Triangulations of a convex polygon. Fix a convex $(n+2)$-gon and choose one side as a distinguished "root side". Given any triangulation, place a new vertex in the interior of each triangle and connect it to the midpoints of the three edges of that triangle. The graph formed by the new vertices and the segments across edges of the triangulation is a rooted ordered binary tree with $n+1$ leaves, rooted at the triangle adjacent to the root side, and with leaves in bijection with the sides of the polygon.

Thus triangulations of a convex $(n+2)$-gon are in bijection with rooted ordered binary trees with $n+1$ leaves; part (i) now gives the result. □

Theorem 3.14 (Catalan recurrence)**.** For $n \geq 1$ we have the recurrence

$$C_n = \sum_{k=0}^{n-1} C_k \, C_{n-1-k}.$$

Proof. We use the interpretation of C_n as the number of rooted ordered *full* binary trees with n internal vertices.

Fix $n \geq 1$ and consider such a tree T with n internal vertices.

The root is an internal vertex and therefore has exactly two children: a left child and a right child. Each child is the root of a (possibly empty) full binary subtree.

Let

$$L = \text{left subtree}, \qquad R = \text{right subtree}.$$

Suppose L has k internal vertices. Then:

$$\#\text{internal vertices in } R = n - 1 - k,$$

since the total is n, and we have already counted the root and the k internal vertices in L.

Thus every tree T with n internal vertices determines a unique integer $k \in \{0, 1, \ldots, n-1\}$ and a pair of trees

$$(L, R) \quad \text{with} \quad L \text{ having } k \text{ internal vertices, } R \text{ having } n - 1 - k \text{ internal vertices.}$$

- There are C_k choices for the left subtree L (any full binary tree with k internal vertices).
- Independently, there are C_{n-1-k} choices for the right subtree R (any full binary tree with $n - 1 - k$ internal vertices).

Once L and R are chosen, attaching them as left and right subtrees of a new root produces a unique full binary tree with n internal vertices. Conversely, any such tree arises in exactly this way from its left and right subtrees.

Summing over all k,

$$C_n = \sum_{k=0}^{n-1} \big(\# \text{ trees with } k \text{ internal vertices in the left subtree}\big) = \sum_{k=0}^{n-1} C_k \, C_{n-1-k},$$

as claimed. □

4 Recurrences

4.1 Fibonacci recurrences

Example 4.1 (Smarts and Cadillacs). Consider a linear parking lot of size n in a row. We have *Smarts* (cars of length 1) and *Cadillacs* (cars of length 2). Let F_n be the number of ways to occupy the n spots (every spot is either occupied by a car or empty, with no overlaps). Can we determine an expression for F_n?

If we look at the leftmost car:

- Either it is occupied by a Smart; remove it and we obtain a configuration of size $n-1$.
- Or it begins a Cadillac; remove the Cadillac and we obtain a configuration of size $n-2$.

Then

$$F_n = F_{n-1} + F_{n-2} \quad (n \geq 2),$$

with initial conditions $F_0 = 1$ (empty lot) and $F_1 = 1$ (only one Smart). Hence the number of configurations satisfies the Fibonacci recurrence. Thus F_n is the nth Fibonacci number $(1, 1, 2, 3, 5, \dots)$.

Lemma 4.1. For $n \geq 0$,

$$\sum_{i=0}^{n} F_i^2 = F_n F_{n+1}.$$

Proof. Consider two parallel lots: the top of length n and the bottom of length $n+1$. A pair of fillings is counted by $F_n F_{n+1}$.

Given such a pair, scan from the right and let i be the *rightmost* position $0 \leq i \leq n$ at which one may place a vertical barrier without cutting a Cadillac in either lot. This i is well-defined and unique.

Then in both lots the segment to the left of the barrier has length i, hence each left segment is an arbitrary filling of length i. These classes for $i = 0, 1, \dots, n$ partition all $F_n F_{n+1}$ pairs, so

$$F_n F_{n+1} = \sum_{i=0}^{n} F_i^2. \qquad \square$$

4.2 Derangements

Definition 4.1 (Derangement). A *derangement* of $[n]$ is a permutation $\sigma \in S_n$ with no fixed point, i.e. $\sigma(i) \neq i$ for all i. Let D_n be the number of derangements of $[n]$.

Theorem 4.2 (Derangements). For $n \geq 1$ the numbers D_n satisfy:

$$D_n = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

$$D_n = (n-1)\big(D_{n-1} + D_{n-2}\big), \qquad n \geq 2,$$

with initial values $D_0 = 1$ and $D_1 = 0$.

Proof of the recurrence. Fix $n \geq 2$ and consider a derangement σ of $[n]$. Look at the functional digraph of n. Since σ is a derangement, the functional digraph has no loops (1-cycles)

- *Case 1:* n is in a 2-cycle $\{i, n\}$, i.e. $\sigma(n) = i$ and $\sigma(i) = n$ for some $i < n$. There are $n-1$ choices for i, and after fixing this 2-cycle the remaining $n-2$ elements must form a derangement. Thus we obtain $(n-1)D_{n-2}$ derangements.
- *Case 2:* n is in a cycle of length at least 3. Then $\sigma(n) = i$ for some $i < n$ and $\sigma(i) \neq n$. If we delete n from the cycle and "splice" the edges appropriately, we obtain a derangement of $[n-1]$, and conversely we can insert n into any cycle of a derangement of $[n-1]$ in $(n-1)$ different ways. Hence we obtain $(n-1)D_{n-1}$ derangements.

Adding the two cases gives

$$D_n = (n-1)D_{n-2} + (n-1)D_{n-1} = (n-1)(D_{n-1} + D_{n-2}).$$

□

4.3 Simple words, set partitions and permutations with cycles

We collect three classical families with similar recurrences.

Example 4.2. Let $P(n,k)$ denote the number of *simple k-words* on $[n]$, i.e. words of length k over alphabet $[n]$ with no repeated letters in a word. Clearly $P(n,k) = n(n-1)\cdots(n-k+1)$. Can we obtain an recurrence relation for $P(n,k)$?

In a simple k-word on $[n]$,

1. either the letter n does not appear, giving $P(n-1,k)$ possibilities
2. n appears in some position $1 \leq i \leq k$ (choose the position for n in k ways) and the remaining $k-1$ positions contain a simple $(k-1)$-word on $[n-1]$, giving $kP(n-1,k-1)$ possibilities.

The recurrence (valid for $n \geq 1$, $k \geq 1$) is

$$P(n,k) = P(n-1,k) + k\,P(n-1,k-1),$$

with initial condition $P(n,0) = 1$ and $P(0,k) = 0$ for $k \geq 1$.

Example 4.3. Let $S(n,k)$ denote the number of ways to partition $[n]$ into k *(nonempty, unlabeled)* blocks. These are the Stirling numbers of the second kind. Derive an recurrence relation for $S(n,k)$.

Consider where the element n goes. There are 2 case:

1. It forms a singleton block (contributing $S(n-1,k-1)$)
2. It joins one of the existing k blocks of a partition of $[n-1]$ into k blocks (contributing $kS(n-1,k)$).

The recurrence (again for $n \geq 1$, $k \geq 1$) is

$$S(n,k) = k\,S(n-1,k) + S(n-1,k-1),$$

with initial conditions $S(0,0) = 1$, $S(n,0) = 0$ for $n \geq 1$, $S(0,k) = 0$ for $k \geq 1$.

Example 4.4. Let $C(n,k)$ denote the number of permutations of $[n]$ with exactly k disjoint cycles in their cycle decomposition (these are called the "signless" Stirling numbers of the first kind). Derive an recurrence relation for $C(n,k)$.

Given a permutation of $[n]$ with k cycles, look at n. There are 2 cases:

1. n is a fixed point, in which case removing n gives a permutation of $[n-1]$ with $k-1$ cycles.
2. n lies in a cycle of length at least 2, in which case deleting n and splicing the cycle gives a permutation of $[n-1]$ with k cycles, and conversely n can be inserted into any of the $n-1$ positions in any cycle.

The recurrence (for $n \geq 1$, $k \geq 1$) is

$$C(n,k) = (n-1)\,C(n-1,k) + C(n-1,k-1),$$

with initial conditions $C(0,0) = 1$, $C(n,0) = 0$ for $n \geq 1$, $C(0,k) = 0$ for $k \geq 1$.

4.4 Delannoy recurrences

Example 4.5. With Delannoy numbers, we have $d_{0,0} = 1$ and $d_{m,0} = d_{0,n} = 1$ for all $m, n \geq 1$. Removing the last step of a path gives the recursion

$$d_{m,n} = d_{m-1,n} + d_{m,n-1} + d_{m-1,n-1} \qquad (m, n \geq 1).$$

Example 4.6. Let $a_{m,n}$ be number of points in the lattice ball of radius m in $\mathbb{Z}^n$ Derive an recurrence relation for $a_{m,n}$.

Consider a point $(x_1, \ldots, x_n)$ with $|x_1| + \cdots + |x_n| \leq m$. Partition these points according to the last coordinate x_n.

- If $x_n = 0$, then $(x_1, \ldots, x_{n-1})$ is a point in the $(n-1)$-dimensional ball of radius m, contributing $a_{m,n-1}$ points.
- If $x_n > 0$, write $x_n' = x_n - 1 \geq 0$; then $|x_1| + \cdots + |x_{n-1}| + x_n' \leq m - 1$, so we obtain a point counted by $a_{m-1,n}$.
- If $x_n < 0$, write $x_n'' = -x_n - 1 \geq 0$; again we get a point with sum of absolute values at most $m - 1$. This gives another $a_{m-1,n}$ points, but the two cases together can be encoded as a contribution $a_{m-1,n-1}$ when we treat the sign separately and focus on the positions of nonzero coordinates.

Then $a_{m,n}$ satisfy

$$a_{0,n} = 1, \quad a_{m,0} = 1 \qquad (m, n \geq 0),$$

and for $m, n \geq 1$,

$$a_{m,n} = a_{m,n-1} + a_{m-1,n} + a_{m-1,n-1}.$$

Comparing with the recursion and initial conditions for $d_{m,n}$, we obtain another proof that the number of points in the lattice-ball of radius m in $\mathbb{Z}^n$ equals the number of Delannoy paths from $(0,0)$ to (m,n).

5 Solution methods for linear recurrences

5.1 Recurrence relation

Definition 5.1 (Sequence). A (real) *sequence* is a list

$$(a_n)_{n\ge 0} = (a_0, a_1, a_2, \dots),$$

where $a_n \in \mathbb{R}$ for each integer $n \ge 0$.

Definition 5.2 (Recurrence relation, order, linear, homogeneous). Let $(a_n)_{n\ge 0}$ be a sequence.

- A *recurrence relation of order* k for (a_n) is a rule of the form

$$a_n = g\big(n, a_{n-1}, a_{n-2}, \dots, a_{n-k}\big) \qquad (n \ge k),$$

  where g is some function of n and the previous k terms.
- The recurrence is called *linear* if it can be written

$$a_n = g_1(n)a_{n-1} + g_2(n)a_{n-2} + \cdots + g_k(n)a_{n-k} + f(n),$$

  for some functions $g_1, \dots, g_k, f$ of n.
- A linear recurrence is called *homogeneous* if $f(n) \equiv 0$ for all n; otherwise it is *non-homogeneous*.

Example 5.1 (Some recurrences).

- Order 1, homogeneous:

$$a_n = 3a_{n-1}.$$

- Order 2, homogeneous:

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

- Order 2, non-homogeneous:

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

- General homogeneous linear recurrence of finite order:

$$a_n = \sum_{k=1}^{m} c_k(n)\, a_{n-k},$$

  where $c_k(n)$ are given coefficient functions.

Example 5.2 (Catalan numbers). The *Catalan numbers* $(C_n)_{n\ge 0}$ are defined by

$$C_0 = 1, \qquad C_n = \sum_{k=0}^{n-1} C_k\, C_{n-1-k} \qquad (n \ge 1).$$

Partition by the last time each Catalan lattice path touches the line $y = x$

5.2 Linear recurrences with constant coefficients

We now focus on *linear recurrences with constant coefficients*.

Definition 5.3 (Linear constant-coefficient recurrence). A recurrence of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k} + f(n), \qquad n \geq k,$$

where $c_1, \ldots, c_k$ are fixed real constants and $f(n)$ is a given function, is called a *linear constant-coefficient recurrence of order* k. It is *homogeneous* if $f(n) \equiv 0$.

Definition 5.4 (Characteristic polynomial and equation). For the homogeneous relation

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k},$$

the *characteristic polynomial* is

$$\phi(x) = x^k - c_1 x^{k-1} - c_2 x^{k-2} - \cdots - c_k.$$

The equation $\phi(x) = 0$ is called the *characteristic equation*; its roots are the *characteristic roots*.

Example 5.3. Consider

$$a_n = \alpha a_{n-1}, \qquad n \geq 1,$$

with initial value $a_0 = c$. The characteristic equation is $x = \alpha$, so α is the only characteristic root. It is easy to check by induction that

$$a_n = c\,\alpha^n$$

for all $n \geq 0$.

More generally, any sequence of the form $a_n = C\alpha^n$ (with C arbitrary) is a solution of the recurrence; the initial condition picks out the particular value $C = c$.

5.3 General homogeneous recurrence with distinct roots

Now consider the homogeneous recurrence of order k:

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k}, \qquad n \geq k,$$

and let its characteristic polynomial be

$$\phi(x) = x^k - c_1 x^{k-1} - \cdots - c_k.$$

Suppose α is a root of $\phi(x)$. Then the sequence $a_n = \alpha^n$ satisfies the recurrence: substituting $a_n = \alpha^n$ gives

$$\alpha^n = c_1 \alpha^{n-1} + \cdots + c_k \alpha^{n-k} \quad \Longleftrightarrow \quad \alpha^k = c_1 \alpha^{k-1} + \cdots + c_k,$$

which is exactly $\phi(\alpha) = 0$.

Thus, for any constant C, the sequence $a_n = C\alpha^n$ is a solution. If β is another root, then $b_n = C'\beta^n$ is also a solution. Because the recurrence is linear, any linear combination

$$a_n = C_1 \alpha^n + C_2 \beta^n$$

is again a solution.

More generally, if $\alpha_1, \ldots, \alpha_k$ are k *distinct* characteristic roots, then each sequence α_i^n is a solution, and any linear combination

$$a_n = C_1\alpha_1^n + C_2\alpha_2^n + \cdots + C_k\alpha_k^n$$

is also a solution. Initial conditions $a_0, \ldots, a_{k-1}$ determine the constants $C_1, \ldots, C_k$ uniquely (the k sequences α_i^n are linearly independent), so this already gives the general solution when all roots are distinct.

Example 5.4 (Fibonacci sequence)**.** The Fibonacci numbers are defined by

$$F_n = F_{n-1} + F_{n-2} \quad (n \geq 2), \qquad F_0 = 0,\ F_1 = 1.$$

The characteristic polynomial is

$$\phi(x) = x^2 - x - 1,$$

with roots

$$\alpha = \frac{1+\sqrt{5}}{2}, \qquad \beta = \frac{1-\sqrt{5}}{2}.$$

Thus every solution of the recurrence has the form

$$a_n = C_1\alpha^n + C_2\beta^n.$$

Imposing $a_0 = F_0 = 0$ and $a_1 = F_1 = 1$ gives

$$C_1 + C_2 = 0, \qquad C_1\alpha + C_2\beta = 1,$$

so $C_1 = \frac{1}{\sqrt{5}}$ and $C_2 = -\frac{1}{\sqrt{5}}$. Hence

$$F_n = \frac{1}{\sqrt{5}}\left(\alpha^n - \beta^n\right),$$

the usual closed form (Binet formula).

This example illustrates a more general phenomenon: the set of all solutions of a linear homogeneous recurrence of order k is a k-dimensional vector space, and suitably many distinct characteristic roots produce a basis of this space.

5.4 General solution with repeated roots

We now describe what happens when the characteristic polynomial has repeated roots.

Theorem 5.1 (General solution with multiple roots). Let

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k}$$

be a linear homogeneous recurrence of order k with constant coefficients, and let its characteristic polynomial factor as

$$\phi(x) = x^k - c_1 x^{k-1} - \cdots - c_k = \prod_{j=1}^{t} (x - \alpha_j)^{d_j},$$

where the α_j are distinct and $d_1 + \cdots + d_t = k$.

Then the space of all solutions $(a_n)_{n \geq 0}$ is spanned by the k sequences

$$n^s \alpha_j^n \qquad (1 \leq j \leq t,\ 0 \leq s \leq d_j - 1).$$

Equivalently, every solution can be written in the form

$$a_n = \sum_{j=1}^{t} P_j(n)\, \alpha_j^n,$$

where each $P_j(n)$ is a polynomial in n of degree at most $d_j - 1$.

Proof sketch. Write the recurrence as a linear operator equation

$$a_n - c_1 a_{n-1} - \cdots - c_k a_{n-k} = 0.$$

Introduce the shift operator E acting on sequences by $(Ea)_n = a_{n+1}$. The recurrence becomes

$$\left(E^k - c_1 E^{k-1} - \cdots - c_k\right) a = 0,$$

i.e. $\phi(E)a = 0$.

Factor $\phi(x) = \prod_{j=1}^{t} (x - \alpha_j)^{d_j}$; formally this gives

$$\prod_{j=1}^{t} (E - \alpha_j)^{d_j} a = 0.$$

For a fixed j, the solutions of $(E - \alpha_j)a = 0$ are exactly the geometric sequences $a_n = C\alpha_j^n$. Solutions of $(E - \alpha_j)^{d_j} a = 0$ are then obtained by taking derivatives with respect to α_j; this produces the additional factors of n and leads to the d_j linearly independent sequences $n^s \alpha_j^n$ for $0 \leq s \leq d_j - 1$.

Taking the product over j shows that the full solution space has dimension $d_1 + \cdots + d_t = k$ and is spanned by these sequences. Finally, any choice of initial values $a_0, \ldots, a_{k-1}$ yields a unique linear combination of these basis solutions, so every solution has the claimed form. □

Consider a sequence $(a_n)_{n \geq 0}$ satisfying a homogeneous linear recurrence of order k with constant coefficients

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k} \qquad (n \geq k),$$

where $c_k \neq 0$. The *characteristic polynomial* is

$$\chi(x) = x^k - c_1 x^{k-1} - \cdots - c_{k-1} x - c_k.$$

If the roots of χ are $\lambda_1, \dots, \lambda_r$ with multiplicities $d_1, \dots, d_r$, then every solution has the form

$$a_n = \sum_{j=1}^{r} P_j(n)\,\lambda_j^{\,n},$$

where P_j is a polynomial of degree at most $d_j - 1$. The coefficients of the P_j's are determined from the k initial values $a_0, \dots, a_{k-1}$.

5.5 Tower of Hanoi

Example 5.5 (Tower of Hanoi). Let h_n be the minimum number of moves required to move a tower of n disks from one peg to another according to the following rules.

Rules of the Tower of Hanoi. We are given three pegs (often called *source*, *auxiliary*, and *target*) and n disks of distinct sizes. Initially, all n disks are stacked on the source peg in increasing order of size from top to bottom (i.e., the smallest disk is on top and the largest disk is on the bottom).

A *legal move* consists of taking the top disk from one peg and placing it onto the top of another peg, subject to the following constraints:

1. Only one disk may be moved at a time.
2. You may only move the top disk of any peg.
3. At all times, no disk may be placed on top of a smaller disk. Equivalently, on each peg the disks must always appear in increasing order of size from top to bottom.

The goal is to start from the initial configuration on the source peg and end with all n disks stacked in the same order on the target peg, using only legal moves.

- We clearly have $h_0 = 0$ and $h_1 = 1$.
- To move a tower of size n:
  1. move the top $n-1$ disks to the spare peg (takes h_{n-1} moves),
  2. move the largest disk (one move),
  3. move the tower of $n-1$ disks from the spare peg onto the largest disk (another h_{n-1} moves).

  Therefore
$$h_n = 2h_{n-1} + 1 \qquad (n \geq 1).$$

We solve this recurrence.

Homogeneous part:
$$h_n^{(h)} = 2h_{n-1}^{(h)} \quad \Longrightarrow \quad h_n^{(h)} = C \cdot 2^n.$$

For a particular solution we try a constant $h_n^{(p)} = A$:
$$A = 2A + 1 \quad \Longrightarrow \quad A = -1.$$

Thus the general solution is
$$h_n = C \cdot 2^n - 1.$$

Using $h_0 = 0$ gives $C = 1$, so
$$h_n = 2^n - 1.$$

5.6 Non-homogeneous recurrences

We now consider recurrences of the form

$$a_n = c_1 a_{n-1} + \cdots + c_k a_{n-k} + F(n), \qquad n \geq k, \tag{1}$$

where $F(n)$ is a polynomial in n.

Theorem 5.2. Let $F(n)$ be a polynomial of degree d. Let $\chi(x)$ be the characteristic polynomial of the associated homogeneous recurrence (obtained from (1) by deleting $F(n)$), and suppose r is a root of χ of multiplicity $w \geq 0$.

Then there exists a particular solution of the form

$$a_n^{(p)} = n^w P(n)\, r^n,$$

where P is a polynomial of degree at most d.

Remark 5.1. In particular, if $r = 1$ is a root of multiplicity w then there is a particular solution of the form
$$a_n^{(p)} = n^w P(n)$$
with $\deg P \leq d$. The full solution is then

$$a_n = a_n^{(h)} + a_n^{(p)},$$

where $a_n^{(h)}$ is the general solution of the homogeneous recurrence.

Proof. For clarity, consider the non–homogeneous recurrence

$$a_n - c_1 a_{n-1} - \cdots - c_k a_{n-k} = F(n) r^n,$$

where F is a polynomial of degree d, and let

$$\chi(x) = x^k - c_1 x^{k-1} - \cdots - c_k$$

be the characteristic polynomial. Assume that r is a root of χ of multiplicity w.

First reduce to the case $r = 1$. Put
$$b_n := r^{-n} a_n.$$
Then (b_n) satisfies
$$b_n - c_1 b_{n-1} - \cdots - c_k b_{n-k} = F(n),$$
and the characteristic polynomial is still χ, but now we are interested only in the root 1 of multiplicity w (corresponding to r). Once a particular solution $b_n^{(p)}$ is found, we obtain $a_n^{(p)} = r^n b_n^{(p)}$. Hence it suffices to prove the theorem for $r = 1$.

Let E be the shift operator $(Ea)_n = a_{n+1}$ and put

$$L := E^k - c_1 E^{k-1} - \cdots - c_k,$$

so that the recurrence is $L(a)_n = F(n)$. Since 1 is a root of χ of multiplicity w, we can factor

$$\chi(x) = (x-1)^w \psi(x), \qquad \psi(1) \neq 0,$$

and correspondingly

$$L = (E-1)^w \psi(E).$$

Let $\Delta := E - 1$ be the forward–difference operator. We work on the finite-dimensional vector spaces

$$\mathcal{P}_{\le m} := \{\, \text{polynomials in } n \text{ of degree} \le m \,\}.$$

In the basis $\{1, n, n^2, \ldots, n^m\}$, the matrix of $\psi(E)$ is upper–triangular with diagonal entries all equal to $\psi(1) \neq 0$ (the leading term of $\psi(E)P$ is $\psi(1)$ times the leading term of P). Hence $\psi(E)$ restricts to an invertible linear map

$$\psi(E) : \mathcal{P}_{\le m} \to \mathcal{P}_{\le m}$$

for every m.

For any polynomial P, $\deg \Delta P = \deg P - 1$, so Δ^w maps $\mathcal{P}_{\le d+w}$ into $\mathcal{P}_{\le d}$. In the binomial basis $\{\binom{n}{0}, \binom{n}{1}, \ldots, \binom{n}{d+w}\}$ one has $\Delta\binom{n}{k} = \binom{n}{k-1}$, so

$$\Delta^w \binom{n}{k+w} = \binom{n}{k} \qquad (k \ge 0).$$

Thus Δ^w sends the $(d+1)$–dimensional subspace $\operatorname{span}\{\binom{n}{w}, \ldots, \binom{n}{d+w}\} \subset \mathcal{P}_{\le d+w}$ *onto* $\mathcal{P}_{\le d}$, so it is surjective from $\mathcal{P}_{\le d+w}$ to $\mathcal{P}_{\le d}$.

Given $F \in \mathcal{P}_{\le d}$, as shown above, there is a unique $G \in \mathcal{P}_{\le d}$ with

$$\psi(E)G = F.$$

As shown, there exists $Q \in \mathcal{P}_{\le d+w}$ such that

$$\Delta^w Q = G.$$

Hence

$$LQ = (E-1)^w \psi(E)Q = \Delta^w \psi(E)Q = \Delta^w G = F.$$

So $b_n^{(p)} := Q(n)$ is a particular solution of the recurrence with $r = 1$, and Q has degree at most $d + w$.

Write $Q(n)$ uniquely as

$$Q(n) = n^w P(n) + R(n),$$

with $\deg P \le d$ and $\deg R < w$ (polynomial division by n^w). Consider the sequence $c_n := R(n)$. Since $\deg R < w$, one checks that

$$(E-1)^w c = 0,$$

so $(E-1)^w$ annihilates c, and hence $Lc = 0$ (because $L = (E-1)^w \psi(E)$). Thus c is a solution of the *homogeneous* recurrence.

Now

$$L(n^w P(n)) = LQ - LR = F - 0 = F,$$

so $b_n^{(p)} := n^w P(n)$ is also a particular solution.

Finally, returning to the original sequence $a_n = r^n b_n$, we obtain a particular solution of the form

$$a_n^{(p)} = n^w P(n)\, r^n$$

with $\deg P \le d$, as desired. □

5.7 Regions of the plane

Example 5.6. Let R_n denote the maximum number of regions into which n distinct lines can divide the plane, if no two lines are parallel and no three lines meet in a single point. Derive a recurrence relation for H_n

When we add the n-th line, it intersects each of the previous $n-1$ lines in a distinct point, so it is chopped into n segments. Each segment lies entirely inside a single region determined by the first $n-1$ lines and splits that region into two new regions. Thus the number of regions increases by exactly n:

$$R_n = R_{n-1} + n \qquad (n \geq 1).$$

Clearly $R_0 = 1$ (with no lines, the plane is one region).

This is a first–order non-homogeneous recurrence with constant coefficients and polynomial forcing $F(n) = n$.

The homogeneous recurrence is

$$R_n^{(h)} = R_{n-1}^{(h)},$$

so the characteristic equation is $x - 1 = 0$, with root $r = 1$ of multiplicity 1. Hence

$$R_n^{(h)} = C \cdot 1^n = C.$$

Here $F(n) = n$ has degree $d = 1$ and $r = 1$ has multiplicity $w = 1$, so the theorem tells us to look for a particular solution of the form

$$R_n^{(p)} = P(n),$$

where P is a polynomial of degree at most $d + w = 2$. Write

$$P(n) = an^2 + bn + c$$

and substitute into the recurrence:

$$an^2 + bn + c = a(n-1)^2 + b(n-1) + c + n.$$

Expand the right-hand side:

$$a(n^2 - 2n + 1) + b(n-1) + c + n = an^2 + (-2a + b + 1)n + (a - b + c).$$

Equating coefficients of n^2, n, and the constant term gives

$$\begin{cases} a = a, \\ b = -2a + b + 1, \\ c = a - b + c. \end{cases}$$

Thus

$$-2a + 1 = 0 \quad \Rightarrow \quad a = \frac{1}{2},$$

and

$$a - b = 0 \quad \Rightarrow \quad b = \frac{1}{2}.$$

The constant c is not determined by the recurrence; it will be fixed using the initial condition.

So

$$R_n = C + \frac{1}{2}n^2 + \frac{1}{2}n + c.$$

Using $R_0 = 1$,

$$1 = C + c.$$

We may absorb C into c and simply write

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

Example 5.7. Consider the recurrence

$$a_0 = 0, \qquad a_n = 2a_{n-1} + n \quad (n \geq 1).$$

The homogeneous part $a_n^{(h)} = 2a_{n-1}^{(h)}$ has solution $a_n^{(h)} = C \cdot 2^n$.
Here $F(n) = n$ is degree 1, and the characteristic root is $r = 2$, which is *not* a root of multiplicity ≥ 1 at $r = 2$ for the polynomial $x - 2 = 0$ besides the obvious single root. Thus $w = 0$, and we seek a particular solution of the form

$$a_n^{(p)} = an + b.$$

Substituting into the recurrence:

$$an + b = 2(a(n-1) + b) + n = 2an - 2a + 2b + n.$$

Equating coefficients:

$$\begin{cases} a = 2a + 1, \\ b = -2a + 2b. \end{cases}$$

Hence $-a = 1$ so $a = -1$, and then $b = -2(-1) + 2b$ gives $b = -2$.
So $a_n^{(p)} = -n - 2$, and the general solution is

$$a_n = C \cdot 2^n - n - 2.$$

Using $a_0 = 0$,

$$0 = C \cdot 1 - 0 - 2 \quad \Rightarrow \quad C = 2,$$

so

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

6 Generating function methods for recurrences

Definition 6.1 (Ordinary generating function)**.** Let $(a_n)_{n\geq 0}$ be a sequence of complex numbers. The *ordinary generating function* (OGF) of (a_n) is the formal power series

$$A(x) := \sum_{n\geq 0} a_n x^n \in \mathbb{C}[[x]].$$

For a formal power series $F(x) = \sum_{n\geq 0} f_n x^n$ we write

$$[x^n]F(x) := f_n$$

for the coefficient of x^n in $F(x)$. Thus $a_n = [x^n]A(x)$.

Example 6.1 (Solving a simple recurrence)**.** Let $(a_n)_{n\geq 0}$ be defined by

$$a_0 = 1, \qquad a_n = a_{n-1} + n \quad (n \geq 1).$$

. Determine an expression for a_n.

Let $A(x) = \sum_{n\geq 0} a_n x^n$ be its generating function. Multiply the recurrence by x^n and sum over $n \geq 1$:

$$\sum_{n\geq 1} a_n x^n = \sum_{n\geq 1} a_{n-1} x^n + \sum_{n\geq 1} n x^n.$$

The left-hand side is $A(x) - a_0 = A(x) - 1$. The first sum on the right is $xA(x)$. Using the standard series

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

we get

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

Thus

$$(1-x)A(x) = 1 + \frac{x}{(1-x)^2}, \qquad A(x) = \frac{1}{1-x} + \frac{x}{(1-x)^3}.$$

If we want an explicit formula for a_n, we expand each term:

$$\frac{1}{1-x} = \sum_{n\geq 0} x^n, \qquad \frac{1}{(1-x)^3} = \sum_{n\geq 0} \binom{n+2}{2} x^n.$$

Hence

$$A(x) = \sum_{n\geq 0} x^n + \sum_{n\geq 0} \binom{n+2}{2} x^{n+1} = \sum_{n\geq 0} \left(1 + \binom{n+1}{2}\right) x^n.$$

Therefore

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

in agreement with solving the recurrence by summation.

6.1 The negative binomial / "stars and bars" generating function

Proposition 6.1 (Negative binomial generating function)**.** Let $k \in \mathbb{N}$ and $c \in \mathbb{C}$. Then

$$\frac{1}{(1-cx)^k} = \sum_{n\geq 0} \binom{n+k-1}{k-1} c^n x^n.$$

Proof. Consider the product

$$\frac{1}{(1-cx)^k} = \underbrace{\frac{1}{1-cx} \cdots \frac{1}{1-cx}}_{k \text{ copies}} = \prod_{j=1}^{k} (1 + cx + c^2x^2 + \cdots).$$

To obtain a term $c^n x^n$ in the product, we must choose from the j-th factor a term $c^{\ell_j} x^{\ell_j}$ for some nonnegative integers $\ell_1, \ldots, \ell_k$ with

$$\ell_1 + \cdots + \ell_k = n.$$

Each such k-tuple contributes $c^n x^n$ to the product, and (**??**) follows once we count how many k-tuples $(\ell_1, \ldots, \ell_k)$ of nonnegative integers have sum n.

By the standard "stars and bars" argument, the number of solutions to $\ell_1 + \cdots + \ell_k = n$ in nonnegative integers is $\binom{n+k-1}{k-1}$, so the coefficient of x^n on the right-hand side is exactly $\binom{n+k-1}{k-1} c^n$. □

Corollary 6.2 (Number of weak compositions)**.** For fixed k, the number of k-tuples of nonnegative integers $(\ell_1, \ldots, \ell_k)$ with $\ell_1 + \cdots + \ell_k = n$ is

$$\binom{n+k-1}{k-1} = [x^n] \frac{1}{(1-x)^k}.$$

6.2 Structure theorem for linear recurrences

Let $(a_n)_{n\geq 0}$ satisfy a linear recurrence with constant coefficients of order t:

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_t a_{n-t} \qquad (n \geq t),$$

where $c_t \neq 0$, and let $Q(x)$ be the associated polynomial

$$Q(x) := 1 - c_1 x - c_2 x^2 - \cdots - c_t x^t.$$

Suppose $Q(x)$ has the factorization

$$Q(x) = \prod_{i=1}^{s} (1 - \alpha_i x)^{d_i}$$

with distinct α_i and positive integers d_i.

Theorem 6.3 (Main structure theorem for recurrences)**.** Let $A(x) = \sum_{n\geq 0} a_n x^n$ be the generating function of $(a_n)_{n\geq 0}$. The following are equivalent:

1. The sequence (a_n) satisfies the recurrence $a_n = c_1 a_{n-1} + \cdots + c_t a_{n-t}$ for all $n \geq t$.
2. $A(x)$ is a rational function of the form
$$A(x) = \frac{P(x)}{Q(x)}$$
for some polynomial $P(x)$ of degree $< t$.
3. $A(x)$ can be written as a linear combination of the basic generating functions
$$\frac{1}{(1-\alpha_i x)^j}, \qquad 1 \leq i \leq s,\ 1 \leq j \leq d_i.$$
4. The terms a_n admit a closed form
$$a_n = \sum_{i=1}^{s} P_i(n)\, \alpha_i^n,$$
where each P_i is a polynomial of degree $< d_i$.

Remark 6.1. The proof is a systematic version of what we did in the example: writing $A(x)$ as a rational function, performing partial fraction decomposition into powers of $(1-\alpha_i x)^{-1}$, and then reading off coefficients using the negative binomial expansion.

6.3 Example: Catalan numbers

Recall that the *Catalan numbers* $(C_n)_{n\geq 0}$ are defined recursively by

$$C_0 = 1, \qquad C_n = \sum_{k=1}^{n} C_{k-1}\, C_{n-k} \quad (n \geq 1).$$

Proposition 6.4. Let
$$C(x) := \sum_{n\geq 0} C_n x^n$$
be the generating function of the Catalan sequence. Then
$$C(x) = \frac{1-\sqrt{1-4x}}{2x},$$
and hence
$$C_n = \frac{1}{n+1}\binom{2n}{n} \quad (n \geq 0).$$

Proof. Multiply the recurrence $C_n = \sum_{k=1}^{n} C_{k-1}C_{n-k}$ by x^n and sum over $n \geq 1$:

$$\sum_{n\geq 1} C_n x^n = \sum_{n\geq 1} \sum_{k=1}^{n} C_{k-1}C_{n-k}x^n.$$

The left-hand side is $C(x) - C_0 = C(x) - 1$. On the right-hand side, make the change of variables $i = k - 1$, $j = n - k$. Then $i, j \geq 0$ and $i + j = n - 1$, so

$$\sum_{n\geq 1} \sum_{k=1}^{n} C_{k-1}C_{n-k}x^n = \sum_{i,j\geq 0} C_i C_j x^{i+j+1} = x\left(\sum_{i\geq 0} C_i x^i\right)\left(\sum_{j\geq 0} C_j x^j\right) = xC(x)^2.$$

Thus

$$C(x) - 1 = xC(x)^2, \quad \text{i.e.} \quad xC(x)^2 - C(x) + 1 = 0.$$

This is a quadratic equation for $C(x)$:

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

As a formal power series, $C(x)$ has constant term $C_0 = 1$, whereas $\frac{1+\sqrt{1-4x}}{2x}$ has a pole at $x = 0$, so we must take the minus sign:

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

To extract the coefficients C_n we expand $\sqrt{1-4x}$ using the binomial series with exponent $\frac{1}{2}$:

$$\sqrt{1-4x} = (1-4x)^{1/2} = \sum_{m\geq 0} \binom{1/2}{m}(-4x)^m.$$

It is more convenient to expand $(1-4x)^{-1/2}$ and then integrate, or directly note that

$$1 - \sqrt{1-4x} = 1 - \sum_{m\geq 0} \binom{1/2}{m}(-4x)^m = \sum_{m\geq 1} (-\binom{1/2}{m}(-4)^m) x^m.$$

Dividing by $2x$ and simplifying the binomial coefficients yields the well-known closed form

$$C_n = \frac{1}{n+1}\binom{2n}{n}, \quad n \geq 0.$$

(One standard way is to use the identity $\binom{1/2}{m} = \frac{(-1)^{m-1}}{4^m} \frac{1}{m} \binom{2m}{m}$.) □

6.4 Main theorem of linear recurrences

Let $c_1, \ldots, c_k \in \mathbb{C}$ with $c_k \neq 0$, and set

$$Q(x) = 1 - c_1 x - \cdots - c_k x^k = \prod_{i=1}^{t} (1 - d_i x)^{e_i},$$

where the d_i are distinct complex numbers and the $e_i \geq 1$ are their multiplicities.

Theorem 6.5 (Main theorem). For a complex sequence $a_0, a_1, \ldots$, the following are equivalent.

(A) The sequence satisfies the linear recurrence

$$a_n = c_1 a_{n-1} + \cdots + c_k a_{n-k} \qquad (n \geq k).$$

(B) The ordinary generating function $A(x) = \sum_{n\geq 0} a_n x^n$ is a rational function of the form

$$A(x) = \frac{P(x)}{Q(x)}$$

where $P(x)$ is a polynomial of degree $< k$.

(C) $A(x)$ can be written as a finite linear combination

$$A(x) = \sum_{i=1}^{t} \sum_{\ell=1}^{e_i} \frac{\alpha_{i,\ell}}{(1 - d_i x)^{\ell}}$$

with complex coefficients $\alpha_{i,\ell}$.

(D) There are polynomials $P_{i,\ell}(n)$ with $\deg P_{i,\ell} \leq e_i - 1$ such that

$$a_n = \sum_{i=1}^{t} \sum_{\ell=1}^{e_i} P_{i,\ell}(n)\, d_i^n \qquad (n \geq 0).$$

Example 6.2 (Catalan numbers). Let $C_0 = 1$ and for $n \geq 0$,

$$C_{n+1} = \sum_{i=0}^{n} C_i C_{n-i}.$$

Let $C(x) = \sum_{n\geq 0} C_n x^n$. Then

$$C(x) - 1 = xC(x)^2,$$

so $xC(x)^2 - C(x) + 1 = 0$. Solving this quadratic for $C(x)$ gives

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

where we choose the minus sign so that $C(0) = 1$. Expand $\sqrt{1 - 4x}$ using the extended binomial theorem:

$$(1 - 4x)^{1/2} = \sum_{n\geq 0} \binom{1/2}{n} (-4x)^n = \sum_{n\geq 0} \binom{2n}{n} \frac{(-1)^n}{4^n} (4x)^n = \sum_{n\geq 0} (-1)^n \binom{2n}{n} x^n.$$

From

$$C(x) = \frac{1 - \sqrt{1 - 4x}}{2x} = \sum_{n \geq 0} \frac{1}{n+1} \binom{2n}{n} x^n$$

we read off the closed form

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

6.5 Substitution Method

Recall a *derangement* of $[n] = \{1, 2, \ldots, n\}$ is a permutation with no fixed points. Let D_n denote the number of derangements of $[n]$. We proved previously that D_n satisfies

$$D_0 = 1, \qquad D_1 = 0, \qquad D_n = (n-1)\big(D_{n-1} + D_{n-2}\big) \ \ (n \geq 2).$$

We will obtain a closed form solution for D_n. First, we convert this into a first order recurrence.

Lemma 6.6 (A first-order recurrence)**.** For all $n \geq 1$,

$$D_n = nD_{n-1} + (-1)^n.$$

Proof. We prove the identity by induction on n.

For $n = 1$, we have $D_1 = 0$ and $1 \cdot D_0 + (-1)^1 = 1 - 1 = 0$, so the formula holds.

Assume $n \geq 2$ and that the statement holds for $n-1$, i.e. $D_{n-1} = (n-1)D_{n-2} + (-1)^{n-1}$. We have

$$D_n = (n-1)D_{n-1} + (n-1)D_{n-2}.$$

Now substitute $(n-1)D_{n-2} = D_{n-1} - (-1)^{n-1}$ from the induction hypothesis:

$$D_n = (n-1)D_{n-1} + \big(D_{n-1} - (-1)^{n-1}\big) = nD_{n-1} - (-1)^{n-1} = nD_{n-1} + (-1)^n,$$

as desired. □

We will now demonstrate the substitution method to convert this first order recurrence relation for D_n into a closed form.

Define $b_n := D_n/n!$. Then for all $n \geq 1$,

$$b_n = b_{n-1} + \frac{(-1)^n}{n!}, \qquad b_0 = 1.$$

Consequently,

$$b_n = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \qquad \text{and hence} \qquad D_n = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

Divide the first-order recurrence from the lemma by $n!$:

$$\frac{D_n}{n!} = \frac{nD_{n-1}}{n!} + \frac{(-1)^n}{n!} \quad \Longrightarrow \quad b_n = b_{n-1} + \frac{(-1)^n}{n!}.$$

Now telescope from 0 to n:

$$b_n = b_0 + \sum_{k=1}^{n} \frac{(-1)^k}{k!} = 1 + \sum_{k=1}^{n} \frac{(-1)^k}{k!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

Multiplying by $n!$ gives the closed form for D_n.

6.6 Stirling's formula

From the exact closed form

$$\frac{D_n}{n!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!},$$

we immediately see what happens as $n \to \infty$: the partial sums converge to the full exponential series

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

Hence

$$\frac{D_n}{n!} \longrightarrow \frac{1}{e} \qquad \text{and therefore} \qquad D_n \sim \frac{n!}{e}.$$

In words: *a uniformly random permutation of* $[n]$ *has probability tending to* $1/e$ *of having no fixed points*. Equivalently, D_n is asymptotically $n!/e$.

This naturally raises the next question: how large is $n!$ itself? The answer is given by *Stirling's formula*, which provides an extremely accurate approximation to factorial growth:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

Remark 6.2. Moreover, one can refine this to an asymptotic expansion (see the book):

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n} + \frac{1}{288n^2} + \cdots\right).$$

Example 6.3. Flip $2n$ fair coins independently, and let H be the number of heads. Then the probability of getting exactly n heads is

$$\mathbb{P}(H = n) = \frac{\binom{2n}{n}}{2^{2n}}.$$

Stirling's formula gives an asymptotic approximation for the probability

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

$$\frac{\left(\frac{2n}{e}\right)^{2n}}{\left(\frac{n}{e}\right)^{2n}} = 2^{2n} = 4^n$$

$$\frac{\sqrt{4\pi n}}{2\pi n} = \frac{1}{\sqrt{\pi n}}.$$

Therefore

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

Dividing by $2^{2n} = 4^n$ gives the desired asymptotic for the probability:

$$\mathbb{P}(H = n) = \frac{\binom{2n}{n}}{4^n} \sim \frac{1}{\sqrt{\pi n}}.$$

7 Ordinary generating functions

7.1 Why generating functions exist

In counting problems, we often have a family of numbers

$$a_0, a_1, a_2, \dots$$

where a_n counts "how many objects of size n" we have. A generating function is a way to package this entire infinite list into *one algebraic object* so that:

- algebraic operations (like adding or multiplying series) correspond to natural combinatorial operations (like disjoint union or building composite objects),
- extracting the coefficient of x^n recovers the quantity we care about.

The key idea is: instead of storing the number a_n in a sequence indexed by n, we store it as the coefficient of x^n in a formal series.

7.2 Combinatorial classes and weights

Definition 7.1 (Combinatorial class and weight). A *combinatorial class* is just a set $\mathcal{S}$ whose elements we think of as combinatorial objects (strings, subsets, graphs, partitions, etc.). A *weight function* on $\mathcal{S}$ is a map

$$w : \mathcal{S} \to \mathbb{Z}_{\geq 0}.$$

The intended meaning is that $w(s)$ measures the "size" of an object s.

We assume a mild finiteness condition: for every $n \geq 0$, there are only finitely many $s \in \mathcal{S}$ with $w(s) = n$.

7.3 Definition of the OGF

Definition 7.2 (Ordinary generating function (OGF)). Let $(\mathcal{S}, w)$ be a weighted combinatorial class. Its *ordinary generating function* is

$$S(x) \;:=\; \sum_{s \in \mathcal{S}} x^{w(s)}.$$

Define

$$a_n := \left|\{s \in \mathcal{S} : w(s) = n\}\right|.$$

Then grouping terms in the sum by weight gives the equivalent form

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

So the coefficient $[x^n]S(x)$ is exactly the number of objects of weight n.

Definition 7.3 (Formal power series). A *formal power series* over a ring R is an expression

$$A(x) = \sum_{n \geq 0} a_n x^n \qquad (a_n \in R),$$

where we treat x as an indeterminate and manipulate the series *purely algebraically* (we do not care about convergence).

We write $[x^n]A(x)$ for the coefficient of x^n.

The set of all formal power series in x with coefficients in $\mathbb{R}$ (or $\mathbb{C}$) forms an infinite-dimensional vector space and a commutative ring under coefficientwise addition and Cauchy product.

- The multiplicative identity is $1 = 1 + 0x + 0x^2 + \ldots$.
- A series $A(x)$ has a multiplicative inverse $A(x)^{-1}$ (i.e., there exists $B(x)$ with $A(x)B(x) = 1$) iff its constant term is nonzero:
$$[x^0]A(x) \neq 0.$$

7.4 Two fundamental combinatorial operations

The real power of OGFs is that basic constructions on classes correspond to simple algebra on generating functions.

7.4.1 Disjoint union ↔ addition

Definition 7.4 (Disjoint union of classes). If $\mathcal{A}$ and $\mathcal{B}$ are weighted classes with the same weight rule, their disjoint union $\mathcal{A} \sqcup \mathcal{B}$ consists of objects from either class, tagged by which class they came from, and the weight is preserved.

Proposition 7.1 (Addition rule). If $A(x)$ and $B(x)$ are the OGFs of $\mathcal{A}$ and $\mathcal{B}$, then the OGF of $\mathcal{A} \sqcup \mathcal{B}$ is

$$A(x) + B(x).$$

Proof. By definition,

$$\sum_{s \in \mathcal{A} \sqcup \mathcal{B}} x^{w(s)} = \sum_{a \in \mathcal{A}} x^{w(a)} + \sum_{b \in \mathcal{B}} x^{w(b)}.$$

□

7.4.2 Product construction ↔ multiplication (convolution)

Definition 7.5 (Product of classes). Given weighted classes $\mathcal{A}, \mathcal{B}$, define their product $\mathcal{A} \times \mathcal{B}$ to be the class of ordered pairs (a, b) with $a \in \mathcal{A}$, $b \in \mathcal{B}$, equipped with the additive weight

$$w(a, b) = w(a) + w(b).$$

Think: "build a composite object by choosing one A-object and one B-object, and size adds."

Proposition 7.2 (Multiplication rule / convolution). If $A(x) = \sum a_n x^n$ and $B(x) = \sum b_n x^n$ are the OGFs of $\mathcal{A}$ and $\mathcal{B}$, then the OGF of $\mathcal{A} \times \mathcal{B}$ is

$$A(x)B(x) = \sum_{n \geq 0} c_n x^n \qquad \text{where} \qquad c_n = \sum_{k=0}^{n} a_k b_{n-k}.$$

Equivalently, c_n counts pairs (a, b) with $w(a) + w(b) = n$.

Proof. Start from the definition of OGF:

$$\sum_{(a,b) \in \mathcal{A} \times \mathcal{B}} x^{w(a,b)} = \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} x^{w(a)+w(b)} = \left(\sum_{a \in \mathcal{A}} x^{w(a)} \right) \left(\sum_{b \in \mathcal{B}} x^{w(b)} \right) = A(x)B(x).$$

If we now write $A(x) = \sum a_n x^n$ and $B(x) = \sum b_n x^n$ and expand, the coefficient of x^n is exactly $\sum_{k=0}^{n} a_k b_{n-k}$. □

Example 7.1. Fix n. Let $\mathcal{S}$ be the class of all subsets $S \subseteq [n]$, and assign weight $w(S) = |S|$.

Step 1: describe a subset as a product of independent choices. For each element $i \in [n]$, we make an independent binary decision: either we *do not* include i (weight contribution 0), or we *do* include i (weight contribution 1). Thus, for a single element i, the local class is

$$\mathcal{S}_i = \{\text{exclude } i, \text{ include } i\}, \qquad S_i(x) = 1 + x.$$

Step 2: combine the n choices using the Product Rule. A subset of $[n]$ is exactly a choice from $\mathcal{S}_1 \times \cdots \times \mathcal{S}_n$, and weights add under products:

$$\mathcal{S} \;=\; \mathcal{S}_1 \times \cdots \times \mathcal{S}_n.$$

Therefore, by the Product Rule,

$$S(x) = \prod_{i=1}^{n} S_i(x) = (1 + x)^n.$$

Step 3: read off coefficients. Since $[x^k]S(x)$ counts weight-k objects, we get

$$[x^k]S(x) = \binom{n}{k},$$

the number of k-element subsets of $[n]$.

Example 7.2. Fix n. Let $\mathcal{M}$ be the class of all multisets of elements from $[n]$. Define the weight $w(M)$ to be the total size of the multiset (counting multiplicity).

Step 1: describe a multiset as n independent multiplicity choices. For each element type $i \in [n]$, we choose a multiplicity

$$m_i \in \{0, 1, 2, \dots\}.$$

Choosing multiplicity m_i contributes weight m_i. So the local class for a single type i is

$$\mathcal{M}_i = \{0, 1, 2, \dots\}, \qquad M_i(x) = \sum_{m \geq 0} x^m = \frac{1}{1-x}.$$

Step 2: combine types using the Product Rule. A multiset is exactly the data of $(m_1, \dots, m_n)$, i.e.

$$\mathcal{M} \ = \ \mathcal{M}_1 \times \cdots \times \mathcal{M}_n,$$

and the total size is $m_1 + \cdots + m_n$. Hence, by the Product Rule,

$$M(x) = \prod_{i=1}^{n} M_i(x) = \left(\frac{1}{1-x}\right)^n.$$

Step 3: interpret the coefficient (stars and bars). The coefficient $[x^k]M(x)$ counts n-tuples $(m_1, \dots, m_n)$ of nonnegative integers with sum k, i.e. the number of k-multisets from $[n]$. Thus

$$[x^k]M(x) = \binom{n+k-1}{n-1}.$$

7.5 Restricted multiplicities

Up to now, a multiset on $[n] = \{1, 2, \dots, n\}$ can be encoded by an n-tuple of multiplicities

$$(m_1, m_2, \dots, m_n), \qquad m_i \in \{0, 1, 2, \dots\},$$

where m_i is how many copies of element i appear. We use the weight

$$w(m_1, \dots, m_n) = m_1 + \cdots + m_n$$

(the total size, counting multiplicity). The ordinary generating function is

$$M(x) = \sum_{(m_1, \dots, m_n)} x^{w(m_1, \dots, m_n)}.$$

Why the "one-type factor" works. If we fix a single type i, and we allow it to appear with multiplicity in some set $B_i \subseteq \mathbb{N} = \{0, 1, 2, \dots\}$, then the contribution of type i alone is the series

$$A_i(x) = \sum_{b \in B_i} x^b,$$

because choosing multiplicity b contributes weight b.

Now the crucial point: the n choices of multiplicities are *independent* across types, and the total weight is the *sum* of the individual weights. Therefore, by the Product Rule (for weighted classes),

$$A(x) = \prod_{i=1}^{n} A_i(x).$$

The coefficient $[x^k]A(x)$ counts the number of allowed multisets of total size k.

Special case: the same restriction for every type. If every element type has the same allowed multiplicity set B, then $A_i(x) = A_{\text{one}}(x)$ for all i, and

$$A(x) = \left(A_{\text{one}}(x)\right)^n.$$

Example 7.3 (Even multiplicities)**.** Suppose each type must appear an even number of times:

$$B = \{0, 2, 4, 6, \dots\}.$$

For one type,

$$A_{\text{one}}(x) = \sum_{t \geq 0} x^{2t} = 1 + x^2 + x^4 + \cdots = \frac{1}{1 - x^2}.$$

Hence, by the Product Rule,

$$A(x) = \left(\frac{1}{1 - x^2}\right)^n.$$

Equivalently, $[x^k]A(x) = 0$ for odd k, and for even k it counts the number of ways to write k as a sum of n even nonnegative integers.

Example 7.4 (Multiplicity at least 2)**.** Suppose each type must appear at least twice:

$$B = \{2, 3, 4, \dots\}.$$

For one type,

$$A_{\text{one}}(x) = \sum_{b \geq 2} x^b = x^2(1 + x + x^2 + \cdots) = \frac{x^2}{1 - x}.$$

Therefore

$$A(x) = \left(\frac{x^2}{1 - x}\right)^n.$$

Here the factor x^{2n} is doing exactly what you think: it forces a baseline of 2 copies of each of the n types before any "extra" copies are distributed.

Fully general restriction. If each type i has its own allowed multiplicity set $B_i \subseteq \mathbb{N}$, then the one-type series is

$$A_i(x) = \sum_{b \in B_i} x^b,$$

and the total OGF for multisets respecting all restrictions is

$$A(x) = \prod_{i=1}^{n} A_i(x).$$

Example 7.5 (Making change (order does *not* matter)). Fix a finite set of coin denominations

$$D = \{d_1, d_2, \dots, d_r\} \subseteq \mathbb{Z}_{>0}.$$

We want to count the number of ways to make total value n using these coins, where a *way* means: for each denomination $d \in D$, we choose how many coins of value d we use (order is irrelevant).

Step 1: Define the combinatorial class. A choice of coins is exactly an r-tuple of nonnegative integers

$$(m_{d_1}, m_{d_2}, \dots, m_{d_r}) \in \mathbb{Z}_{\geq 0}^r,$$

where m_d means "how many d-coins we take." So define the class

$$\mathcal{C} = \{(m_d)_{d\in D} : m_d \in \mathbb{Z}_{\geq 0}\}.$$

Step 2: Define the weight function. The most natural weight is the *total value* of the chosen coins:

$$w\big((m_d)_{d\in D}\big) = \sum_{d\in D} d\, m_d.$$

For example, if $D = \{1, 2, 5\}$ and we choose $(m_1, m_2, m_5) = (2, 1, 1)$, then

$$w(2, 1, 1) = 1 \cdot 2 + 2 \cdot 1 + 5 \cdot 1 = 9.$$

Step 3: Define the OGF from the class and the weight. By definition, the ordinary generating function of $(\mathcal{C}, w)$ is

$$C(x) = \sum_{c\in\mathcal{C}} x^{w(c)}.$$

So the coefficient $[x^n]C(x)$ counts how many choices of coin multiplicities produce total value exactly n.

Step 4: Break the class into independent pieces (Product Rule). For each denomination $d \in D$, define the single-denomination class

$$\mathcal{C}_d = \{m_d : m_d \in \mathbb{Z}_{\geq 0}\},$$

with weight $w_d(m_d) = d\, m_d$.

Choosing a full coin-multiset is the same as choosing m_d independently for each $d \in D$, and the total value is the *sum* of the values contributed by each denomination. This is exactly the condition for the Product Rule.

Step 5: Compute each factor. For a fixed d, the OGF is

$$C_d(x) = \sum_{m_d \geq 0} x^{w_d(m_d)} = \sum_{m_d\geq 0} x^{d m_d} = 1 + x^d + x^{2d} + x^{3d} + \cdots = \frac{1}{1 - x^d}.$$

Step 6: Multiply the factors (Product Rule). Therefore the total OGF is

$$C(x) = \prod_{d\in D} C_d(x) = \prod_{d\in D} \frac{1}{1 - x^d}.$$

Conclusion:

$$[x^n] \prod_{d\in D} \frac{1}{1 - x^d}$$

equals the number of ways to make value n using denominations D, where order does *not* matter.

7.6 Bivariate OGFs (tracking two statistics)

Sometimes objects have two natural statistics (say size and number of parts). Then we use two variables and track both at once.

Definition 7.6 (Bivariate OGF)**.** Let $\mathcal{S}$ be a class with two weight functions $w_1 : \mathcal{S} \to \mathbb{Z}_{\geq 0}$ and $w_2 : \mathcal{S} \to \mathbb{Z}_{\geq 0}$. Its bivariate OGF is

$$S(x, y) = \sum_{s \in \mathcal{S}} x^{w_1(s)} y^{w_2(s)}.$$

Equivalently, if $a_{n,k}$ counts objects with $w_1 = n$ and $w_2 = k$, then

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

Example 7.6 (Pascal's identity via a bivariate OGF)**.** Let $\mathcal{S}$ be the class of pairs (n, S) where $n \geq 0$ and $S \subseteq [n]$. Define two weights:

$$w_1(n, S) = n, \qquad w_2(n, S) = |S|.$$

Then the bivariate OGF is

$$S(x, y) = \sum_{n \geq 0} \sum_{S \subseteq [n]} x^n y^{|S|} = \sum_{n \geq 0} x^n (1 + y)^n = \frac{1}{1 - x(1 + y)}.$$

Expanding coefficients gives $[x^n y^k] S(x, y) = \binom{n}{k}$.
Now the identity

$$(1 - x - xy)\, S(x, y) = 1$$

implies that the coefficient of $x^n y^k$ for $n, k \geq 1$ is zero, i.e.

$$\binom{n}{k} - \binom{n-1}{k} - \binom{n-1}{k-1} = 0,$$

which is Pascal's identity.

7.7 Extracting coefficients

Definition 7.7. Let

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

be an ordinary generating function. The *formal derivative* of A is defined term-by-term by

$$A'(x) = \sum_{n \geq 1} n a_n x^{n-1}.$$

Everything here is purely algebraic: we are differentiating a *formal* power series, so there are no convergence assumptions.

Basic rules. The formal derivative satisfies the usual identities:

$$(A+B)' = A' + B', \qquad (AB)' = A'B + AB'.$$

Multiplying by x shifts the exponents back up:

$$xA'(x) = \sum_{n\geq 1} n a_n x^n.$$

So for every $n \geq 0$,

$$[x^n]\big(xA'(x)\big) = n\, a_n.$$

In words: $x\frac{d}{dx}$ *multiplies the* n*th coefficient by* n.

Example 7.7 (Differentiating the geometric series)**.**

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

Start with the geometric series identity

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

Differentiate formally:

$$\left(\frac{1}{1-x}\right)' = \frac{1}{(1-x)^2} \qquad \text{and} \qquad \left(\sum_{n\geq 0} x^n\right)' = \sum_{n\geq 1} n x^{n-1}.$$

Hence

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

Multiplying by x gives the cleaner, more commonly used form

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

(And if you want the sum from $n \geq 0$, the $n = 0$ term is just 0 anyway.)

Example 7.8. For each integer $m \geq 1$,

$$\frac{1}{(1-x)^m} = \sum_{n\geq 0} \binom{n+m-1}{m-1} x^n.$$

One way to prove this is by induction on m: the case $m = 1$ is the geometric series, and differentiating both sides of the mth case produces the $(m+1)$st case after a short coefficient simplification.

7.8 Shifting indices

Given an OGF

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

multiplication by a power of x shifts the coefficients:

$$x^k A(x) = \sum_{n\geq 0} a_n x^{n+k} = \sum_{t\geq k} a_{t-k} x^t.$$

Thus

$$[x^t]\big(x^k A(x)\big) = \begin{cases} a_{t-k}, & t \geq k, \\ 0, & t < k. \end{cases}$$

Example 7.9.

$$\sum_{t\geq k} \binom{t}{k} x^t = \frac{x^k}{(1-x)^{k+1}}.$$

Recall

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots,$$

so

$$\frac{1}{(1-x)^{k+1}} = \left(1 + x + x^2 + \cdots\right)^{k+1}.$$

When we expand this product, choosing x^{a_i} from the ith factor produces the monomial $x^{a_0+\cdots+a_k}$. Hence the coefficient of x^n counts the number of $(k+1)$-tuples $(a_0, \ldots, a_k) \in \mathbb{Z}_{\geq 0}^{k+1}$ with $a_0 + \cdots + a_k = n$, which is $\binom{n+k}{k}$ by stars and bars. Therefore

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

Multiplying by x^k shifts every exponent up by k:

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

Writing $t = n + k$ (so $t \geq k$) gives $\binom{n+k}{k} = \binom{t}{k}$, so

$$\sum_{t\geq k} \binom{t}{k} x^t = \frac{x^k}{(1-x)^{k+1}}.$$

7.9 OGF Vandermonde convolution

Theorem 7.3 (Vandermonde convolution)**.** For all integers $m, n, r \geq 0$,

$$\sum_k \binom{m}{k}\binom{n}{r-k} = \binom{m+n}{r}.$$

Generating functions as subset-counters. Let A and B be disjoint sets with $|A| = m$ and $|B| = n$. For a set A, the polynomial

$$(1+x)^m = \prod_{a\in A}(1+x)$$

encodes choosing a subset of A: from each element we either choose it (contributing a factor x) or do not choose it (contributing 1). Thus $[x^k](1+x)^m = \binom{m}{k}$ counts k-subsets of A.

Similarly, $[x^j](1+x)^n = \binom{n}{j}$ counts j-subsets of B. Therefore the product $(1+x)^m(1+x)^n$ encodes choosing a subset of $A \cup B$ by choosing independently a subset of A and a subset of B. To obtain a subset of total size r, we must choose k elements from A and $r-k$ elements from B, which can be done in $\binom{m}{k}\binom{n}{r-k}$ ways. Summing over all k gives

$$[x^r]\big((1+x)^m(1+x)^n\big) = \sum_k \binom{m}{k}\binom{n}{r-k}.$$

On the other hand, $A \cup B$ has $m+n$ elements, so

$$(1+x)^{m+n} = \prod_{u\in A\cup B}(1+x)$$

and $[x^r](1+x)^{m+n} = \binom{m+n}{r}$ counts r-subsets of $A \cup B$. Since $(1+x)^m(1+x)^n = (1+x)^{m+n}$, the coefficients of x^r agree, yielding Vandermonde's identity. □

7.10 Catalan recurrence

Let $(C_n)_{n\geq 0}$ be the Catalan numbers defined by $C_0 = 1$ and the recurrence

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}.$$

Theorem 7.4. The OGF $C(x) = \sum_{n\geq 0} C_n x^n$ satisfies

$$C(x) = 1 + xC(x)^2.$$

Equivalently, the recurrence above holds for all $n \geq 0$.

Proof. Using the product rule for OGFs, we have

$$C(x)^2 = \sum_{n\geq 0}\Big(\sum_{k=0}^{n} C_k C_{n-k}\Big)x^n.$$

Multiplying by x shifts indices:

$$xC(x)^2 = \sum_{n\geq 0}\Big(\sum_{k=0}^{n} C_k C_{n-k}\Big)x^{n+1} = \sum_{m\geq 1}\Big(\sum_{k=0}^{m-1} C_k C_{m-1-k}\Big)x^m.$$

Now

$$C(x) = C_0 + \sum_{m\geq 1} C_m x^m = 1 + \sum_{m\geq 1} C_m x^m.$$

Hence the functional equation $C(x) = 1 + xC(x)^2$ is equivalent (by coefficient comparison) to

$$C_m = \sum_{k=0}^{m-1} C_k C_{m-1-k} \quad (m \geq 1),$$

which is the recurrence. □

7.11 How to manipulate OGFs for coefficients?

Theorem 7.5. 1. For all n,

$$b_n = \begin{cases} a_{n-r}, & n \geq r, \\ 0, & n < r, \end{cases} \qquad \Longleftrightarrow \qquad B(x) = x^r A(x).$$

2. For all n,

$$b_n = n\, a_n \qquad \Longleftrightarrow \qquad B(x) = x\, A'(x).$$

3. For all n,

$$c_n = \sum_{i=0}^{n} a_i \qquad \Longleftrightarrow \qquad C(x) = \frac{A(x)}{1-x} = A(x)\,(1 + x + x^2 + \cdots).$$

4. (Even/odd parts.)

$$b_n = \begin{cases} a_n, & n \text{ even}, \\ 0, & n \text{ odd}, \end{cases} \qquad \Longleftrightarrow \qquad B(x) = \frac{A(x) + A(-x)}{2},$$

$$b_n = \begin{cases} a_n, & n \text{ odd}, \\ 0, & n \text{ even}, \end{cases} \qquad \Longleftrightarrow \qquad B(x) = \frac{A(x) - A(-x)}{2}.$$

5. Let $m \geq 1$. For all n,

$$b_n = \begin{cases} a_{n/m}, & m \mid n, \\ 0, & m \nmid n, \end{cases} \qquad \Longleftrightarrow \qquad B(x) = A(x^m).$$

Example 7.10.

$$\sum_{k=0}^{n} k \binom{n}{k} = n 2^{n-1}.$$

Let

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

Then

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

Comparing coefficients of x^k gives

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

Example 7.11.

$$\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

Recall

$$\sum_{k\geq 0} x^k = \frac{1}{1-x}.$$

Differentiate and multiply by x:

$$\sum_{k\geq 0} kx^k = \frac{x}{(1-x)^2}.$$

Differentiate once more and multiply by x:

$$\sum_{k\geq 0} k^2x^k = x\left(\frac{x}{(1-x)^2}\right)' = \frac{x+x^2}{(1-x)^3} =: A(x).$$

Let $a_k = k^2$, so $A(x) = \sum_{k\geq 0} a_k x^k$. Define

$$C(x) := \frac{A(x)}{1-x} = \sum_{n\geq 0} c_n x^n,$$

then

$$c_n = \sum_{k=0}^{n} a_k = \sum_{k=0}^{n} k^2 = [x^n]\, C(x).$$

Since

$$C(x) = \frac{x+x^2}{(1-x)^4} = x(1-x)^{-4} + x^2(1-x)^{-4},$$

we get

$$[x^n]\, C(x) = \binom{n-1+3}{3} + \binom{n-2+3}{3} = \frac{n(n+1)(2n+1)}{6}.$$

Hence

$$\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

Example 7.12. We will extract the *even-index* part of $(1+x)^n$, i.e. find a closed form for

$$E(x) := \sum_{i\geq 0} \binom{n}{2i} x^{2i},$$

and as a consequence compute $\sum_{i\geq 0} \binom{n}{2i}$.

Let

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

Then the even–index part is

$$\sum_{i\ge 0} \binom{n}{2i} x^{2i} = \frac{1}{2}\big((1+x)^n + (1-x)^n\big),$$

since the odd–index terms cancel when we add $(1+x)^n$ and $(1-x)^n$.

Setting $x = 1$ (and assuming $n \ge 1$) gives

$$\sum_{i\ge 0} \binom{n}{2i} = \frac{1}{2}\big((1+1)^n + (1-1)^n\big) = \frac{1}{2}\,(2^n + 0) = 2^{n-1}.$$

(For $n = 0$ the sum is $\binom{0}{0} = 1$.)

Combinatorial proof. There are 2^n subsets of $[n] := \{1, \ldots, n\}$. Pair each subset $S \subseteq [n]$ with $S \triangle \{n\}$ (toggle the element n); in each pair exactly one subset has even cardinality and one has odd cardinality. Thus exactly half of all subsets are even, so the number of even subsets is 2^{n-1}, i.e.

$$\sum_{i\ge 0} \binom{n}{2i} = 2^{n-1}.$$

Example 7.13. Evaluate the sum

$$S_n := \sum_{k=0}^{n} k(n-k)$$

in closed form.

Let

$$a_k = b_k = k.$$

Then

$$A(x) = B(x) = \sum_{k\ge 0} k x^k = \frac{x}{(1-x)^2},$$

so

$$C(x) := A(x)B(x) = \frac{x^2}{(1-x)^4} = \sum_{k\ge 0} \binom{k+4-1}{4-1} x^{k+2}.$$

Thus

$$C_n = \sum_{k=0}^{n} k\,(n-k) = [x^n]\,C(x) = \binom{(n-2)+3}{3} = \binom{n+1}{3}.$$

Example 7.14. Prove the identity

$$S_n := \sum_{k=0}^{n} \binom{2k}{k}\binom{2n-2k}{n-k} = 4^n.$$

Consider

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

Let

$$A(x) := \sum_{k\geq 0} \binom{2k}{k} x^k.$$

Let the Catalan generating function be

$$C(x) := \sum_{k\geq 0} C_k x^k = \sum_{k\geq 0} \frac{1}{k+1}\binom{2k}{k} x^k = \frac{1-\sqrt{1-4x}}{2x}.$$

Then

$$xC(x) = \frac{1-\sqrt{1-4x}}{2} \quad\Longrightarrow\quad A(x) = (xC(x))' = (1-4x)^{-1/2}.$$

Hence

$$A(x)^2 = (1-4x)^{-1} \quad\Longrightarrow\quad [x^n]\,A(x)^2 = [x^n]\frac{1}{1-4x} = 4^n.$$

But

$$A(x)^2 = \left(\sum_{k\geq 0}\binom{2k}{k}x^k\right)\left(\sum_{j\geq 0}\binom{2j}{j}x^j\right) = \sum_{n\geq 0}\left(\sum_{k=0}^{n}\binom{2k}{k}\binom{2n-2k}{n-k}\right)x^n,$$

so

$$S_n = \sum_{k=0}^{n}\binom{2k}{k}\binom{2n-2k}{n-k} = [x^n]\,A(x)^2 = 4^n.$$

7.12 Snake Oil

Herbert Wilf's *snake oil* method is a reliable trick for evaluating sums that involve binomial coefficients. The slogan is:

> *If you can sum it, you can generate it.*

Concretely, suppose you have a sequence defined by a sum

$$a_n := \sum_{k\geq 0} T(n,k),$$

where $T(n,k)$ is some expression in n and k (often binomial coefficients). The method is:

1. Form the ordinary generating function (OGF)

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

2. Substitute the definition of a_n and (formally) swap the order of summation:

$$A(x) = \sum_{n\geq 0}\sum_{k\geq 0} T(n,k)x^n = \sum_{k\geq 0}\sum_{n\geq 0} T(n,k)x^n.$$

   This is legal in *formal* power series.

3. For each fixed k, rewrite $\sum_{n\geq 0} T(n,k)x^n$ into a known closed form. This usually requires a reindexing so that the binomial coefficient becomes something like $\binom{m+k}{k}$, whose OGF we know.

4. After that, the k–sum typically becomes a geometric series. Finish the algebra, get a closed form for $A(x)$, and then identify its coefficients using a known generating function (or a recurrence derived from the denominator).

Consider the following example to illustrate the Snake Oil method.

Example 7.15. Define

$$a_n := \sum_{k\geq 0}\binom{n-k}{k} \qquad (n \geq 0),$$

where we adopt the usual convention $\binom{r}{k} = 0$ if $r < k$ or $r < 0$. We will show

$$a_n = F_{n+1},$$

where $F_0 = 0$, $F_1 = 1$, and $F_{m+2} = F_{m+1} + F_m$.

The term $\binom{n-k}{k}$ is nonzero only when $n - k \geq k$, i.e. $n \geq 2k$. So for fixed n, the sum is actually finite: $0 \leq k \leq \lfloor n/2 \rfloor$. This is exactly the kind of shape snake oil likes, because the constraint $n \geq 2k$ suggests substituting $n = m + 2k$.

Step 1: build the OGF and swap sums. Let

$$A(x) := \sum_{n\geq 0} a_n x^n = \sum_{n\geq 0}\sum_{k\geq 0}\binom{n-k}{k}x^n.$$

Swap the order:

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

Step 2: reindex to remove the constraint $n \geq 2k$. For a fixed k, the inner term is nonzero only when $n \geq 2k$. Write

$$n = m + 2k \qquad (m \geq 0).$$

Then $n - k = m + k$, so

$$\binom{n-k}{k} = \binom{m+k}{k}, \qquad x^n = x^{m+2k} = x^{2k}x^m.$$

Thus

$$A(x) = \sum_{k\geq 0}\sum_{m\geq 0}\binom{m+k}{k}x^{m+2k} = \sum_{k\geq 0} x^{2k}\sum_{m\geq 0}\binom{m+k}{k}x^m.$$

Step 3: use a known binomial OGF.

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

can be proved by stars and bars. Plugging this in gives

$$A(x) = \sum_{k\geq 0}\frac{x^{2k}}{(1-x)^{k+1}} = \frac{1}{1-x}\sum_{k\geq 0}\left(\frac{x^2}{1-x}\right)^k.$$

Step 4: finish with a geometric series. Since $\sum_{k\geq 0} r^k = \frac{1}{1-r}$ in formal power series,

$$A(x) = \frac{1}{1-x}\cdot\frac{1}{1-\frac{x^2}{1-x}} = \frac{1}{1-x-x^2}.$$

Step 5: recognize Fibonacci. The Fibonacci numbers satisfy

$$\sum_{n\geq 0} F_{n+1}x^n = \frac{1}{1-x-x^2},$$

so comparing coefficients yields $a_n = F_{n+1}$ for all $n \geq 0$. Equivalently,

$$\sum_{k\geq 0}\binom{n-k}{k} = F_{n+1}.$$

Proposition 7.6 (A Delannoy identity)**.** For all integers $m, n \geq 0$,

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

Proof. Fix m and let

$$L_n := \sum_{k\geq 0}\binom{m}{k}\binom{n+m-k}{m}, \qquad R_n := \sum_{k\geq 0}\binom{m}{k}\binom{n}{k}2^k.$$

We show that (L_n) and (R_n) have the same generating function in n.

Left-hand side. Let

$$L(x) := \sum_{n\geq 0} L_n x^n = \sum_{n\geq 0}\sum_{k\geq 0}\binom{m}{k}\binom{n+m-k}{m}x^n.$$

Swap sums and use the change of variable $N = n+m-k$:

$$L(x) = \sum_{k\geq 0}\binom{m}{k}\sum_{n\geq 0}\binom{n+m-k}{m}x^n = \sum_{k\geq 0}\binom{m}{k}\sum_{N\geq m}\binom{N}{m}x^{N-k}.$$

Thus

$$L(x) = \sum_{k\geq 0}\binom{m}{k}x^{-k}\Big(\sum_{N\geq m}\binom{N}{m}x^N\Big).$$

Using

$$\sum_{N\geq m}\binom{N}{m}x^N = \frac{x^m}{(1-x)^{m+1}},$$

we get

$$L(x) = \frac{x^m}{(1-x)^{m+1}}\sum_{k\geq 0}\binom{m}{k}x^{-k} = \frac{x^m}{(1-x)^{m+1}}(1+x^{-1})^m = \frac{(1+x)^m}{(1-x)^{m+1}}.$$

Right-hand side. Similarly, let

$$R(x) := \sum_{n\geq 0}R_n x^n = \sum_{n\geq 0}\sum_{k\geq 0}\binom{m}{k}\binom{n}{k}2^k x^n.$$

Swap sums:

$$R(x) = \sum_{k\geq 0}\binom{m}{k}2^k\sum_{n\geq 0}\binom{n}{k}x^n.$$

We use the standard generating function

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

to get

$$R(x) = \sum_{k\geq 0}\binom{m}{k}2^k\frac{x^k}{(1-x)^{k+1}} = \frac{1}{1-x}\sum_{k\geq 0}\binom{m}{k}\left(\frac{2x}{1-x}\right)^k = \frac{1}{1-x}\left(1+\frac{2x}{1-x}\right)^m = \frac{(1+x)^m}{(1-x)^{m+1}}.$$

Since $L(x) = R(x)$, the sequences (L_n) and (R_n) are identical, and the stated identity follows. □

Proposition 7.7. Let $c(n,k)$ denote the number of permutations of $[n]$ with exactly k cycles (Stirling numbers of the first kind). Then for integers $n \geq m \geq 0$,

$$\sum_{k=m}^{n} c(n,k)\binom{k}{m} = c(n+1, m+1).$$

Proof. Recall the permutation cycle generating function identity

$$\sum_{k=0}^{n} c(n,k)x^k = x^{(n)} := x(x+1)\cdots(x+n-1),$$

the rising factorial.

Consider

$$\sum_{k=m}^{n} c(n,k)\binom{k}{m}x^m.$$

Using the binomial expansion of $(1+x)^k$ we have

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

so

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

Now use the generating function with x replaced by $(1+x)$:

$$\sum_{k=0}^{n} c(n,k)(1+x)^k = (1+x)^{(n)} = (1+x)(2+x)\cdots(n+x).$$

But

$$(1+x)^{(n)} = \frac{x^{(n+1)}}{x} = \frac{1}{x}\sum_{m=0}^{n+1} c(n+1,m)x^m.$$

Taking the coefficient of x^m on both sides,

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

as claimed. □

8 Permutations statistics

8.1 Inversions

Definition 8.1. The *symmetric group* S_n is the set of all permutations of $\{1, 2, \dots, n\}$, with composition as the group operation.

Definition 8.2 (One–line notation). A permutation $\pi \in S_n$ can be written as the word

$$(\pi(1), \pi(2), \dots, \pi(n)).$$

This is called *one–line notation*.

Definition 8.3 (Inversion). Let $\pi \in S_n$. An *inversion* of π is a pair (i, j) with $1 \le i < j \le n$ such that $\pi(i) > \pi(j)$. Denote by $\mathrm{inv}(\pi)$ the number of inversions of π.

Example 8.1. For $\pi = (3, 1, 4, 2) \in S_4$ (in one–line notation), the inversions are

$$(1, 2), (1, 4), (3, 4),$$

so $\mathrm{inv}(\pi) = 3$.

Definition 8.4 (Inversion generating polynomial). For $n \ge 0$ define

$$A_n(x) = \sum_{\pi \in S_n} x^{\mathrm{inv}(\pi)}.$$

Theorem 8.1. For every $n \ge 1$,

$$A_n(x) = \prod_{i=1}^{n} (1 + x + \cdots + x^{i-1}).$$

Proof. **Goal.** We prove a recurrence

$$A_n(x) = A_{n-1}(x)\,(1 + x + \cdots + x^{n-1}),$$

and then iterate it.

Step 1: insert n into a permutation of $[n-1]$. Fix $\sigma \in S_{n-1}$ and write it in one–line form

$$(\sigma(1), \sigma(2), \dots, \sigma(n-1)).$$

We create a permutation $\pi \in S_n$ by inserting the symbol n into this word. There are exactly n insertion slots: before the first entry, between consecutive entries, or after the last entry.

Step 2: count how many *new* inversions are created. Since n is the largest value, it can only create inversions where it appears *on the left*. If we insert n so that there are r elements to its right, then n forms an inversion with each of those r elements (because every one of them is $< n$). Therefore the number of new inversions created is exactly r.

As we vary the insertion slot from "far right" to "far left", the number r runs through

$$0, 1, 2, \ldots, n-1.$$

Step 3: translate this into generating functions. For the fixed σ, the n resulting permutations contribute

$$x^{\mathrm{inv}(\sigma)}(1 + x + \cdots + x^{n-1})$$

to $A_n(x)$, because we add $0, 1, \ldots, n-1$ inversions depending on the slot.

Summing over all $\sigma \in S_{n-1}$ gives

$$A_n(x) = \sum_{\sigma \in S_{n-1}} x^{\mathrm{inv}(\sigma)}(1 + x + \cdots + x^{n-1}) = A_{n-1}(x)(1 + x + \cdots + x^{n-1}).$$

Step 4: solve the recurrence. Since $A_1(x) = 1$, iterating yields

$$A_n(x) = \prod_{i=1}^{n}(1 + x + \cdots + x^{i-1}),$$

as claimed. □

Definition 8.5 (Two–line and word form). A permutation $\pi \in S_n$ can be written in *two–line form*

$$\pi = \begin{pmatrix} 1 & 2 & \ldots & n \\ \pi(1) & \pi(2) & \ldots & \pi(n) \end{pmatrix}$$

or in *word form* as

$$(\pi(1), \pi(2), \ldots, \pi(n)).$$

8.2 Permutation Cycles

Definition 8.6 (Cycle decomposition). A permutation $\pi \in S_n$ decomposes uniquely into *disjoint cycles*. For example,

$$\pi = (1\,4\,7\,3)(2)(5\,8)(6\,9)$$

means $\pi(1) = 4$, $\pi(4) = 7$, $\pi(7) = 3$, $\pi(3) = 1$, and so on. Cycles are disjoint, so the order of writing them does not affect the permutation.

Definition 8.7 (Canonical cycle representation). To make the cycle notation unique as a *written string*, we impose conventions:

- within each cycle, write the smallest element first;
- order the cycles by increasing smallest element.

The resulting cycle product is the *canonical cycle representation*.

Lemma 8.2 (Uniqueness of canonical cycle form)**.** Every permutation $\pi \in S_n$ has a unique canonical cycle representation.

Proof. The disjoint cycle decomposition of π is unique up to: (i) rotating the entries within each cycle, and (ii) permuting the order of the cycles. Putting the smallest element first in each cycle fixes (i), and ordering cycles by their smallest elements fixes (ii). □

Definition 8.8 (Unsigned Stirling numbers of the first kind)**.** Let $c(n, k)$ be the number of permutations in S_n having exactly k cycles in their (equivalently, any) disjoint cycle decomposition. These are the *unsigned Stirling numbers of the first kind*.

Definition 8.9 (Cycle index polynomial)**.** For $n \geq 1$, define

$$C_n(x) = \sum_{k=1}^{n} c(n, k)\, x^k.$$

Thus $[x^k]C_n(x) = c(n, k)$ counts permutations of $[n]$ with exactly k cycles.

Theorem 8.3 (Product formula for cycle counts)**.** For all $n \geq 1$,

$$C_n(x) = x(x + 1)(x + 2) \cdots (x + n - 1).$$

Equivalently, the coefficient of x^k in $x(x + 1) \cdots (x + n - 1)$ is $c(n, k)$.

Proof. We prove the polynomial identity by counting the same set of objects in two ways.

Objects being counted. Fix a positive integer x. A *coloured permutation* means: take $\pi \in S_n$ and assign to each cycle one of x colours (colours may repeat between cycles).

First count (group by number of cycles). If π has k cycles, then there are x^k ways to colour those cycles. Therefore

$$\#\{\text{coloured permutations of } [n]\} = \sum_{k=1}^{n} c(n, k)\, x^k = C_n(x).$$

Second count (insert m into a coloured permutation on $[m - 1]$). We build the coloured permutation by adding elements $1, 2, \ldots, n$ one at a time.

Start with $m = 1$. The only permutation is the 1–cycle (1), and we may choose its colour in x ways. So there are x possibilities.

Now suppose we have already formed a coloured permutation on $\{1, \ldots, m - 1\}$. We insert m as follows:

- *Start a new cycle:* create the 1–cycle (m) and choose its colour. This gives x possibilities.
- *Insert into an existing cycle:* in cycle notation, inserting m means: choose an existing element $t \in \{1, \ldots, m - 1\}$ and declare that m comes *right after* t in its cycle. There are exactly $m - 1$ choices of t. (No new colour choice is needed because we are not creating a new cycle.)

Hence, at step m, there are exactly $x + (m - 1)$ ways to place m.
Multiplying over $m = 1, 2, \ldots, n$ gives

$$\#\{\text{coloured permutations of } [n]\} = x(x+1)(x+2)\cdots(x+n-1).$$

□

8.3 Eulerian numbers

This section is about a different permutation statistic than inversions. Instead of counting pairs out of order, we count the places where a permutation *drops* when written in one–line notation.

Definition 8.10 (Descents). For $\pi \in S_n$ the *descent set* of π is

$$\mathrm{Des}(\pi) = \{i \in \{1, \dots, n-1\} : \pi(i) > \pi(i+1)\}.$$

The number of descents of π is $|\mathrm{Des}(\pi)|$.

Example 8.2. If $\pi = (3, 1, 4, 2) \in S_4$, then

$$3 > 1 \Rightarrow 1 \in \mathrm{Des}(\pi), \qquad 1 < 4 \Rightarrow 2 \notin \mathrm{Des}(\pi), \qquad 4 > 2 \Rightarrow 3 \in \mathrm{Des}(\pi),$$

so $\mathrm{Des}(\pi) = \{1, 3\}$ and $\mathrm{Des}(\pi) = 2$.

Definition 8.11 (Eulerian numbers and polynomials). For $0 \le k \le n-1$ let $A(n,k)$ be the number of permutations $\pi \in S_n$ with exactly k descents. The numbers $A(n,k)$ are the *Eulerian numbers*. The associated *Eulerian polynomial* is Define the *Eulerian polynomial*

$$E_n(x) := \sum_{k=0}^{n-1} A(n,k)\, x^k.$$

8.4 Worpitzky's Identity

Theorem 8.4 (Worpitzky, 1883). For every integer $n \ge 1$ and every positive integer x,

$$x^n = \sum_{k \ge 0} A(n,k) \binom{x+k}{n},$$

where we understand $A(n,k) = 0$ for $k \notin \{0, \dots, n-1\}$.

The identity is best understood as a change of basis: the polynomials $\binom{x}{n}, \binom{x+1}{n}, \binom{x+2}{n}, \dots$ form a very natural "binomial basis" for degree-n polynomials, and Worpitzky says the coefficients of x^n in that basis are the Eulerian numbers.

Proof of Worpitzky's Theorem. We give a combinatorial proof by double counting.

Fix n and $x \in \mathbb{N}$. On the left-hand side, x^n counts the number of functions

$$f : [n] \longrightarrow [x] = \{1, 2, \dots, x\},$$

since each of the n elements may be sent independently to one of x values.

We will now classify such functions according to a permutation with k descents and an additional choice counted by $\binom{x+k}{n}$.

Given a function $f : [n] \to [x]$, group together elements with the same image and, inside each fibre, arrange the elements in increasing order. If we then list the fibres in increasing order of the value in $[x]$, and concatenate these increasing lists, we obtain a word

$$\pi(1)\, \pi(2)\, \dots\, \pi(n)$$

which is a permutation $\pi \in S_n$ written in word form. The places where a new fibre begins are exactly the positions where the sequence of values $f(\pi(i))$ increases. Between two fibres the function values strictly increase; inside a fibre the values are equal.

The permutation π decomposes uniquely into *increasing runs* (maximal consecutive segments on which π is increasing). The number of such runs is $k+1$ if π has k descents, and the boundaries between runs occur precisely at the positions of descents.

Conversely, given a permutation with k descents, the positions between runs are fixed, but we are free to assign the function values in $[x]$ to each run, as long as they weakly increase from run to run.

Fix a permutation $\pi \in S_n$ with k descents, hence $k+1$ increasing runs. To obtain f, we must assign to each run a value in $[x]$ so that the run values form a weakly increasing sequence of length $k+1$. Let these run values be $1 \le v_1 \le v_2 \le \cdots \le v_{k+1} \le x$. We may think of these as choosing $k+1$ (not necessarily distinct) numbers between 1 and x in weakly increasing order.

By the standard "stars and bars" bijection, such weakly increasing sequences are in bijection with subsets of size n in a set of size $x+k$; more concretely, they are counted by

$$\binom{x+k}{n}.$$

(Equivalently, we can encode the k increases between successive run values by inserting k bars among x positions.)

Thus, for a fixed π with k descents, there are $\binom{x+k}{n}$ ways to choose f whose sorted word yields π.

For each k, there are $A(n,k)$ permutations of S_n with exactly k descents. Each such permutation corresponds to exactly $\binom{x+k}{n}$ functions $f : [n] \to [x]$. Therefore

$$x^n = \#\{f : [n] \to [x]\} = \sum_{k=0}^{n-1} A(n,k)\binom{x+k}{n},$$

which proves the identity. □

For completeness we record a generating–function reformulation of Worpitzky's identity.

Definition 8.12. Fix $n \ge 1$ and set

$$A_n(x) = \sum_{k\ge 0} A(n,k)x^k, \qquad C_n(x) = \sum_{m\ge 0}\binom{m+n}{n}x^m = \frac{1}{(1-x)^{n+1}}.$$

Proposition 8.5. For every $n \ge 1$,

$$\frac{A_n(x)}{(1-x)^{n+1}} = \sum_{m\ge 0} m^n x^m.$$

Equivalently,

$$A_n(x) = (1-x)^{n+1}\sum_{m\ge 0} m^n x^m,$$

and expanding $(1-x)^{n+1}$ recovers Worpitzky's identity.

Proof. Start from Worpitzky's identity in the form

$$m^n = \sum_{k\geq 0} A(n,k)\binom{m+k}{n} \qquad (m \in \mathbb{N}).$$

Multiply both sides by x^m and sum over $m \geq 0$. On the right-hand side, interchange the order of summation and use

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

to obtain

$$\sum_{m\geq 0} m^n x^m = C_n(x)\sum_{k\geq 0} A(n,k)x^k = \frac{A_n(x)}{(1-x)^{n+1}}.$$

Rearranging gives the desired formula. □

9 Exponential generating functions

9.1 Why EGFs exist

Ordinary generating functions (OGFs) are the right language when we are counting *unlabeled* objects by a size/weight: if an object has weight n, it contributes x^n.

Exponential generating functions (EGFs) are the right language when our objects are *labeled*: the labels matter, and when we build larger objects by combining smaller ones, we must count the ways to distribute labels among the parts. That label bookkeeping is exactly where the factorials and binomial coefficients come from.

Definition 9.1 (Labeled object)**.** Fix $n \geq 0$ and let $[n] = \{1, 2, \ldots, n\}$ (with $[0] = \varnothing$).

A *labeled object of size n* is a pair (S, Φ) where:

- S is an underlying "shape" with exactly n distinguished *atoms* (the pieces that are being labeled), and
- $\Phi : \mathrm{At}(S) \longrightarrow [n]$ is a bijection.

In other words, the labels $1, 2, \ldots, n$ are assigned to the atoms of S in a one-to-one way: every atom gets exactly one label, no label is repeated, and every label is used exactly once.

Definition 9.2. A *labeled combinatorial class* $\mathcal{A}$ is specified by giving, for each $n \geq 0$, a set $\mathcal{A}[n]$ of labeled objects of size n (i.e. pairs (S, Φ) as above). We define the counting sequence

$$a_n := |\mathcal{A}[n]|.$$

Remark 9.1 (Graphs: labels are vertex names)**.** For graphs, the *atoms* are the vertices. An unlabeled graph-shape is $S = (V, E)$ with $|V| = n$. A labeling is a bijection

$$\Phi : V \longrightarrow [n],$$

i.e. each vertex gets a unique name $1, 2, \ldots, n$ (no repeats, all used). Thus a labeled graph is the pair (S, Φ).

Because Φ is a bijection, there are exactly $n!$ possible labelings of a fixed n-vertex shape S. This "n distinct vertices $\Rightarrow$ $n!$ ways to name them" is the basic source of the factorials in EGFs.

Definition 9.3 (Exponential generating function)**.** The *exponential generating function* (EGF) of the labeled class $\mathcal{A}$ (equivalently, of the sequence $(a_n)_{n\geq 0}$) is

$$A(x) := \sum_{n\geq 0} a_n \frac{x^n}{n!}.$$

Definition 9.4 (Disjoint union of labeled classes)**.** Let $\mathcal{A}$ and $\mathcal{B}$ be labeled combinatorial classes, and assume they are disjoint as sets of objects (no object belongs to both classes). Define their *disjoint union* (or *sum*) to be

$$\mathcal{C} := \mathcal{A} \uplus \mathcal{B},$$

the class consisting of objects that are either an $\mathcal{A}$-object or a $\mathcal{B}$-object.

For each $n \geq 0$, the size-n objects are exactly the union of the size-n objects from each class:

$$\mathcal{C}[n] = \mathcal{A}[n] \uplus \mathcal{B}[n].$$

Hence, writing $a_n := |\mathcal{A}[n]|$, $b_n := |\mathcal{B}[n]|$, and $c_n := |\mathcal{C}[n]|$, we have

$$c_n = a_n + b_n \qquad (n \geq 0).$$

Theorem 9.1 (Addition rule for EGFs)**.** Let $\mathcal{A}$ and $\mathcal{B}$ be disjoint labeled combinatorial classes, and let

$$A(x) = \sum_{n\geq 0} a_n \frac{x^n}{n!}, \qquad B(x) = \sum_{n\geq 0} b_n \frac{x^n}{n!}$$

be their EGFs, where $a_n := |\mathcal{A}[n]|$ and $b_n := |\mathcal{B}[n]|$. If $\mathcal{C} = \mathcal{A} \uplus \mathcal{B}$, then the EGF of $\mathcal{C}$ is

$$C(x) = A(x) + B(x),$$

equivalently $c_n = a_n + b_n$ for all $n \geq 0$.

Proof. Since $\mathcal{C} = \mathcal{A} \uplus \mathcal{B}$ and the union is disjoint, for each $n \geq 0$ we have $\mathcal{C}[n] = \mathcal{A}[n] \uplus \mathcal{B}[n]$, hence $c_n = a_n + b_n$. Therefore

$$C(x) = \sum_{n\geq 0} c_n \frac{x^n}{n!} = \sum_{n\geq 0} (a_n + b_n) \frac{x^n}{n!} = \sum_{n\geq 0} a_n \frac{x^n}{n!} + \sum_{n\geq 0} b_n \frac{x^n}{n!} = A(x) + B(x).$$

□

9.2 The labeled product construction

We now define the most important operation.

Definition 9.5 (Labeled star product). Let $\mathcal{A}$ and $\mathcal{B}$ be labeled combinatorial classes. Their *(labeled) star product* $C = \mathcal{A} \star \mathcal{B}$ is defined as follows.

For each $n \geq 0$, an object of $C[n]$ is obtained by:

1. choosing an integer k with $0 \leq k \leq n$,
2. choosing a subset $U \subseteq [n]$ with $|U| = k$ (these labels go to the $\mathcal{A}$-part),
3. choosing an $\mathcal{A}$-object on the label set U (i.e. a copy of an object of $\mathcal{A}[k]$ whose atom-labels are exactly the elements of U),
4. choosing a $\mathcal{B}$-object on the complementary label set $[n] \setminus U$ (i.e. a copy of an object of $\mathcal{B}[n-k]$ whose atom-labels are exactly the elements of $[n] \setminus U$),
5. and recording the ordered pair (α, β).

Equivalently, $C[n]$ consists of all ordered pairs (α, β) where α uses some subset $U \subseteq [n]$ of labels, β uses the remaining labels, and together they use each label in $[n]$ exactly once.

9.3 Product rule for EGFs

Theorem 9.2 (Product rule for EGFs). Let $\mathcal{A}, \mathcal{B}$ be labeled classes with counts $a_n = |\mathcal{A}[n]|$ and $b_n = |\mathcal{B}[n]|$, and EGFs

$$A(x) = \sum_{n \geq 0} a_n \frac{x^n}{n!}, \qquad B(x) = \sum_{n \geq 0} b_n \frac{x^n}{n!}.$$

Let $C = \mathcal{A} \star \mathcal{B}$ be their labeled product, with $c_n = |C[n]|$ and EGF

$$C(x) = \sum_{n \geq 0} c_n \frac{x^n}{n!}.$$

Then

$$C(x) = A(x)\, B(x),$$

and equivalently, for each $n \geq 0$,

$$c_n = \sum_{j=0}^{n} \binom{n}{j} a_j\, b_{n-j}.$$

Combinatorial proof. Fix n and count $C[n]$.

To build an ordered pair $(\alpha, \beta) \in C[n]$, we do:

1. Choose how many labels go to the $\mathcal{A}$-part: say j labels.
2. Choose *which* j labels from $[n]$ go to α: $\binom{n}{j}$ choices.
3. Build α on those chosen labels: a_j choices.
4. Build β on the remaining $n - j$ labels: b_{n-j} choices.

Thus the number of objects of $C[n]$ with an $\mathcal{A}$-part of size j is $\binom{n}{j} a_j b_{n-j}$, and summing over j

gives

$$c_n = \sum_{j=0}^{n} \binom{n}{j} a_j b_{n-j}.$$

Finally, the identity $C(x) = A(x)B(x)$ is just the generating-function way to package that coefficient formula. □

A tiny sanity-check example

Let $\mathcal{A}$ be the class "a labeled set of size n with no extra structure." Then $a_n = 1$ for all n and

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

Now $\mathcal{A} \star \mathcal{A}$ is: split the labels into two groups, and record the ordered pair of groups. For fixed n, choosing the first group determines the second, so there are 2^n such ordered splits, i.e. $c_n = 2^n$. The product rule predicts

$$C(x) = A(x)^2 = e^{2x} = \sum_{n \geq 0} 2^n \frac{x^n}{n!},$$

9.5 Basic examples

Example 9.1. For each $n \geq 0$, define $\mathcal{G}[n]$ to be the set of all simple graphs with vertex set exactly $[n] = \{1, 2, \ldots, n\}$. Concretely, an element of $\mathcal{G}[n]$ is obtained by deciding, for every pair $\{i, j\}$ with $1 \leq i < j \leq n$, whether we include the edge between vertex i and vertex j.

What is the labeled object? A labeled object in $\mathcal{G}[n]$ is just a graph where the vertices already come with unique names $1, 2, \ldots, n$.

There are $\binom{n}{2}$ possible edges, and each edge is either present or not. Hence

$$a_n := |\mathcal{G}[n]| = 2^{\binom{n}{2}}.$$

EGF. Therefore the exponential generating function for labeled graphs is

$$G(x) = \sum_{n \geq 0} a_n \frac{x^n}{n!} = \sum_{n \geq 0} 2^{\binom{n}{2}} \frac{x^n}{n!}.$$

Example 9.2 (Words over an alphabet)**.** Let an alphabet have m letters. The number of words of length n over this alphabet is $a_n = m^n$. Hence the EGF is

$$A(x) = \sum_{n \geq 0} m^n \frac{x^n}{n!} = e^{mx}.$$

For the English alphabet ($m = 26$) we get $A(x) = e^{26x}$.

Example 9.3 (Decomposing into vowels and consonants)**.** Let V be the set of vowels ($|V| = 5$) and C the set of consonants ($|C| = 21$). Let $A(x)$ be the EGF for all words over $V \cup C$, $A_V(x)$ the EGF for words over V, and $A_C(x)$ the EGF for words over C.

We have

$$A_V(x) = e^{5x}, \qquad A_C(x) = e^{21x}.$$

Every word over $V \cup C$ is uniquely determined by:

- the set of positions occupied by vowels and by consonants, and
- the vowel word and consonant word on those positions.

On the level of EGFs this is just the product construction, so

$$A(x) = A_V(x)A_C(x) = e^{5x}e^{21x} = e^{26x},$$

in agreement with the direct count $a_n = 26^n$.

Example 9.4 (Permutations)**.** Let $a_n = n!$ be the number of permutations of $[n]$. The EGF is

$$A(x) = \sum_{n\geq 0} n!\,\frac{x^n}{n!} = \sum_{n\geq 0} x^n = \frac{1}{1-x}.$$

Example 9.5 (Placing flags on poles)**.** Let $r \geq 1$ be fixed. Consider n distinct flags and r distinct flagpoles. On each pole the flags are arranged in a linear order; poles may be empty. Let $a_n^{(r)}$ be the number of such arrangements on n flags.

For a single pole the number of arrangements is $n!$, so the EGF is

$$A(x) = \sum_{n\geq 0} n!\,\frac{x^n}{n!} = \frac{1}{1-x}.$$

Arrangements on r poles are an ordered r-tuple of independent arrangements on one pole, so the EGF is

$$A_r(x) = A(x)^r = \frac{1}{(1-x)^r}.$$

Hence

$$A_r(x) = \sum_{n\geq 0} a_n^{(r)}\frac{x^n}{n!} = \frac{1}{(1-x)^r} = \sum_{n\geq 0}\binom{n+r-1}{r-1}x^n,$$

and therefore

$$a_n^{(r)} = n!\binom{n+r-1}{r-1}.$$

Fix an alphabet of size m. We consider two kinds of objects:

- *multisets* of letters, counted by ordinary GFs;
- *words* of length n, counted by EGFs.

For one letter, the ordinary and exponential GFs under various multiplicity conditions are:

condition on multiplicity	OGF (one letter)	EGF (one letter)
unrestricted $0,1,2,\dots$	$1+x+x^2+\cdots=\dfrac{1}{1-x}$	$\sum_{n\ge 0}\dfrac{x^n}{n!}=e^x$
≤ 1	$1+x$	$1+x$
≥ 1	$x+x^2+\cdots=\dfrac{x}{1-x}$	e^x-1
even $(0,2,4,\dots)$	$1+x^2+x^4+\cdots=\dfrac{1}{1-x^2}$	$\dfrac{e^x+e^{-x}}{2}$
odd $(1,3,5,\dots)$	$x+x^3+\cdots=\dfrac{x}{1-x^2}$	$\dfrac{e^x-e^{-x}}{2}$

For an alphabet of size m we take the m-th power (letters behave independently):

condition on multiplicities	OGF (multisets)	EGF (words)
unrestricted	$\left(\dfrac{1}{1-x}\right)^m$	$(e^x)^m=e^{mx}$
each letter used at most once	$(1+x)^m$	$(1+x)^m$
each letter used at least once	$\left(\dfrac{x}{1-x}\right)^m$	$(e^x-1)^m$
each letter used an even number of times	$\left(\dfrac{1}{1-x^2}\right)^m$	$\left(\dfrac{e^x+e^{-x}}{2}\right)^m$
each letter used an odd number of times	$\left(\dfrac{x}{1-x^2}\right)^m$	$\left(\dfrac{e^x-e^{-x}}{2}\right)^m$

Example 9.6 (Ternary words with parity constraints)**.** Let the alphabet be $\{0,1,2\}$. We count words in which the number of 0's is even, the number of 1's is odd, and the number of 2's is arbitrary.

The one-letter EGFs are:

$$E_0(x)=\frac{e^x+e^{-x}}{2},\quad E_1(x)=\frac{e^x-e^{-x}}{2},\quad E_2(x)=e^x.$$

Hence the EGF of the desired words is

$$A(x)=E_0(x)E_1(x)E_2(x)=\frac{1}{4}(e^x+e^{-x})(e^x-e^{-x})e^x=\frac{1}{4}(e^{3x}-e^{-x}).$$

Thus the number a_n of such words of length n is

$$a_n=\frac{1}{4}(3^n-(-1)^n).$$

9.6 Stirling numbers of the second kind

Definition 9.6. For integers $n,k\ge 0$, the *Stirling number of the second kind* $S(n,k)$ is the number of partitions of the set $[n]=\{1,\dots,n\}$ into k nonempty unlabeled blocks.

Theorem 9.3 (EGF for $S(n,k)$)**.** For fixed $k \geq 0$,

$$\sum_{n\geq 0} S(n,k)\frac{x^n}{n!} = \frac{(e^x-1)^k}{k!}.$$

Proof. Instead of k unlabeled blocks, consider partitions of $[n]$ into k numbered boxes $1, 2, \ldots, k$, each required to be nonempty.

Giving a partition into k unlabeled blocks and then naming the blocks "box 1, box 2, ..., box k" produces a partition into labeled blocks. Conversely, forgetting the box names turns a labeled-block partition into an unlabeled-block partition.

For any fixed unlabeled partition into k blocks, there are exactly $k!$ ways to assign the names $1, 2, \ldots, k$ to its k blocks. Therefore,

$$\#\{\text{partitions of } [n] \text{ into } k \text{ numbered boxes}\} = k!\, S(n,k).$$

So it suffices to compute the EGF for *numbered-box* partitions; we can divide by $k!$ at the end. First consider *numbered* boxes: put the n elements of $[n]$ into k numbered boxes, each nonempty. A partition of $[n]$ into k numbered boxes is the same thing as a function

$$f : [n] \to [k]$$

such that every value $1, 2, \ldots, k$ is used at least once.

Indeed:

- Given such a function f, define box i to be $f^{-1}(i)$. The blocks are nonempty exactly when every $i \in [k]$ is hit by f.
- Given k labeled nonempty blocks $B_1, \ldots, B_k$, define $f(x) = i$ whenever $x \in B_i$. This is a well-defined function and it is automatically onto.

So "partition into k labeled nonempty blocks" $\iff$ "surjection $[n] \to [k]$".

Thus

$$\#\{\text{surjections } [n] \to [k]\} = k!\, S(n,k).$$

Given a surjection $f : [n] \to [k]$, define

$$B_i := f^{-1}(i) \qquad (1 \leq i \leq k).$$

Then $(B_1, \ldots, B_k)$ is a k-tuple of *nonempty* disjoint subsets whose union is $[n]$. Conversely, any such k-tuple determines a unique surjection by sending every element of B_i to i.

So a surjection is exactly: "split the labels $[n]$ into k *nonempty* parts, and remember the order $1, 2, \ldots, k$."

Let $\mathcal{S}_{\geq 1}$ be the class "a nonempty labeled set." For each $r \geq 1$ there is exactly one such object on $[r]$ (namely the set $[r]$), and for $r = 0$ there are none. Hence

$$S_{\geq 1}(x) = \sum_{r\geq 1} \frac{x^r}{r!} = e^x - 1.$$

A surjection consists of an ordered k-tuple of nonempty labeled sets (one for B_1, one for B_2, ..., one for B_k) whose labels are disjoint and together form $[n]$. By the labeled product rule, the EGF for an ordered k-tuple is the product of the EGFs, so the EGF for surjections is

$$(e^x - 1)^k.$$

By definition of EGF, saying "the EGF for surjections is $(e^x - 1)^k$" means

$$(e^x - 1)^k = \sum_{n\geq 0} \big(\#\{\text{surjections } [n] \to [k]\}\big) \frac{x^n}{n!}.$$

Substitute $\#\{\text{surjections } [n] \to [k]\} = k!\, S(n,k)$ to get

$$(e^x - 1)^k = \sum_{n\geq 0} k!\, S(n,k) \frac{x^n}{n!}.$$

Divide by $k!$ and we obtain

$$\sum_{n\geq 0} S(n,k) \frac{x^n}{n!} = \frac{(e^x - 1)^k}{k!},$$

which is exactly the desired EGF identity. □

9.7 Stirling numbers of the first kind

Definition 9.7. For integers $n, k \geq 0$, the *signless Stirling number of the first kind* $c(n,k)$ is the number of permutations of $[n]$ having exactly k cycles in their cycle decomposition.

The (signed) Stirling numbers of the first kind $s(n,k)$ are defined by

$$s(n,k) = (-1)^{n-k} c(n,k).$$

We set $s(0,0) = 1$ and $s(0,k) = s(n,0) = 0$ for $n, k > 0$.

Theorem 9.4. For all $n \in \mathcal{N}$,

$$x^n = \sum_{k=0}^{n} S(n,k)\, x^{\underline{k}}, \tag{2}$$

$$x^{\underline{n}} = \sum_{k=0}^{n} s(n,k)\, x^k. \tag{3}$$

Combinatorial proof of (2). Fix n and let x be a positive integer. Interpret x^n as the number of functions $f : [n] \to [x]$, i.e. words of length n over the alphabet $[x] = \{1, \ldots, x\}$.

Partition all such functions f according to the size k of the image $f([n])$. For a fixed k, we must:

- choose a k–element subset $S \subseteq [x]$ to be the image;
- choose a surjection $[n] \twoheadrightarrow S$.

There are $\binom{x}{k}$ ways to choose S, and $S(n,k)\, k!$ surjections onto a fixed k–set (as before). Hence the number of functions with image size k is

$$\binom{x}{k} k!\, S(n,k) = x^{\underline{k}}\, S(n,k).$$

Summing over $k = 0, \ldots, n$ yields

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

Since both sides are polynomials in x that agree for all sufficiently many integer values of x, the identity holds as a polynomial identity. □

Algebraic proof of (3). The polynomials $x^{\underline{0}}, \dots, x^{\underline{n}}$ form a basis of the vector space of polynomials of degree at most n, so the expansion (2) expresses the monomials x^n in that basis with coefficients $S(n,k)$. The matrix $S = (S(n,k))_{n,k\geq 0}$ is therefore invertible, and its inverse has entries $w(n,k)$, which gives the inverse expansion (3). (Equivalently, multiply the two identities and compare the coefficient of $x^{\underline{m}}$ on both sides.) □

Corollary 9.5 (Matrix inverse relation)**.** For all $n, m \geq 0$,

$$\sum_{k=0}^{n} S(n,k)\, s(k,m) = \delta_{n,m},$$

where $\delta_{n,m}$ is the Kronecker delta. Equivalently, the infinite lower-triangular matrices $(S(n,k))_{n,k\geq 0}$ and $(s(n,k))_{n,k\geq 0}$ are mutual inverses.

Proof. Start from (2):

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

Now substitute (3) (with n replaced by k) into each $x^{\underline{k}}$:

$$x^n = \sum_{k=0}^{n} S(n,k) \left(\sum_{m=0}^{k} s(k,m)\, x^m \right) = \sum_{m=0}^{n} \left(\sum_{k=m}^{n} S(n,k)\, s(k,m) \right) x^m.$$

On the other hand,

$$x^n = \sum_{m=0}^{n} \delta_{n,m}\, x^m.$$

Since the polynomials $1, x, x^2, \dots$ are linearly independent, the coefficient of x^m must agree for each m, giving

$$\sum_{k=m}^{n} S(n,k)\, s(k,m) = \delta_{n,m}.$$

Finally, extending the sum to $k = 0, \dots, n$ does not change anything because $s(k,m) = 0$ for $k < m$, so

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

□

9.8 Binomial inversion

Theorem 9.6 (Binomial inversion). Let $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ be sequences, and define EGFs

$$A(x) = \sum_{n\geq 0} a_n \frac{x^n}{n!}, \qquad B(x) = \sum_{n\geq 0} b_n \frac{x^n}{n!}.$$

Then the following are equivalent:

$$\text{(i)} \quad a_n = \sum_{k=0}^{n} \binom{n}{k} b_{n-k} \qquad (n \geq 0),$$

$$\text{(ii)} \quad b_n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} a_{n-k} \qquad (n \geq 0).$$

Proof. **(i) ⇒ EGF identity.** Assume (i). Then for each n,

$$a_n = \sum_{k=0}^{n} \binom{n}{k} b_{n-k} = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} b_{n-k}.$$

Multiply by $x^n/n!$ and sum over $n \geq 0$:

$$A(x) = \sum_{n\geq 0} a_n \frac{x^n}{n!} = \sum_{n\geq 0} \sum_{k=0}^{n} b_{n-k} \frac{x^n}{k!(n-k)!}.$$

Rewrite with $m = n - k$ (so $m \geq 0$ and $k \geq 0$):

$$A(x) = \sum_{k\geq 0} \sum_{m\geq 0} b_m \frac{x^{m+k}}{k!\, m!} = \left(\sum_{k\geq 0} \frac{x^k}{k!}\right)\left(\sum_{m\geq 0} b_m \frac{x^m}{m!}\right) = e^x B(x).$$

So (i) implies

$$A(x) = e^x B(x).$$

EGF identity ⇒ (ii). From $A(x) = e^x B(x)$ we get

$$B(x) = e^{-x} A(x).$$

Expand $e^{-x} = \sum_{k\geq 0} (-1)^k x^k/k!$ and multiply:

$$B(x) = \left(\sum_{k\geq 0} (-1)^k \frac{x^k}{k!}\right)\left(\sum_{m\geq 0} a_m \frac{x^m}{m!}\right).$$

By the product rule for EGFs, the coefficient of $x^n/n!$ in this product is

$$b_n = \sum_{k=0}^{n} \binom{n}{k} (-1)^k a_{n-k},$$

which is exactly (ii).

(ii) ⇒ (i). The same argument with x replaced by $-x$ reverses the steps, or equivalently: (ii) gives $B(x) = e^{-x} A(x)$, hence $A(x) = e^x B(x)$, and extracting coefficients yields (i). □

Example 9.7 (Derangements)**.** Let D_n be the number of derangements of $[n]$. Using the binomial inversion formula, obtain the EGF of $D(x)$

Every permutation of $[n]$ can be obtained by first choosing the set of fixed points and then deranging the rest, so

$$n! = \sum_{k=0}^{n} \binom{n}{k} D_{n-k}.$$

By binomial inversion,

$$D_n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)! = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

The EGF of the derangement numbers is therefore

$$D(x) = \sum_{n \geq 0} D_n \frac{x^n}{n!} = \frac{e^{-x}}{1-x}.$$

9.9 Exponential formula and connected structures

We now look at labelled combinatorial structures that decompose into *connected components*. The exponential formula describes their EGFs.

Let c_n be the number of structures of size n (on label set $[n]$), and define the exponential generating function

$$C(x) = \sum_{n\geq 1} c_n \frac{x^n}{n!}.$$

We call C *connected* if its elements are taken to be connected objects in some sense (graphs, permutations as products of cycles, set partitions as blocks, etc.).

Definition 9.8. Given a connected class $\mathcal{C}$, let $\mathcal{G}$ be the class of finite sets of components from $\mathcal{C}$, taken on disjoint label sets. Equivalently, $\mathcal{G}$ consists of all finite (possibly empty) structures obtained by taking a finite collection of connected components from $\mathcal{C}$ and relabelling them with the same label set.

Let g_n be the number of $\mathcal{G}$–structures on $[n]$, and let

$$G(x) = \sum_{n\geq 0} g_n \frac{x^n}{n!}$$

be its EGF (note that the empty structure is allowed, so $g_0 = 1$).

Theorem 9.7 (Exponential formula)**.** With notation as above, the EGFs satisfy

$$G(x) = \exp\big(C(x)\big).$$

Proof. Fix n and consider a $\mathcal{G}$–structure on $[n]$. It consists of a set $\{C_1, \ldots, C_r\}$ of connected components, where the label sets form a partition of $[n]$ into r (unordered) blocks of sizes $n_1, \ldots, n_r$ summing to n, and C_i is a $\mathcal{C}$–structure on the ith block.

If we temporarily regard the components as *labelled* $1, \ldots, r$, then the EGF for an ordered r–tuple of components is $C(x)^r/r!$, by the product and set constructions for EGFs. Summing over all $r \geq 0$, we obtain

$$G(x) = \sum_{r\geq 0} \frac{C(x)^r}{r!} = e^{C(x)}.$$

The factor $1/r!$ exactly compensates for the ordering of the components, so this counts unordered sets of components. □

Example 9.8 (Bell numbers)**.** The *nth Bell number* B_n is the number of ways to partition the labeled set $[n] = \{1, 2, \ldots, n\}$ into nonempty blocks (i.e. a set partition of $[n]$).

Goal. Find the exponential generating function

$$B(x) := \sum_{n\geq 0} B_n \frac{x^n}{n!}.$$

A *block* is a nonempty set of labels. If a structure consists of exactly one block on $[n]$, there is only one possibility: the block is $[n]$ itself. So $c_n = 1$ for $n \geq 1$ (and $c_0 = 0$), hence

$$C(x) = \sum_{n\geq 1} \frac{x^n}{n!} = e^x - 1.$$

A set partition is an unordered *set of blocks*. By the exponential formula for labeled classes, "a set of $\mathcal{C}$-objects" has EGF $\exp(C(x))$. Therefore

$$\sum_{n\geq 0} B_n \frac{x^n}{n!} = \exp(e^x - 1).$$

Example 9.9 (Permutations and cycles)**.** A *cycle* on $[n]$ means a single cyclic ordering of the labels $1, 2, \ldots, n$. Let $\mathcal{C}$ be the labeled class of single cycles, and let $\mathcal{G}$ be the labeled class of permutations (i.e. disjoint unions of cycles). Compute $C(x)$ for cycles, then use the exponential formula (permutations = union of cycle components) to recover the EGF for permutations.

Fix 1 as the start; then the remaining $n-1$ labels can be arranged in any order, so there are $(n-1)!$ cycles. Hence

$$C(x) = \sum_{n\geq 1} (n-1)! \frac{x^n}{n!} = \sum_{n\geq 1} \frac{x^n}{n} = -\log(1-x).$$

A permutation is a set of disjoint cycles. By the exponential formula,

$$G(x) = \exp(C(x)) = \exp(-\log(1-x)) = \frac{1}{1-x}.$$

Since there are $n!$ permutations of $[n]$, this matches

$$G(x) = \sum_{n\geq 0} n! \frac{x^n}{n!} = \sum_{n\geq 0} x^n = \frac{1}{1-x}.$$

Example 9.10 (Involutions)**.** An *involution* is a permutation π of $[n]$ such that $\pi^2 = \mathrm{id}$, equivalently: every cycle has length 1 or 2. Use the exponential formula (involutions = union of 1-cycles and 2-cycle components) to find the EGF

$$I(x) := \sum_{n\geq 0} I_n \frac{x^n}{n!},$$

where I_n is the number of involutions on $[n]$.

There is exactly 1 labeled 1-cycle on [1] (a fixed point), and exactly 1 labeled 2-cycle on [2] (a transposition). Thus the "allowed cycle" class has EGF

$$C(x) = x + \frac{x^2}{2}.$$

By the exponential formula (set of allowed cycles),

$$I(x) = \exp(C(x)) = \exp\left(x + \frac{x^2}{2}\right).$$

Example 9.11 (Connected graphs). Let G_n be the number of labeled simple graphs on vertex set $[n]$, and let C_n be the number of connected labeled simple graphs on vertex set $[n]$. Compute

$$C(x) := \sum_{n\geq 1} C_n \frac{x^n}{n!}$$

using the exponential formula (a graph is a set of connected components).

A labeled graph on $[n]$ is determined by choosing which of the $\binom{n}{2}$ possible edges are present, so

$$G_n = 2^{\binom{n}{2}}.$$

Define EGFs

$$G(x) = \sum_{n\geq 0} G_n \frac{x^n}{n!}, \qquad C(x) = \sum_{n\geq 1} C_n \frac{x^n}{n!}.$$

Every graph decomposes uniquely into a *set* of connected components. By the exponential formula,

$$G(x) = \exp(C(x)).$$

Hence

$$C(x) = \log G(x),$$

which determines the connected counts C_n by extracting coefficients of $\log\left(\sum_{n\geq 0} 2^{\binom{n}{2}} \frac{x^n}{n!}\right)$.

9.10 Lagrange Inversion Formula

We state the version we will use.

Theorem 9.8 (Lagrange inversion)**.** Let $\varphi(z)$ be a formal power series with $\varphi(0) \neq 0$. Suppose $y = y(x)$ is defined implicitly as the unique formal power series with $y(0) = 0$ satisfying

$$x = \frac{y}{\varphi(y)}.$$

Then for all integers $n \geq 1$ and $m \geq 1$,

$$[x^n]\, y(x)^m = \frac{m}{n}\, [z^{\,n-m}]\, \varphi(z)^n.$$

Proof. Substitute $x = y/\varphi(y)$ and regard y as an indeterminate. Using formal differentiation and the identity

$$\frac{d}{dx} y(x)^m = m y(x)^{m-1} y'(x),$$

one can write

$$y(x)^m = \frac{m}{n}\, x^n \frac{d}{dx}\big(\varphi(y(x))^n\big)$$

and then compare coefficients of x^n on both sides (or use Cauchy's integral formula on formal Laurent series). Rearranging gives the stated coefficient identity. For the full proof, see the textbook. □

9.11 Cayley's Formula from Lagrange Inversion

Let R_n be the number of rooted labelled trees on vertex set $[n]$, and let $r_n := R_n$ for brevity. Let

$$R(x) := \sum_{n \geq 1} r_n \frac{x^n}{n!}$$

be the EGF for rooted trees.

Functional equation for rooted trees

Consider a rooted tree on a labelled vertex set. From the root, remove the edges from the root to its neighbours; each neighbour then becomes the root of a rooted subtree. Thus a rooted tree is:

(a distinguished root vertex) + a set of rooted trees.

By the exponential formula this structure translates to the functional equation

$$R(x) = x \exp\big(R(x)\big).$$

Equivalently,

$$x = \frac{R(x)}{\exp(R(x))}.$$

Counting rooted trees via Lagrange inversion

Here $\varphi(z) = e^z$ and $y(x) = R(x)$ satisfies $x = y/\varphi(y)$, so by Lagrange inversion with $m = 1$ we obtain

$$[x^n]\, R(x) = \frac{1}{n}\, [z^{\,n-1}]\, e^{nz} = \frac{1}{n}\, \frac{n^{n-1}}{(n-1)!} = \frac{n^{n-1}}{n!}.$$

Therefore

$$r_n = n^{n-1}$$

rooted labelled trees on $[n]$.

Unrooted trees

Every labelled tree on $[n]$ has exactly n possible choices of root, so $r_n = n \cdot t_n$, where t_n is the number of (unrooted) labelled trees on $[n]$. Hence

$$t_n = \frac{r_n}{n} = n^{n-2}.$$

Theorem 9.9 (Cayley's formula)**.** The number of labelled trees on vertex set $[n]$ is

$$t_n = n^{n-2}.$$

10 Integer Partitions

Definition 10.1. A *partition* of a nonnegative integer n is a sequence $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ of positive integers with

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell \geq 1, \qquad \lambda_1 + \cdots + \lambda_\ell = n.$$

The λ_i are the *parts* of the partition.

Let $p(n)$ be the number of partitions of n. We encode these in an ordinary generating function.

Example 10.1 (OGF for integer partitions)**.** Let $p(n)$ be the number of (integer) partitions of n. Find the ordinary generating function

$$P(x) := \sum_{n \geq 0} p(n)\, x^n.$$

Idea: "unlimited coin change": A partition of n is the same thing as choosing how many 1's you use, how many 2's you use, how many 3's you use, and so on, with the total sum coming out to n. So a partition is determined by a sequence of multiplicities

$$(m_1, m_2, m_3, \ldots) \quad \text{where } m_j \in \{0, 1, 2, \ldots\}$$

and

$$1 \cdot m_1 + 2 \cdot m_2 + 3 \cdot m_3 + \cdots = n.$$

(Only finitely many m_j are nonzero for a given partition.)

What does $1 + x^j + x^{2j} + \cdots$ **mean?** Fix a part size j. You are allowed to use j:

$$0 \text{ times}, \ 1 \text{ time}, \ 2 \text{ times}, \ \ldots$$

If you use j exactly m times, it contributes mj to the total sum. In an OGF, contributing mj corresponds to multiplying by x^{mj}. So the "menu of choices" for part size j is exactly

$$1 \ + \ x^j \ + \ x^{2j} \ + \ x^{3j} \ + \ \cdots,$$

where:

- the term 1 means "use j zero times",
- the term x^j means "use j once",
- the term x^{2j} means "use j twice", etc.

This is a geometric series, so

$$1 + x^j + x^{2j} + \cdots = \frac{1}{1 - x^j}.$$

Why do we multiply over all j**?** Now we make *all* the choices at once: choose how many 1's, how many 2's, how many 3's, etc. These choices are independent (picking the number of 5's does not restrict how many 2's you pick), so we multiply the choice series:

$$P(x) = \prod_{j \geq 1} (1 + x^j + x^{2j} + \cdots) = \prod_{j \geq 1} \frac{1}{1 - x^j}.$$

Why does the coefficient equal $p(n)$? When you expand the product, you pick *one term* from each factor. Picking $x^{m_j j}$ from the jth factor means "use j exactly m_j times." The product of all chosen terms is

$$x^{m_1 \cdot 1}\, x^{m_2 \cdot 2}\, x^{m_3 \cdot 3} \cdots = x^{\,1m_1 + 2m_2 + 3m_3 + \cdots}.$$

So you get a contribution to x^n precisely when the chosen multiplicities satisfy

$$1m_1 + 2m_2 + 3m_3 + \cdots = n,$$

which is exactly the condition for a partition of n. Moreover, each partition corresponds to exactly one such choice of terms. Therefore the coefficient of x^n in $P(x)$ is $p(n)$, i.e.

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

[Bounded largest part]

Example 10.2. Let $p_k(n)$ be the number of partitions of n in which every part is at most k. Find the ordinary generating function

$$P_k(x) := \sum_{n \ge 0} p_k(n)\, x^n.$$

In the full partition OGF

$$P(x) = \prod_{j \ge 1} \frac{1}{1 - x^j},$$

the factor $\frac{1}{1-x^j} = 1 + x^j + x^{2j} + \cdots$ is the "menu" for how many j's you use. If we require *all parts* $\le k$, then parts $k+1, k+2, \ldots$ are forbidden, meaning their multiplicities must be 0.

In OGF language, "must use 0 of size j" means the only allowed term is 1, so we simply omit those factors. Equivalently, we keep only the factors for $j = 1, 2, \ldots, k$:

$$\sum_{n \ge 0} p_{\le k}(n)\, x^n = \prod_{j=1}^{k} \left(1 + x^j + x^{2j} + \cdots\right) = \prod_{j=1}^{k} \frac{1}{1 - x^j}.$$

Example 10.3 (Partitions into distinct parts)**.** Let $q(n)$ be the number of partitions of n into *distinct* parts (no part size is repeated). Find the ordinary generating function

$$Q(x) := \sum_{n \ge 0} q(n)\, x^n.$$

Use the same "one factor per part size" idea as for $P(x)$, but now each part size j can be used *at most once*. So for each $j \ge 1$ the only allowed choices are:

$$\text{use } j \text{ zero times} \;\Rightarrow\; 1, \qquad \text{use } j \text{ once} \;\Rightarrow\; x^j.$$

Thus the jth factor becomes $1 + x^j$. Multiplying over all j gives

$$Q(x) = \prod_{j \ge 1} (1 + x^j),$$

and by construction the coefficient of x^n counts exactly the partitions of n into distinct parts.

Partitions with restricted part sizes

More generally, fix a set $S \subseteq \mathbb{N}$ of allowed part sizes. Let $p_S(n)$ be the number of partitions of n whose parts all belong to S. The OGF is

$$\sum_{n\geq 0} p_S(n)x^n = \prod_{j\in S} \frac{1}{1-x^j}.$$

These generating functions are the starting point for many further identities and asymptotic results about $p(n)$ and its variants.

10.2 Hardy–Ramanujan asymptotics and a simple upper bound

The famous Hardy–Ramanujan formula (1918) gives the asymptotic

$$p(n) \sim \frac{1}{4n\sqrt{3}} \exp\left(\pi\sqrt{\frac{2n}{3}}\right).$$

A much simpler inequality (due to Lint, 1971) says that for all n,

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

Let

$$P(x) := \sum_{n\geq 0} p(n)x^n = \prod_{k\geq 1} \frac{1}{1-x^k}$$

be the ordinary generating function of the partition numbers.
Taking logarithms,

$$\log P(x) = -\sum_{k\geq 1} \log(1-x^k) = \sum_{k\geq 1}\sum_{j\geq 1} \frac{x^{kj}}{j} = \sum_{m\geq 1}\left(\frac{1}{m}\sum_{k|m} k\right)x^m.$$

From the estimate

$$\sum_{k|m} k \leq m(1+\log m)$$

one obtains, after some calculus, the bound

$$\log P(x) \leq \frac{\pi^2}{6(1-x)} + \frac{1}{2}\log\frac{1}{1-x} + O(1)$$

as $x \uparrow 1$; extracting coefficients yields (4).

10.3 Ferrers diagrams and conjugation

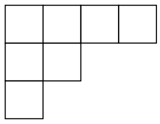
Definition 10.2 (Ferrers diagram). A *partition* of n is a sequence

$$\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell) \quad \text{with} \quad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell \geq 1, \qquad \sum_{i=1}^{\ell} \lambda_i = n.$$

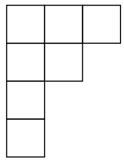
The *Ferrers diagram* of λ is the left-justified array of dots with λ_i dots in row i.

Definition 10.3 (Conjugate partition). The *conjugate* λ' of a partition λ is obtained by reflecting its Ferrers diagram across the main diagonal. Equivalently, λ'_j is the number of parts of λ of size at least j.

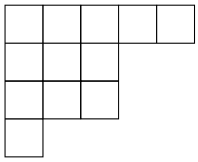
$\lambda = (4, 2, 1)$



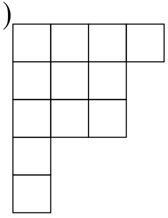
$\lambda' = (3, 2, 1, 1)$



$\mu = (5, 3, 3, 1)$



$\mu' = (4, 3, 3, 1, 1)$



Conjugation is an involution: $(\lambda')' = \lambda$.

Proposition 10.1. For each $k \geq 1$, the number of partitions of n whose largest part is k equals the number of partitions of n with exactly k parts.

Proof. If λ has largest part k, then in λ' the number of parts equals k (there are k columns in the Ferrers diagram). Thus conjugation is a bijection between the two classes of partitions. □

10.4 Distinct parts versus odd parts (Frobenius 1882)

Theorem 10.2 (Euler–Frobenius).

$$\prod_{i \geq 1} (1 + x^i) = \prod_{j \geq 1} \frac{1}{1 - x^{2j-1}}.$$

The left-hand side is the generating function for partitions into *distinct* parts; the right-hand side is the generating function for partitions into *odd* parts. Hence:

Corollary 10.3. For every n, the number of partitions of n into distinct parts equals the number of partitions of n into odd parts.

Proof. Let $\mathcal{O}_n$ be the set of partitions of n into odd parts and $\mathcal{D}_n$ the set of partitions of n into distinct parts. We construct a bijection $\Phi : \mathcal{O}_n \to \mathcal{D}_n$.

Every odd integer can be written uniquely as $m = 2^t(2u + 1)$ with $t \geq 0$ and $u \geq 0$. Take $\lambda \in \mathcal{O}_n$ and fix one odd number $m = 2u + 1$. Suppose m occurs in λ with multiplicity $q \geq 0$. Write q in binary:

$$q = \varepsilon_0 2^0 + \varepsilon_1 2^1 + \cdots + \varepsilon_s 2^s, \qquad \varepsilon_j \in \{0, 1\}.$$

For this odd part m we replace the q copies of m by the (at most $s + 1$) parts

$$2^j m = 2^j(2u + 1) \quad \text{for all } j \text{ with } \varepsilon_j = 1.$$

Do this independently for each distinct odd number m that occurs in λ.

The resulting multiset of parts has the following properties.

- All parts are distinct: for fixed m, the powers $2^j m$ are distinct because the ε_j's are 0 or 1, and for different odd m we cannot have $2^{t_1}m_1 = 2^{t_2}m_2$ since that would force m_1 and m_2 to have the same odd part.
- The sum of all parts is preserved: we simply regrouped q copies of m into $\sum_j \varepsilon_j$ copies of $2^j m$ whose total is still qm.

Thus $\Phi(\lambda)$ is a partition of n into distinct parts, so $\Phi : \mathcal{O}_n \to \mathcal{D}_n$ is well-defined.

Conversely, given $\mu \in \mathcal{D}_n$, write each part of μ in the form $2^t(2u+1)$ with $t \geq 0$ and $2u+1$ odd. For each fixed odd integer $m = 2u+1$, consider all parts of μ whose odd component is m: they have the form $2^t m$ with distinct exponents t, because the parts of μ are distinct. Replace each such part $2^t m$ by 2^t copies of m. Summing over all m we obtain a partition of n into odd parts; this is the inverse map $\Psi : \mathcal{D}_n \to \mathcal{O}_n$.

It is immediate from the definitions that Ψ is inverse to Φ. Hence Φ is a bijection and $|\mathcal{O}_n| = |\mathcal{D}_n|$. □

10.5 Integer triangles and partitions

Let a_n be the number of integer-sided triangles (a, b, c) with perimeter n, counted up to permutation of the side lengths and satisfying the triangle inequalities.

Theorem 10.4. The ordinary generating function of $(a_n)_{n\geq 0}$ is

$$\sum_{n\geq 0} a_n x^n = \frac{x^3}{(1-x^2)(1-x^3)(1-x^4)}.$$

Proof. Since we count triangles up to permutation of the side lengths, every triangle can be written uniquely with

$$a \geq b \geq c \geq 1, \qquad a < b + c, \qquad a + b + c = n.$$

Introduce new nonnegative integers

$$t := c - 1 \geq 0, \qquad u := b - c \geq 0, \qquad w := a - b \geq 0.$$

Then

$$c = t + 1, \quad b = c + u = t + u + 1, \quad a = b + w = t + u + w + 1.$$

The perimeter is therefore

$$n = a + b + c = (t + u + w + 1) + (t + u + 1) + (t + 1) = 3t + 2u + w + 3.$$

The only remaining condition is the triangle inequality $a < b + c$:

$$a < b + c \iff t + u + w + 1 < (t + u + 1) + (t + 1) = 2t + u + 2 \iff w < t + 1.$$

Since w and t are integers, this is equivalent to $0 \leq w \leq t$.

Now write t as $t = s + w$ with $s \geq 0$; this parametrises all pairs (t, w) with $0 \leq w \leq t$. Substituting into the expression for n, we obtain

$$n = 3(s + w) + 2u + w + 3 = 3s + 2u + 4w + 3.$$

Thus each triangle corresponds uniquely to a triple (s, u, w) of nonnegative integers, and the perimeter of the triangle is

$$n = 3s + 2u + 4w + 3.$$

Hence the ordinary generating function is

$$\sum_{s,u,w\geq 0} x^{3s+2u+4w+3} = x^3 \left(\sum_{s\geq 0} x^{3s}\right)\left(\sum_{u\geq 0} x^{2u}\right)\left(\sum_{w\geq 0} x^{4w}\right) = \frac{x^3}{(1-x^3)(1-x^2)(1-x^4)}.$$

This is the claimed rational function. □

10.6 Euler's identity for self-conjugate partitions

Theorem 10.5 (Euler)**.** The generating function for self-conjugate partitions (those equal to their conjugates) is

$$\prod_{i\geq 1}(1 + x^{2i-1}) = 1 + \sum_{k\geq 1} \frac{x^{k^2}}{(1-x^2)(1-x^4)\cdots(1-x^{2k})}.$$

Proof. We prove both equalities combinatorially.

Step 1: Decomposition by Durfee square. Let λ be a self-conjugate partition and draw its Ferrers diagram. Let k be the size of its Durfee square, i.e. the largest integer such that the Young diagram contains a $k \times k$ square in the top left corner. Because λ is self-conjugate, this square is symmetric with respect to the main diagonal, and all cells outside this square come in symmetric pairs: if a cell appears to the right of the square in row i, a matching cell appears below the square in column i.

For each $i = 1, \ldots, k$, let r_i be the number of cells to the right of the Durfee square in row i; self-conjugacy implies that there are also r_i cells below the square in column i. Since the row lengths are weakly decreasing, we have

$$r_1 \geq r_2 \geq \cdots \geq r_k \geq 0.$$

Thus $(r_1, \ldots, r_k)$ is a partition (possibly with zero parts) with at most k parts.

The size of λ is

$$|\lambda| = k^2 + 2\sum_{i=1}^{k} r_i,$$

the k^2 cells of the Durfee square plus r_i cells in row i and r_i cells in column i for each i.

Fix k. The contribution to the generating function from all self-conjugate partitions with Durfee square of size k is therefore

$$x^{k^2} \sum_{r_1\geq\cdots\geq r_k\geq 0} x^{2(r_1+\cdots+r_k)}.$$

But the inner sum is exactly the generating function for partitions with at most k parts, with weight x^{2m} for a total of m cells. It is well known (and easy to check by the usual Ferrers-diagram bijection) that the generating function for partitions with at most k parts is $\prod_{i=1}^{k}(1-x^i)^{-1}$. Replacing x by x^2 gives

$$\sum_{r_1\geq\cdots\geq r_k\geq 0} x^{2(r_1+\cdots+r_k)} = \prod_{i=1}^{k}\frac{1}{1-x^{2i}}.$$

Thus the total generating function for self-conjugate partitions is

$$\sum_{k\geq 0} x^{k^2}\prod_{i=1}^{k}\frac{1}{1-x^{2i}} = 1+\sum_{k\geq 1}\frac{x^{k^2}}{(1-x^2)(1-x^4)\cdots(1-x^{2k})},$$

proving the second equality.

Step 2: Bijection with partitions into distinct odd parts. Let $\mathcal{S}$ be the set of self-conjugate partitions and $\mathcal{D}_{\text{odd}}$ the set of partitions into distinct odd parts. We construct a bijection

$$\Phi : \mathcal{D}_{\text{odd}} \longrightarrow \mathcal{S}.$$

Given $\mu \in \mathcal{D}_{\text{odd}}$, write its parts in decreasing order and denote them by

$$\mu_1 > \mu_2 > \cdots > \mu_\ell, \qquad \text{each } \mu_j \text{ odd}.$$

Write $\mu_j = 2a_j + 1$ with $a_j \geq 0$. We now build a Ferrers diagram for a self-conjugate partition by successively adding *hooks* centered on the main diagonal.

Start with a single cell on the diagonal. For the largest part $\mu_1 = 2a_1 + 1$, attach a hook of arm length a_1 to the right and a leg length a_1 downward from this central cell; this gives a symmetric "cross" of $2a_1+1$ cells. For $\mu_2 = 2a_2+1$, place a similar hook strictly inside the previous one (one step closer to the diagonal), and so on. Because the parts are strictly decreasing, these hooks nest properly and produce a Ferrers diagram that is symmetric about the diagonal. The total number of cells equals the sum of the parts μ_j, so we have obtained a self-conjugate partition $\Phi(\mu)$ of the same integer.

Conversely, given a self-conjugate partition λ, examine its Ferrers diagram and look at the cells lying on the main diagonal. Around each diagonal cell there is a maximal symmetric hook: move right until the diagram ends, and move down until the diagram ends; by self-conjugacy these two legs have the same length, say a. This hook contains $2a + 1$ cells. Remove all these hooks, starting with the outermost and proceeding inward; what remains is again a (possibly empty) self-conjugate Ferrers diagram, and the lengths of the hooks removed form a strictly decreasing sequence of odd integers $2a_1+1 > 2a_2+1 > \cdots$. This sequence is precisely a partition into distinct odd parts. This defines the inverse map $\Psi : \mathcal{S} \to \mathcal{D}_{\text{odd}}$.

It is straightforward to check that Ψ is the inverse of Φ, so Φ is a bijection. Therefore the generating function of self-conjugate partitions equals the generating function of partitions into distinct odd parts, which is

$$\prod_{i\geq 1}(1+x^{2i-1}),$$

since each odd part $2i-1$ can be chosen at most once.

Combining Step 1 and Step 2 gives both equalities in the statement of the theorem. □

11 Inclusion–Exclusion Principle (PIE)

11.1 Basic statement

Let U be a finite universal set and let $A_1, \ldots, A_n \subseteq U$.

Theorem 11.1 (Inclusion–Exclusion Principle)**.**

$$\left|A_1 \cup \cdots \cup A_n\right| = \sum_{\emptyset \neq T \subseteq [n]} (-1)^{|T|+1} \Big|\bigcap_{i \in T} A_i\Big|.$$

Equivalently,

$$\left|U \setminus (A_1 \cup \cdots \cup A_n)\right| = \sum_{T \subseteq [n]} (-1)^{|T|} \Big|\bigcap_{i \in T} A_i\Big|.$$

Proof. Fix $x \in U$ and let t be the number of sets A_i that contain x. On the right-hand side, x is counted in exactly $\binom{t}{j}$ intersections of size j, with sign $(-1)^j$. Thus its total contribution is

$$\sum_{j=0}^{t} (-1)^j \binom{t}{j} = (1-1)^t = \begin{cases} 1, & t = 0, \\ 0, & t \geq 1. \end{cases}$$

Hence x is counted once iff it lies in none of the A_i, and not counted otherwise. Summing over all $x \in U$ gives the formula for $|U \setminus \bigcup_i A_i|$; the formula for $|\bigcup_i A_i|$ follows. □

11.2 Derangements

Example 11.1 (Derangements)**.** Let $U = S_n$ be the set of all permutations of $[n]$. For each $i \in [n]$, let A_i be the set of permutations with i as a fixed point. Then $|A_i| = (n-1)!$ and

$$\Big|\bigcap_{i \in T} A_i\Big| = (n - |T|)!.$$

The number D_n of derangements (permutations with no fixed points) is therefore

$$D_n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)!.$$

11.3 Euler's totient function

Definition 11.1. For $m \geq 1$, Euler's totient function

$$\varphi(m) := |\{1 \leq i \leq m : \gcd(i, m) = 1\}|$$

counts integers mod m that are relatively prime to m.

Let $p_1, \ldots, p_s$ be the distinct prime divisors of m.

Theorem 11.2.

$$\varphi(m) = m \prod_{i=1}^{s} \left(1 - \frac{1}{p_i}\right) = m \sum_{T \subseteq [s]} (-1)^{|T|} \frac{1}{\prod_{i \in T} p_i}.$$

Proof. Let $U = [m]$. For $i = 1, \ldots, s$, let A_i be the set of integers in $[m]$ divisible by p_i. Then $\varphi(m) = |U \setminus \bigcup_i A_i|$. For a nonempty $T \subseteq [s]$,

$$\left|\bigcap_{i\in T} A_i\right| = \left\lfloor \frac{m}{\prod_{i\in T} p_i} \right\rfloor = \frac{m}{\prod_{i\in T} p_i},$$

since $\prod_{i\in T} p_i$ divides m. PIE gives the stated formula. □

11.4 A PIE formula for Stirling numbers

Let $S(n,k)$ be the Stirling number of the second kind, the number of set partitions of $[n]$ into k nonempty blocks.

Theorem 11.3. For integers $n, k \geq 0$,

$$S(n,k) = \frac{1}{k!}\sum_{j=0}^{k}(-1)^j\binom{k}{j}(k-j)^n.$$

Equivalently,

$$k!S(n,k) = \sum_{j=0}^{k}(-1)^j\binom{k}{j}(k-j)^n.$$

Proof. Fix n and k. Let U be the set of all functions $f : [n] \longrightarrow [k]$, so $|U| = k^n$. For $i \in [k]$, let A_i be the set of functions that *miss* the value i (no element is mapped to i). Then functions that hit all k values are precisely $U \setminus \bigcup_i A_i$.

Moreover,

$$\left|\bigcap_{i\in T} A_i\right| = (k - |T|)^n.$$

By PIE,

$$|U \setminus \bigcup_i A_i| = \sum_{j=0}^{k}(-1)^j\binom{k}{j}(k-j)^n.$$

Each surjection $[n] \twoheadrightarrow [k]$ corresponds to a partition of $[n]$ into k labelled blocks; forgetting labels yields a factor of $k!$. Hence

$$k!S(n,k) = \sum_{j=0}^{k}(-1)^j\binom{k}{j}(k-j)^n.$$

□

11.5 Multisets via inclusion–exclusion

Let $a_{m,n,r}$ be the number of multisets of size m drawn from an n-element type set, where each type may appear at most r times. Equivalently, $a_{m,n,r}$ counts integer solutions of

$$x_1 + \cdots + x_n = m, \qquad 0 \leq x_i \leq r.$$

Proposition 11.4.

$$a_{m,n,r} = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \binom{m - k(r+1) + n - 1}{n - 1},$$

where the binomial coefficient is interpreted as 0 if the top is $< n - 1$.

Proof. Let U be the set of all solutions in nonnegative integers of $x_1 + \cdots + x_n = m$; then

$$|U| = \binom{m + n - 1}{n - 1}$$

by the stars-and-bars argument. For $i \in [n]$, let A_i be the set of solutions with $x_i \geq r + 1$. Then $a_{m,n,r} = |U \setminus \bigcup_i A_i|$.

For $T \subseteq [n]$, $|T| = k$, we have $x_i \geq r + 1$ for all $i \in T$. Writing $x_i' = x_i - (r + 1)$ for $i \in T$ and $x_j' = x_j$ otherwise,

$$\sum_{j=1}^{n} x_j' = m - k(r + 1),$$

and the number of such solutions is

$$\binom{m - k(r+1) + n - 1}{n - 1},$$

provided $m - k(r + 1) \geq 0$, and 0 otherwise. PIE now yields the claimed formula. □

11.6 PIE as an evaluation tool for sums

Many alternating binomial sums can be interpreted using PIE. For instance:

Example 11.2.

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

Interpret 2^{n-k} as the number of subsets of $[n]$ that avoid a fixed k-element set and apply PIE with U the power set of $[n]$ and A_i the family of subsets containing element i.

Example 11.3. For integers $m, n \geq 0$,

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

Let U be the family of r-subsets of $[m + n]$ and let A_i be the subsets containing element i for $i \in [n]$. Then the right-hand side counts r-subsets disjoint from $[n]$, while the left-hand side is the PIE expansion for $|U \setminus \bigcup_i A_i|$.

Example 11.4 (Derangements again)**.**

$$D_n = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

Apply PIE with $U = S_n$ and A_i the permutations fixing i, as in the previous section.

Generalization of derangements

Definition 11.2 (Permutation matrix). Let G be a permutation of $\{1,\dots,n\}$. The *permutation matrix* of G is the $n \times n$ matrix $A = (a_{ij})$ with entries in $\{0,1\}$ defined by

$$a_{ij} = \begin{cases} 1, & \text{if } j = G(i), \\ 0, & \text{otherwise.} \end{cases}$$

Equivalently, every row and every column of A contains exactly one entry equal to 1.

This point of view lets us extend the notion of derangements: instead of forbidding the diagonal entries (i,i), one may forbid an arbitrary set of matrix positions for the 1's.

Definition 11.3 (Derangement as a forbidden–positions problem). A classical derangement of $\{1,\dots,n\}$ is a permutation G such that $G(i) \neq i$ for all i, i.e. in the permutation matrix of G no 1 is allowed on the diagonal squares (i,i).

Definition 11.4 (Forbidden squares and sets A_s). Let s be a square (matrix position) $s = (i,j)$ in an $n \times n$ matrix.

- Let A_s be the set of permutations G of $\{1,\dots,n\}$ such that the permutation matrix of G has a 1 in position s; equivalently, $G(i) = j$. Then

$$|A_s| = (n-1)!.$$

- For a set I of squares, write

$$\bigcap_{s \in I} A_s$$

for the set of permutations whose permutation matrix has 1's in all positions in I. We call the squares in I *independent* if no two of them lie in the same row or in the same column. In that case, the conditions $G(i) = j$ for $(i,j) \in I$ fix $|I|$ values of the permutation, and

$$\left| \bigcap_{s \in I} A_s \right| = (n - |I|)!.$$

11.8 Rook polynomials

Definition 11.5 (Boards and rook numbers). Let B be a board (a subset of squares of an $n \times n$ grid). A subset $B' \subseteq B$ is called *independent* if no two squares of B' lie in the same row or the same column.

For $k \geq 0$ define $r_k(B)$ to be the number of independent k-subsets of B. By convention $r_0(B) = 1$ (the empty set).

Definition 11.6 (Rook polynomial). The *rook polynomial* of a board B is

$$R_B(x) := \sum_{k\geq 0} r_k(B)\, x^k.$$

Example 11.5. For the 4×4 board B indicated in the figure in the notes (with certain forbidden squares), one finds

$$r_0(B) = 1, \quad r_1(B) = 5, \quad r_2(B) = 7, \quad r_3(B) = 2, \quad r_4(B) = 0,$$

so

$$R_B(x) = 1 + 5x + 7x^2 + 2x^3.$$

Proposition 11.5 (Product rule). Suppose $B = B_1 \cup B_2$ where B_1 and B_2 lie in disjoint sets of rows and columns (no row or column of B contains squares from both B_1 and B_2). Then

$$R_B(x) = R_{B_1}(x)\, R_{B_2}(x).$$

Proof. An independent k-subset of B is obtained by choosing, for some i, an independent i-subset of B_1 and an independent $(k-i)$-subset of B_2. Thus

$$r_k(B) = \sum_{i=0}^{k} r_i(B_1)\, r_{k-i}(B_2),$$

and the stated identity is exactly the Cauchy product for the series R_{B_1} and R_{B_2}. □

Proposition 11.6 (Recursion). Let s be a square of B. Let $B \setminus s$ be the board obtained from B by deleting s, and let $B - s$ be the board obtained by deleting s together with its entire row and column. Then

$$R_B(x) = R_{B\setminus s}(x) + x\, R_{B-s}(x).$$

Proof. Independent rook sets on B either avoid s or contain s. Those avoiding s are precisely the independent sets of $B \setminus s$. Those containing s correspond to independent sets on $B - s$ (place one rook at s, delete its row and column, and choose the remaining rooks). In the generating polynomial this gives

$$R_B(x) = \sum_{k\geq 0} r_k(B)x^k = \sum_{k\geq 0} r_k(B\setminus s)x^k + \sum_{k\geq 1} r_{k-1}(B-s)x^k = R_{B\setminus s}(x) + xR_{B-s}(x).$$

□

11.9 Polynomial Inclusion–Exclusion

Let U be a finite set and $A_1, \ldots, A_n \subseteq U$.

Definition 11.7. For $p \geq 0$ let

$$a_p := \left|\{u \in U : u \text{ lies in exactly } p \text{ of the } A_i\}\right|.$$

For $k \geq 0$ let

$$b_k := \sum_{\substack{I \subseteq [n] \\ |I|=k}} \left|\bigcap_{i \in I} A_i\right|$$

(with the convention that for $k = 0$ the sum has one term $|U|$).

Theorem 11.7 (Polynomial form of P.I.E.). For an indeterminate x we have the identity of polynomials

$$\sum_{p \geq 0} a_p x^p \;=\; \sum_{k \geq 0} b_k (x-1)^k.$$

Proof. Fix $u \in U$ and suppose u lies in exactly p of the sets A_i. Then u contributes x^p to the left-hand side. On the right-hand side, u belongs to $\bigcap_{i \in I} A_i$ precisely when I is a subset of the p indices of sets containing u. For each $k \leq p$ there are $\binom{p}{k}$ such subsets I, so the contribution of u to the right-hand side is

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

by the binomial theorem. Since every $u \in U$ contributes the same amount to both sides, the polynomials are equal. □

11.10 Fixed points of a random permutation

Let $U = S_n$ be the set of all permutations of $[n]$, and for each $i \in [n]$ let A_i be the set of permutations that fix i.

For $\sigma \in S_n$ let $p(\sigma)$ be the number of fixed points of σ; then a_p is the number of permutations with exactly p fixed points. Define

$$N(x) := \sum_{\sigma \in S_n} x^{p(\sigma)} = \sum_{p \geq 0} a_p x^p.$$

The polynomial P.I.E. gives an alternative expression for $N(x)$.

First compute b_k: for a fixed k and a subset $I \subseteq [n]$ with $|I| = k$,

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

since the k points in I must be fixed and the remaining $n-k$ points can be permuted arbitrarily. There are $\binom{n}{k}$ choices of I, so

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

Therefore

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

The expected number of fixed points of a random permutation is

$$\mathbb{E}[p(\sigma)] = \frac{N'(1)}{n!}.$$

Differentiate:

$$N'(x) = \sum_{k=1}^{n} \binom{n}{k}(n-k)!\,k(x-1)^{k-1}.$$

Evaluating at $x = 1$ only the $k = 1$ term survives, giving

$$N'(1) = \binom{n}{1}(n-1)! = n!.$$

Hence

$$\mathbb{E}[p(\sigma)] = \frac{N'(1)}{n!} = 1.$$

So a random permutation of $[n]$ has on average one fixed point.

Many permutation–avoidance problems can be converted to rook polynomials by interpreting forbidden positions of the permutation matrix as the squares of a board B: a rook in (i, j) means that $\sigma(i) = j$.

Independent sets of k squares in B correspond to permutations that violate exactly k of the forbidden positions. Combining rook polynomials with P.I.E. yields formulas for the number of permutations avoiding all forbidden positions (e.g. derangements as the special case where the forbidden squares are the diagonal).

12 Symmetric counting

12.1 Signed permutations, parity, and determinants

Definition 12.1 (Involution). A permutation $\pi \in S_n$ is an *involution* if $\pi^2 = \mathrm{id}$. Equivalently, every cycle of π has length 1 or 2.

Definition 12.2 (Parity and sign of a permutation). A permutation $\sigma \in S_n$ is *even* if it can be written as a product of an even number of transpositions; otherwise it is *odd*. The *sign* of σ is

$$\mathrm{sign}(\sigma) := \begin{cases} +1, & \sigma \text{ even}, \\ -1, & \sigma \text{ odd}. \end{cases}$$

Equivalent characterisations of parity include:

- parity of the number of inversions of σ;
- parity of the number of cycles of even length, etc.

Definition 12.3 (Determinant). For an $n \times n$ matrix $A = (a_{i,j})$ over a commutative ring, the determinant is

$$\det A = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}.$$

Example 12.1. If A is the $n \times n$ all-1 matrix, then $\det A = 0$ for $n \geq 2$. Indeed,

$$\det A = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) = \#\{\text{even permutations}\} - \#\{\text{odd permutations}\},$$

and these two numbers are equal, so the sum vanishes.

Consider lattice paths in $\mathbb{Z}^2$ that use only unit steps $(1,0)$ (right) and $(0,1)$ (up). For points $P = (p_1, p_2)$ and $Q = (q_1, q_2)$ with $q_1 \geq p_1$, $q_2 \geq p_2$, the number of such paths from P to Q is

$$\binom{(q_1 - p_1) + (q_2 - p_2)}{q_1 - p_1}.$$

Given sources $x_1, \ldots, x_m$ and sinks $y_1, \ldots, y_m$, let $a_{i,j}$ be the number of lattice paths from x_i to y_j and form the $m \times m$ matrix $A = (a_{i,j})$.

In the 2×2 example from the notes (two sources and two sinks), there are 12 ordered pairs of paths in total, but only 8 ordered pairs of vertex-disjoint paths. The path–count matrix is

$$A = \begin{pmatrix} \binom{2}{1} & \binom{4}{1} \\ \binom{4}{3} & \binom{6}{3} \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 4 & 20 \end{pmatrix},$$

and

$$\det A = 2 \cdot 20 - 4 \cdot 4 = 40 - 16 = 24.$$

After dividing by appropriate symmetries (depending on how we count ordered vs. unordered systems), this matches the number of vertex-disjoint path systems. This is a special case of the Lindström–Gessel–Viennot lemma.

Theorem 12.1 (LGV lemma)**.** For suitable acyclic directed graphs with sources $x_1, \ldots, x_m$ and sinks $y_1, \ldots, y_m$, the determinant of the matrix $A = (a_{i,j})$ of single-path counts equals a signed sum over all m-tuples of vertex-disjoint paths from the x_i to the y_j.

Consider a regular hexagon of side length n, subdivided into equilateral triangles of unit side length. A *rhombus tiling* of this hexagon is a tiling by rhombi consisting of two unit triangles.

Theorem 12.2 (MacMahon, 1916)**.** The number of rhombic tilings of a regular hexagon of side length n is equal to the determinant of the $n \times n$ matrix whose (i, j)–entry counts certain lattice walks (equivalently, non-intersecting paths) associated with the tiling, as illustrated in the notes. In particular this leads to MacMahon's famous product formula for plane partitions fitting in an $n \times n \times n$ box.

12.2 Burnside's Lemma

Let G be a finite group acting on a finite set C. For $\pi \in G$ let

$$\psi(\pi) := \left|\{c \in C : \pi \cdot c = c\}\right|$$

be the number of elements of C fixed by π.

Lemma 12.3 (Burnside)**.** The number of orbits of G on C equals

$$\frac{1}{|G|} \sum_{\pi \in G} \psi(\pi).$$

Proof. Let

$$S := \{(\pi, c) \in G \times C : \pi \cdot c = c\}.$$

We count $|S|$ in two ways.

Fix π. For a given $\pi \in G$, there are exactly $\psi(\pi)$ elements $c \in C$ with $\pi \cdot c = c$. Hence

$$|S| = \sum_{\pi \in G} \psi(\pi).$$

Fix c. For a given $c \in C$, let $G_c := \{\pi \in G : \pi \cdot c = c\}$ be the stabilizer of c. Then $|G_c|$ is the number of $\pi \in G$ with $(\pi, c) \in S$, so

$$|S| = \sum_{c \in C} |G_c|.$$

The orbit–stabilizer theorem says that

$$|G_c| = \frac{|G|}{|\operatorname{Orb}(c)|},$$

where $\operatorname{Orb}(c) = \{\pi \cdot c : \pi \in G\}$ is the orbit of c.

Let $\mathcal{O}$ be the set of orbits. Then

$$\sum_{c \in C} |G_c| = \sum_{O \in \mathcal{O}} \sum_{c \in O} |G_c| = \sum_{O \in \mathcal{O}} \sum_{c \in O} \frac{|G|}{|O|} = \sum_{O \in \mathcal{O}} |G| = |G| \cdot |\mathcal{O}|.$$

Equating the two expressions for $|S|$ gives

$$\sum_{\pi \in G} \psi(\pi) = |G| \cdot |\mathcal{O}| \quad \Longrightarrow \quad |\mathcal{O}| = \frac{1}{|G|} \sum_{\pi \in G} \psi(\pi),$$

as claimed. □

12.3 Colorings and cycle structure

Let X be a finite set and fix an integer $k \geq 1$. A *k-coloring* of X is a function $f : X \to \{1, \ldots, k\}$. Suppose G acts on X; the induced action on colorings is given by

$$(\pi \cdot f)(x) := f(\pi^{-1} x) \qquad (\pi \in G,\ x \in X).$$

Lemma 12.4. Let π be a permutation of X with t cycles in its cycle decomposition. Then the number of k-colorings $f : X \to [k]$ fixed by π is

$$\psi(\pi) = k^t.$$

Proof. Write the cycles of π as

$$C_1, \dots, C_t.$$

If f is fixed by π then f must be constant on each cycle: for $x \in C_j$ we have $\pi^m x = x$ for some $m \geq 1$, and

$$f(x) = f(\pi^m x) = f(\pi^{m-1} x) = \cdots = f(\pi x),$$

so all vertices in C_j have the same color. Conversely any coloring that is constant on each cycle is fixed by π.

Thus to specify a fixed coloring we may choose an arbitrary color in $[k]$ for each cycle C_j, independently. There are k choices per cycle, so in total k^t fixed colorings. □

Example 12.2 (Vertices of a square). Let X be the set of four vertices of a square, and let $G = D_4$ be the dihedral group of order 8 (all rotations and reflections of the square). We want the number of different k-colorings of the vertices up to symmetry.

We list the cycle structure of each type of symmetry on the vertex set:

type	# elements	cycle structure on vertices
identity	1	4 cycles of length 1
rot. by ±90°	2	1 cycle of length 4
rot. by 180°	1	2 cycles of length 2
reflection in a diagonal	2	1 fixed vertex, 1 2-cycle, 1 fixed vertex (i.e. 2 1-cycles ar
reflection in a vertical/horizontal axis	2	2 fixed vertices and 1 2-cycle

Using the lemma, we get

$$\psi(\text{id}) = k^4, \quad \psi(90^\circ) = \psi(270^\circ) = k, \quad \psi(180^\circ) = k^2, \quad \psi(\text{each reflection}) = k^3 \text{ or } k^2,$$

and summing over types gives

$$\#\text{orbits of colorings} = \frac{1}{8}(k^4 + 2k + 3k^2 + 2k^3).$$

12.4 Cycle index

Let G be a permutation group acting on a finite set X of size n. For $\pi \in G$, let $c_j(\pi)$ be the number of cycles of length j in the cycle decomposition of π.

Definition 12.4 (Cycle index). The *cycle index* of G is the polynomial

$$Z_G(x_1, \dots, x_n) := \frac{1}{|G|} \sum_{\pi \in G} x_1^{c_1(\pi)} x_2^{c_2(\pi)} \cdots x_n^{c_n(\pi)}.$$

Example 12.3 (Vertices of a square). For the action of D_4 on the 4 vertices, the cycle structures above yield

$$Z_{D_4}^{(\text{vertices})}(x_1, x_2, x_4) = \frac{1}{8}\Big(x_1^4 + 2x_4 + x_2^2 + 3x_1^2 x_2\Big).$$

Example 12.4 (Edges of a square). Let $G = D_4$ act on the 4 edges of a square. One checks that the cycle index for this action is

$$Z_{D_4}^{(\text{edges})}(x_1, x_2, x_4) = \frac{1}{8}(x_1^4 + x_2^2 + 2x_4 + 4x_1^2 x_2).$$

12.5 Pólya–Redfield counting

Let G act on X and consider colorings $f : X \to \{1, \ldots, k\}$. Give color i a weight y_i, and define the *weight* of a coloring by

$$w(f) = \prod_{x \in X} y_{f(x)}.$$

For each orbit of colorings, the weight of all colorings in that orbit is the same; let $P_G(y_1, \ldots, y_k)$ be the sum of these orbit weights.

Theorem 12.5 (Pólya–Redfield). With notation as above,

$$P_G(y_1, \ldots, y_k) = Z_G\big(y_1 + \cdots + y_k,\ y_1^2 + \cdots + y_k^2,\ \ldots,\ y_1^n + \cdots + y_k^n\big).$$

In particular, the number of orbits of k-colorings is obtained by substituting $y_1 = \cdots = y_k = 1$:

$$\#\{\text{inequivalent } k\text{-colorings}\} = Z_G(k, k, \ldots, k).$$

Proof. For a fixed $\pi \in G$ with cycle structure $c_1(\pi), \ldots, c_n(\pi)$, the colorings fixed by π are exactly those that are constant on each cycle. For a cycle of length j, the possible weights it contributes are $y_1^j, \ldots, y_k^j$, so the generating polynomial of colorings on that cycle is $y_1^j + \cdots + y_k^j$. Since different cycles are independent, the total weight of all colorings fixed by π is

$$\prod_{j=1}^{n} (y_1^j + \cdots + y_k^j)^{c_j(\pi)}.$$

By Burnside's lemma, the *sum of weights per orbit* equals

$$P_G(y_1, \ldots, y_k) = \frac{1}{|G|} \sum_{\pi \in G} \prod_{j=1}^{n} (y_1^j + \cdots + y_k^j)^{c_j(\pi)}.$$

This is exactly the cycle index Z_G with the substitution $x_j \leftarrow y_1^j + \cdots + y_k^j$, proving the formula. □

Example 12.5 (Edges of a square in two colors). Consider 2-colorings of the 4 edges of a square with colors $\{\text{red}, \text{blue}\}$, under the action of D_4.

Take $y_1 = r$ (for red) and $y_2 = b$ (for blue). Substituting into the cycle index above gives

$$P_{D_4}(r, b) = \frac{1}{8}\Big((r + b)^4 + (r^2 + b^2)^2 + 2(r^4 + b^4) + 4(r + b)^2(r^2 + b^2)\Big).$$

Setting $r = b = 1$ we get the number of inequivalent 2-colorings:

$$P_{D_4}(1, 1) = \frac{1}{8}(2^4 + 2^2 + 2 \cdot 2^2 + 4 \cdot 2^3) = 6.$$

So there are 6 distinct edge colorings up to symmetry.

12.6 Cube rotation group

Let G be the rotation group of a cube in $\mathbb{R}^3$. It has $|G| = 24$ elements.

Action on vertices, faces, and edges

The group G acts transitively on the 8 vertices, the 6 faces, and the 12 edges. One can classify rotations by their axes and compute cycle structures in each action. The resulting cycle indices are:

- Vertices ($|X| = 8$):
$$Z_V(G) = \frac{1}{24}\Big(x_1^8 + 9x_1^2x_2^3 + 8x_3^2 + 6x_4^2\Big).$$
- Faces ($|X| = 6$):
$$Z_F(G) = \frac{1}{24}\Big(x_1^6 + 6x_1^2x_4 + 3x_1^2x_2^2 + 8x_3^2 + 6x_2^3\Big).$$
- Edges ($|X| = 12$):
$$Z_E(G) = \frac{1}{24}\Big(x_1^{12} + 3x_1^4x_2^4 + 6x_1^4x_2^4 + 8x_3^4 + 6x_2^6\Big),$$
where the individual terms arise from rotations about face-centres, edge-centres and body diagonals.

Example 12.6. Using $Z_V(G)$ we can count colorings of cube vertices with k colors up to rotation:

$$\#\{\text{inequivalent } k\text{-colorings of vertices}\} = Z_V(G)(k,k,k,k,k,k,k,k) = \frac{1}{24}\big(k^8 + 9k^5 + 8k^2 + 6k^2\big).$$

Similar formulas hold for faces and edges.

12.8 Graphs up to isomorphism

Fix $n \geq 1$ and let $X = \binom{[n]}{2}$ be the set of unordered pairs of vertices, i.e. the edge set of the complete graph K_n. Any simple graph on vertex set $[n]$ is a subset $E \subseteq X$, so we can think of graphs as $\{0,1\}$-colorings of X, where color 1 means "edge present" and color 0 means "edge absent".

The symmetric group S_n acts on X and hence on graphs by relabelling vertices. Orbits under this action are exactly isomorphism classes of graphs on n vertices.

Let $Z^{(2)}_{S_n}$ denote the cycle index of the S_n-action on the set of 2-subsets X. Pólya–Redfield with colors $\{0,1\}$ then gives

$$\#\{\text{non-isomorphic graphs on } n \text{ vertices}\} = Z^{(2)}_{S_n}(2,2,\ldots,2).$$

Example 12.7 ($n = 4$). For $n = 4$, one computes

$$Z^{(2)}_{S_4}(x_1,x_2,x_3,x_4,x_6) = \frac{1}{24}\Big(x_1^6 + 3x_1^2x_2^2 + 8x_3^2 + 6x_1^2x_4 + 6x_2^3\Big).$$

Setting $x_j = 2$ for all j gives

$$\#\{\text{non-isomorphic graphs on 4 vertices}\} = \frac{1}{24}\big(2^6 + 3\cdot 2^4 + 8\cdot 2^2 + 6\cdot 2^3 + 6\cdot 2^3\big) = 11.$$

This matches the well-known fact that there are 11 unlabeled graphs on 4 vertices.

13 Basics of Graph Theory

Definition 13.1. A *graph* is a pair $G = (V, E)$ where $V = V(G)$ is a finite set of *vertices* and $E = E(G)$ is a set of 2–element subsets of V, called *edges*. If $uv \in E(G)$ we say that u and v are *adjacent*.

For a vertex $v \in V(G)$ the *degree* of v is

$$d_G(v) := |\{ u \in V(G) : uv \in E(G) \}|.$$

Definition 13.2. A graph G is

- *k-regular* if $d_G(v) = k$ for every $v \in V(G)$.
- of *order* n if $|V(G)| = n$, and of *size* m if $|E(G)| = m$.

Clearly $|E(G)| \le \binom{|V(G)|}{2}$.

Remark 13.1 (Conventions). Unless I explicitly say otherwise, graphs are finite, undirected, and simple (no loops, no parallel edges). When we later allow multigraphs, I will say so out loud.

Definition 13.3 (Path and cycle). Let $n \ge 1$.

- A *path* of length $n-1$ is a graph P_n with distinct vertices $v_1, \dots, v_n$ and edges v_iv_{i+1} for $i = 1, \dots, n-1$.
- A *cycle* of length n is a graph C_n with distinct vertices $v_1, \dots, v_n$ and edges v_iv_{i+1} for $i = 1, \dots, n-1$, together with v_nv_1.

A "path of length k" means k edges (not k vertices). So P_n has n vertices but length $n-1$.

13.1 Subgraphs and basic operations

Definition 13.4 (Subgraph). A graph H is a *subgraph* of G (written $H \subseteq G$) if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. If $V(H) = V(G)$ we call H a *spanning* subgraph of G.

Definition 13.5 (Induced subgraph). Given G and $X \subseteq V(G)$, the *induced subgraph* $G[X]$ is the graph with vertex set X and edge set

$$E(G[X]) = \{ uv \in E(G) : u, v \in X \}.$$

A subgraph H of G is *induced* if $H = G[V(H)]$.

Remark 13.2. A subgraph allows deleting some vertices and/or edges, however you like. An induced subgraph $G[X]$ is: you choose the vertex set X, and then you are *forced* to keep every edge of G whose endpoints both lie in X.

Definition 13.6 (Spanning subgraph)**.** A subgraph H of G is *spanning* if it keeps *all* the vertices:

$$V(H) = V(G) \qquad \text{and} \qquad E(H) \subseteq E(G).$$

In other words: you're allowed to delete edges, but you're not allowed to delete vertices.

Definition 13.7 (Vertex and edge deletion)**.** For $e \in E(G)$ let $G - e$ be the graph obtained by deleting e but keeping all vertices. For $v \in V(G)$ let $G - v$ be the graph obtained by deleting v and all edges incident with v.

Definition 13.8 (Neighborhood, isolated vertex)**.** For $v \in V(G)$ the *neighborhood* of v is

$$N_G(v) := \{ u \in V(G) : uv \in E(G) \}.$$

If $N_G(v) = \varnothing$ (equivalently, $d_G(v) = 0$) we call v an *isolated* vertex.

13.2 Complements, cliques and independent sets

Definition 13.9 (Complement)**.** The *complement* $\overline{G}$ of a graph G is the graph with

$$V(\overline{G}) = V(G), \qquad E(\overline{G}) = \{ uv : u \neq v,\ uv \notin E(G) \}.$$

Equivalently,

$$|E(\overline{G})| = \binom{|V(G)|}{2} - |E(G)|.$$

Definition 13.10 (Complete and empty graphs, cliques)**.** The *complete graph* K_n is the graph on n vertices with all possible edges.

A *clique* in G is a set $X \subseteq V(G)$ such that $G[X] \cong K_{|X|}$.

The *empty* graph on n vertices is the graph with vertex set of size n and no edges.

Definition 13.11 (Independent set)**.** A set $I \subseteq V(G)$ is an *independent set* if $G[I]$ has no edges (equivalently, no two vertices of I are adjacent).

Lemma 13.1 (Handshake Lemma)**.** For any finite (undirected) graph $G = (V, E)$,

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

In particular, the number of vertices of odd degree is even.

Proof. Count the set of *incidences*

$$I := \{(v, e) \in V \times E : e \text{ is incident to } v\}.$$

Fixing v, there are exactly $\deg(v)$ edges incident to v, so

$$|I| = \sum_{v \in V} \deg(v).$$

Fixing $e = \{u, v\}$, it is incident to exactly two vertices, so it contributes exactly 2 incidences; hence $|I| = 2|E|$. Therefore $\sum_{v \in V} \deg(v) = 2|E|$.

For the parity claim, reduce mod 2:

$$\sum_{v \in V} \deg(v) \equiv 0 \pmod 2.$$

Even-degree vertices contribute 0 mod 2 and odd-degree vertices contribute 1, so the number of odd-degree vertices is even. □

13.3 Bipartite and multipartite graphs

Definition 13.12 (Bipartite graphs)**.** A graph G is *bipartite* if its vertex set can be written as a disjoint union $V(G) = A \cup B$ such that both A and B are independent sets. In this case we also say that G is (A, B)*-bipartite* and write $G[A, B]$.

The *complete bipartite graph* $K_{a,b}$ has a bipartition A, B with $|A| = a$, $|B| = b$, and all possible edges between A and B.

Definition 13.13 (Complete k-partite graphs)**.** Let $U_1, \ldots, U_k$ be a partition of a finite set V. The *complete k-partite graph* with parts $U_1, \ldots, U_k$ is the graph G with vertex set V and edge set

$$E(G) = \{\, uv : u \in U_i,\ v \in U_j,\ i \neq j \,\}.$$

Each U_i is independent, and every pair of vertices in different parts are adjacent.

13.4 Matrices associated to a graph

Let G be a graph of order n with vertices $v_1, \ldots, v_n$ and edges $e_1, \ldots, e_m$.

Definition 13.14 (Adjacency matrix)**.** The *adjacency matrix* of G is the $n \times n$ matrix $A(G) = (a_{ij})$ where

$$a_{ij} = \begin{cases} 1, & v_i v_j \in E(G), \\ 0, & \text{otherwise.} \end{cases}$$

For a simple graph $A(G)$ is a symmetric $\{0, 1\}$-matrix with zeros on the diagonal.

Definition 13.15 (Incidence matrix). The *incidence matrix* of G is the $n \times m$ matrix $M(G) = (m_{ve})$ with rows indexed by vertices and columns by edges, where

$$m_{ve} = \begin{cases} 1, & v \text{ is an endpoint of edge } e, \\ 0, & \text{otherwise.} \end{cases}$$

13.5 Isomorphisms and automorphisms

Definition 13.16 (Graph isomorphism). Graphs G and H are *isomorphic* (written $G \cong H$) if there is a bijection $f : V(G) \to V(H)$ such that for all $u, v \in V(G)$

$$uv \in E(G) \iff f(u)f(v) \in E(H).$$

Such a map f is called an *isomorphism*.

Isomorphic graphs have the same order and size, and every isomorphism preserves degrees, paths, cycles, cliques, independent sets, etc.

Definition 13.17 (Automorphisms). An *automorphism* of a graph G is an isomorphism $f : V(G) \to V(G)$ from G to itself. The set of all automorphisms, with composition, forms a group $\mathrm{Aut}(G)$.

Definition 13.18 (Vertex- and edge-transitive). A graph G is

- *vertex-transitive* if for all $u, v \in V(G)$ there exists $\varphi \in \mathrm{Aut}(G)$ with $\varphi(u) = v$.
- *edge-transitive* if for all $e, f \in E(G)$ there exists $\varphi \in \mathrm{Aut}(G)$ with $\varphi(e) = f$.

Let C_n be the n-cycle with vertex set $\{1, \dots, n\}$ and edges $12, 23, \dots, (n-1)n, n1$.
Every automorphism of C_n is determined by the image of a single edge (or of the ordered pair $(1, 2)$), so

$$\mathrm{Aut}(C_n) \cong D_{2n},$$

the *dihedral group* of order $2n$, generated by a rotation σ of order n and a reflection δ with $\delta^2 = 1$, $\delta\sigma\delta = \sigma^{-1}$.

13.6 The Petersen graph

Let $[5] := \{1, 2, 3, 4, 5\}$ and let

$$V(P) = \binom{[5]}{2} = \{\{i, j\} : 1 \le i < j \le 5\}$$

be the set of 2-element subsets of $[5]$. Two vertices $A, B \in V(P)$ are adjacent if and only if they are disjoint as sets:

$$E(P) = \left\{ AB : A, B \in \tbinom{[5]}{2},\ A \cap B = \varnothing \right\}.$$

The resulting graph P is the *Petersen graph*.

Proposition 13.2. The Petersen graph P is 3-regular of order 10 and size 15.

Proof. There are $\binom{5}{2} = 10$ vertices (all 2-subsets of [5]). Fix a vertex $A \subset [5]$ with $|A| = 2$. A vertex B is adjacent to A exactly when $B \cap A = \varnothing$, i.e. B is a 2-subset of $[5] \setminus A$, which has size 3. Thus

$$d_P(A) = \binom{3}{2} = 3,$$

so P is 3-regular. By the Handshake Lemma,

$$2|E(P)| = \sum_{v \in V(P)} d_P(v) = 10 \cdot 3,$$

so $|E(P)| = 15$. □

Remark 13.3. The automorphism group of P contains all permutations of the ground set [5]; in fact

$$\mathrm{Aut}(P) \cong S_5.$$

Indeed, each $\pi \in S_5$ induces a permutation of the 2-subsets, preserving disjointness.

Remark 13.4. The Petersen graph is the smallest, most famous "counterexample graph": it is highly symmetric and 3-regular, yet it breaks many tempting conjectures (e.g. about Hamilton cycles and edge-colourings).

13.7 Girth and circumference

Definition 13.19. The *girth* $g(G)$ of a graph G is the length of a shortest cycle of G (or $+\infty$ if G is acyclic). The *circumference* $c(G)$ is the length of a longest cycle of G.

Proposition 13.3. The Petersen graph P has girth $g(P) = 5$.

Proof. First we show that P has no 3- or 4-cycles.

No triangles: Suppose A, B, C form a 3-cycle in P. Then $A, B, C \in \binom{[5]}{2}$ are pairwise disjoint (since adjacent vertices correspond to disjoint subsets). Hence

$$|A \cup B \cup C| = |A| + |B| + |C| = 6,$$

so $A \cup B \cup C$ would be a 6-element subset of [5], which is impossible. Thus P is triangle-free.

No 4-cycles. Suppose A, B, C, D form a 4-cycle

$$A \sim B \sim C \sim D \sim A.$$

Then $A \cap B = B \cap C = C \cap D = D \cap A = \varnothing$. Because P has no triangles, A and C cannot be adjacent, so $A \cap C \neq \varnothing$; similarly $B \cap D \neq \varnothing$.

Relabel the ground set so that $A = \{1, 2\}$. Then any neighbor of A is a 2-subset of $\{3, 4, 5\}$, so we may take $B = \{3, 4\}$ (the other choices are symmetric). Since $B \sim C$, the set C must be a 2-subset of $\{1, 2, 5\}$ disjoint from B, so $C \in \{\{1, 5\}, \{2, 5\}\}$ (it cannot be $\{1, 2\} = A$).

Assume $C = \{1,5\}$; the other case is analogous. Because $C \sim D$ and $D \sim A$, the set D must be disjoint from both $\{1,5\}$ and $\{1,2\}$, so

$$D \subseteq [5] \setminus \{1,5\} = \{2,3,4\}, \qquad D \subseteq [5] \setminus \{1,2\} = \{3,4,5\}.$$

Thus $D \subseteq \{3,4\}$ and $|D| = 2$, so $D = \{3,4\} = B$, contradicting that the four vertices on the cycle are distinct. Hence P has no 4-cycle.

We have shown that $g(P) \geq 5$. On the other hand, P contains the 5-cycle

$$12 \sim 34 \sim 15 \sim 23 \sim 45 \sim 12,$$

so $g(P) = 5$. □

13.8 Kneser Graph

Definition 13.20 (Kneser graph). Fix integers $n \geq 1$ and $1 \leq k \leq n$. The *Kneser graph* $K(n,k)$ has

$$V(K(n,k)) = \binom{[n]}{k}$$

(the set of all k-subsets of $[n] = \{1,2,\dots,n\}$), and two vertices $A, B \in \binom{[n]}{k}$ are adjacent iff they are disjoint:

$$A \sim B \iff A \cap B = \varnothing.$$

(*Note:* if $n < 2k$, then there are no disjoint k-subsets, so $K(n,k)$ has no edges.)

Immediate observations.

- If $n < 2k$, then no two k-subsets are disjoint, hence $E(K(n,k)) = \varnothing$.
- If $n = 2k$, then each $A \in \binom{[n]}{k}$ has a unique disjoint partner $[n] \setminus A$, so $K(2k,k)$ is a 1-regular graph.

Basic parameters (for $n \geq 2k$).

- Number of vertices:
$$|V(K(n,k))| = \binom{n}{k}.$$
- Regular degree: for any $A \in \binom{[n]}{k}$,
$$\deg_{K(n,k)}(A) = \binom{n-k}{k},$$
since a neighbor of A is a k-subset chosen from the remaining $n-k$ elements.
- Number of edges:
$$|E(K(n,k))| = \frac{1}{2}|V(K(n,k))| \cdot \deg(K(n,k)) = \frac{1}{2}\binom{n}{k}\binom{n-k}{k}.$$
- Vertex-transitivity: the symmetric group S_n acts on $\binom{[n]}{k}$ by permuting $[n]$, and preserves disjointness, so $K(n,k)$ is vertex-transitive.

The Petersen graph as a special case. Let $n = 5$ and $k = 2$. Then

$$|V(K(5,2))| = \binom{5}{2} = 10, \qquad \deg(K(5,2)) = \binom{5-2}{2} = \binom{3}{2} = 3, \qquad |E(K(5,2))| = \tfrac{1}{2} \cdot 10 \cdot 3 = 15.$$

Thus $K(5,2)$ is a 3-regular graph on 10 vertices with 15 edges; this graph is the *Petersen graph*.

13.9 The k-dimensional hypercube.

Definition 13.21 (Hypercube Q_k). Fix $k \geq 1$. The *k-dimensional hypercube Q_k* is the graph with

$$V(Q_k) = \{0,1\}^k,$$

and for $x = (x_1, \dots, x_k)$ and $y = (y_1, \dots, y_k)$,

$$x \sim y \quad \Longleftrightarrow \quad x \text{ and } y \text{ differ in exactly one coordinate.}$$

Identify each vertex $x \in \{0,1\}^k$ with the subset

$$A_x := \{\, i \in [k] : x_i = 1 \,\} \subseteq [k],$$

so that $V(Q_k) \cong 2^{[k]}$ via characteristic vectors. Under this identification, for $A, B \subseteq [k]$,

$$A \sim B \quad \Longleftrightarrow \quad |A \triangle B| = 1 \quad \Longleftrightarrow \quad \big(A \subseteq B \text{ or } B \subseteq A\big) \text{ and } \big||A| - |B|\big| = 1.$$

Remark 13.5. Any permutation of coordinates is an automorphism of Q_k, and any independent bit-flip

$$(x_1, \dots, x_k) \mapsto (x_1 \oplus \varepsilon_1, \dots, x_k \oplus \varepsilon_k), \qquad \varepsilon \in \{0,1\}^k,$$

is also an automorphism. In particular,

$$|\mathrm{Aut}(Q_k)| = 2^k\, k!.$$

Definition 13.22 (Cartesian product $G \square H$). Let G and H be graphs. The *Cartesian product* $G \square H$ is the graph with

$$V(G \square H) = V(G) \times V(H),$$

and

$$(u,v) \sim (u',v') \quad \Longleftrightarrow \quad \big(u = u' \text{ and } vv' \in E(H)\big) \text{ or } \big(v = v' \text{ and } uu' \in E(G)\big).$$

The hypercube is an iterated Cartesian product:

$$Q_k \cong \underbrace{K_2 \square K_2 \square \cdots \square K_2}_{k \text{ factors}}.$$

For every $k \geq 1$,

$$Q_{k+1} \cong Q_k \square K_2.$$

Definition 13.23 (Union of (labeled) graphs). Let G and H be graphs (think: their vertex sets are actual labels, not "up to isomorphism"). The *union* $G \cup H$ is the graph with

$$V(G \cup H) = V(G) \cup V(H), \qquad E(G \cup H) = E(G) \cup E(H).$$

Remark 13.6. If $V(G)$ and $V(H)$ overlap, then $G \cup H$ *identifies* those common vertices (same labels). If you want two separate copies with no identification, use disjoint union.

Definition 13.24 (Disjoint union). If $V(G) \cap V(H) = \varnothing$, the *disjoint union* (also written $G+H$) is just the union:

$$G + H := G \cup H \quad (\text{when } V(G) \cap V(H) = \varnothing).$$

Equivalently, $G + H$ consists of two connected components isomorphic to G and H.

Definition 13.25 (Join). If $V(G) \cap V(H) = \varnothing$, the *join* of G and H, denoted $G \vee H$, is the graph with

$$V(G \vee H) = V(G) \cup V(H),$$

and

$$E(G \vee H) = E(G) \ \cup \ E(H) \ \cup \ E\big[V(G), V(H)\big],$$

where

$$E\big[V(G), V(H)\big] := \big\{ uv : \ u \in V(G), \ v \in V(H) \big\}$$

is the set of *all cross-edges* between $V(G)$ and $V(H)$.

Definition 13.26 (m copies of a graph). For an integer $m \geq 1$, define

$$m \cdot G \ := \ \underbrace{G + G + \cdots + G}_{m \text{ times}},$$

i.e. the disjoint union of m vertex-disjoint copies of G.

14 Vertex Degrees

Recall the Handshake Lemma:

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

As a result, the sum of all vertex degrees is even, and therefore the number of vertices of odd degree is even.

Here is a geometric problem that looks like it should involve coordinates and area, but actually collapses under a simple parity argument.

Theorem 14.1 (Integer side forced by an axis-parallel tiling)**.** Let R be a rectangle in the plane whose sides are parallel to the coordinate axes. Suppose R is partitioned (tiled) into rectangles $R_1, \dots, R_m$ whose interiors are disjoint, whose union is R, and whose sides are also parallel to the axes. Assume that for every i, the rectangle R_i has *at least one* side of integer length. Then R has at least one side of integer length.

We are going to build a bipartite graph whose degrees encode "how many integer lattice corners" each tile has. All tile-vertices will have even degree. The Handshake Lemma then forces the lattice-point side to have an even sum of degrees. Since one specific lattice point has odd degree, some other lattice point must also have odd degree. Finally, we show that the *only* lattice points that can have odd degree are the corners of R, which forces one of those corners to be an integer lattice point, and therefore forces W or H to be an integer.

Proof. Write the big rectangle as

$$R = [0, W] \times [0, H] \qquad \text{for some } W, H > 0,$$

so $(0, 0)$ is its lower-left corner.

Build a bipartite graph. Let $A = \{R_1, \dots, R_m\}$. Let

$$B := R \cap \mathbb{Z}^2$$

be the set of integer grid points inside (or on the boundary of) R. Define a bipartite graph $G = (A, B; E)$ by joining $R_i \in A$ to $p \in B$ iff p is a corner of R_i.

Every rectangle vertex has even degree. Every tile has even degree. Fix a tile R_i. Its corners have the form

$$(x_1, y_1),\ (x_2, y_1),\ (x_1, y_2),\ (x_2, y_2),$$

with $x_1 < x_2$ and $y_1 < y_2$. The hypothesis says that at least one of the side lengths $x_2 - x_1$ or $y_2 - y_1$ is an integer.

If R_i has *any* integer lattice corner, say $(x_1, y_1) \in \mathbb{Z}^2$, then:

- if $x_2 - x_1 \in \mathbb{Z}$, then $(x_2, y_1) = (x_1 + (x_2 - x_1), y_1) \in \mathbb{Z}^2$;
- if $y_2 - y_1 \in \mathbb{Z}$, then $(x_1, y_2) = (x_1, y_1 + (y_2 - y_1)) \in \mathbb{Z}^2$.

So an integer corner forces a *second* integer corner. In particular, a tile cannot have exactly 1 or 3 integer corners. Therefore

$$\deg_G(R_i) \in \{0, 2, 4\} \qquad \text{for every } R_i \in A,$$

and so $\sum_{R_i \in A} \deg_G(R_i)$ is even.

Because G is bipartite, the handshake lemma gives

$$\sum_{R_i \in A} \deg_G(R_i) = \sum_{p \in B} \deg_G(p).$$

The left-hand sum is even, hence the right-hand sum is even as well.

Claim: $(0,0)$ has odd degree. Exactly one tile R_i contains the corner $(0,0)$ of the big rectangle R, so $(0,0)$ is a corner of exactly one R_i and thus

$$\deg_G((0,0)) = 1.$$

Since $\sum_{p \in B} \deg_G(p)$ is even but includes the odd term $\deg_G((0,0)) = 1$, there must exist another grid point $q \in B$ with odd degree.

Claim: Any grid point that is *not* a corner of R has even degree. Let $p \in B$ be a grid point that is not a corner of R. Looking in a small neighborhood of p, the tiling is by axis-parallel rectangles, so rectangles can meet at p as corners in 0, 2, or 4 local quadrants only; in particular,

$$\deg_G(p) \in \{0, 2, 4\}.$$

Thus every non-corner grid point has even degree.

Therefore the odd-degree point q must be a *corner* of R. So one of $(W,0)$, $(0,H)$, (W,H) lies in $\mathbb{Z}^2$. In particular, either $W \in \mathbb{Z}$ or $H \in \mathbb{Z}$ (or both), so R has an integer-length side. □

14.1 Graphic Sequences

A degree sequence is what you get when you forget everything about a graph except how many neighbors each vertex has. The natural inverse problem is: *given a list of degrees, does any simple graph realize it?*

Definition 14.1 (Degree sequence). Let G be a graph, $V(G) = \{v_1, \dots, v_n\}$. The *degree sequence* of G is the list

$$(\deg(v_1), \deg(v_2), \dots, \deg(v_n)).$$

Usually we sort it in nonincreasing order and write

$$d_1 \geq d_2 \geq \cdots \geq d_n,$$

and call $(d_1, \dots, d_n)$ *the* degree sequence.

Two immediate sanity checks. If $(d_1, \dots, d_n)$ is the degree sequence of a simple graph, then:

- Bounds: $0 \leq d_i \leq n-1$ for all i.
- Handshake parity: $\sum_{i=1}^{n} d_i = 2|E|$ is even, so the number of odd d_i is even.

These are necessary conditions. They are *not* sufficient.

Example 14.1. $(3, 3, 3, 1)$ passes the parity test (sum is 10 even) and the bounds (≤ 3), but it is impossible: the vertex of degree 1 can only connect to one of the three degree-3 vertices, and then the remaining two degree-3 vertices can't both reach degree 3.

Given a nonincreasing integer sequence

$$d_1 \geq d_2 \geq \cdots \geq d_n \geq 0,$$

when does there exist a simple graph G on n vertices whose degrees are exactly these numbers?

Definition 14.2 (Graphic sequence). A sequence $(d_1, \ldots, d_n)$ of nonnegative integers is *graphic* if there exists a simple graph G with degree sequence $(d_1, \ldots, d_n)$. Such a graph G is called a *realization* of the sequence.

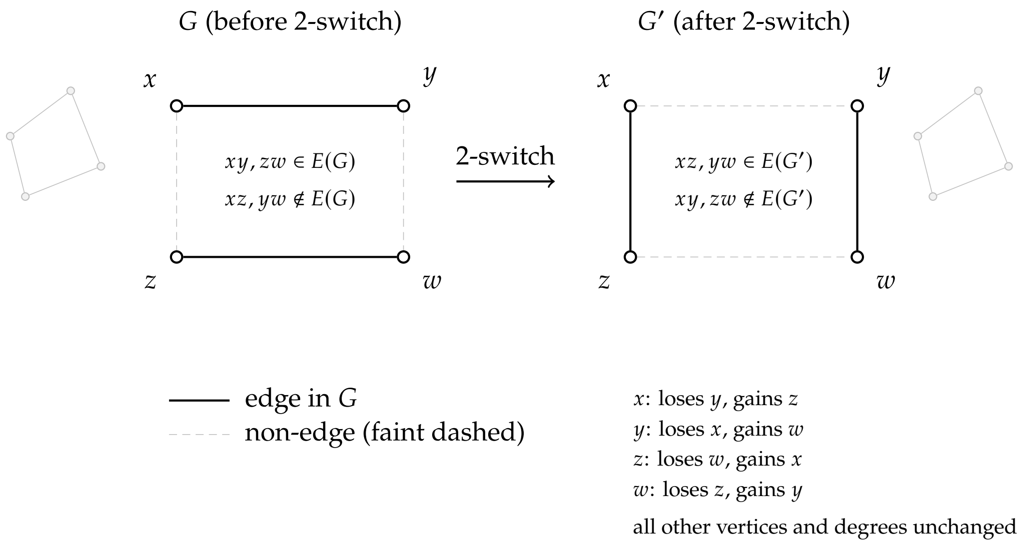
Definition 14.3 (2-switch). A *2-switch* in G is the following operation: pick four *distinct* vertices x, y, z, w such that

$$xy,\ zw \in E(G) \qquad \text{and} \qquad xz,\ yw \notin E(G).$$

Form a new graph G' by *deleting* the edges xy and zw and *adding* the edges xz and yw:

$$E(G') = \big(E(G) \setminus \{xy, zw\}\big) \cup \{xz, yw\}.$$

We say G' is obtained from G by a 2-switch on (x, y, z, w).



Remark 14.1 (Degrees do not change). A 2-switch preserves degrees:

$$\deg_{G'}(v) = \deg_G(v) \qquad \text{for all vertices } v.$$

We reduce the problem "is $d = (d_1, \ldots, d_n)$ graphic?" to a smaller instance. The obvious move is: take a vertex of degree d_1, connect it to d_1 other vertices, then delete it and decrease those d_1 degrees by 1. The Havel–Hakimi theorem says this greedy step is both necessary and sufficient: d is graphic exactly when the resulting shorter sequence is graphic.

14.2 Havel-Hakimi Theorem for graphic sequences

Theorem 14.2 (Havel–Hakimi)**.** Let $d = (d_1, \ldots, d_n)$ be a nonincreasing sequence of nonnegative integers:

$$d_1 \geq d_2 \geq \cdots \geq d_n \geq 0.$$

Define the *reduced sequence* $d' = (d'_1, \ldots, d'_{n-1})$ by

$$d' = \operatorname{sort}\big(d_2 - 1,\ d_3 - 1,\ \ldots,\ d_{d_1+1} - 1,\ d_{d_1+2}, \ldots, d_n\big),$$

i.e. subtract 1 from the next d_1 entries and then re-sort into nonincreasing order. (If $d_1 > n-1$ or some $d_i - 1 < 0$, then d' is declared *invalid*.)

Then d is graphic if and only if d' is graphic.

The problem is that the degree sequence only tells us *how many* neighbors each vertex has, not *which* ones, so in an arbitrary realization the vertex of degree d_1 might be joined to some other set T of d_1 vertices, not necessarily $\{v_2, \ldots, v_{d_1+1}\}$.

The following lemma says we can always rearrange edges by 2-switches so that this is true, without changing any degrees (and hence, the degree sequence).

Lemma 14.3. Let G be a simple graph on vertices $v_1, \ldots, v_n$ with degree sequence

$$d_1 \geq d_2 \geq \cdots \geq d_n, \qquad \text{where } \deg_G(v_i) = d_i.$$

Let $w := v_1$ and let

$$S := \{v_2, v_3, \ldots, v_{d_1+1}\}.$$

Then there exists a graph $\widetilde{G}$ obtained from G by a finite sequence of 2-switches such that

$$\deg_{\widetilde{G}}(v_i) = d_i \text{ for all } i \qquad \text{and} \qquad N_{\widetilde{G}}(w) = S.$$

Proof. Among all graphs obtainable from G by 2-switches (hence with the same degree sequence), choose one (call it again G) that maximizes the quantity

$$t := |N_G(w) \cap S|.$$

We show $t = d_1$, i.e. $S = N_G(w)$.

Assume for contradiction that $t < d_1$. Then there exist

$$x \in S \setminus N_G(w) \qquad \text{and} \qquad y \in N_G(w) \setminus S.$$

Indeed, $|N_G(w)| = |S| = d_1$, so if $S \neq N_G(w)$, some neighbor of w must lie outside S.

Since $x \in S$ and $y \notin S$ and the degree sequence is nonincreasing, we have

$$\deg_G(x) \geq \deg_G(y).$$

Now consider the set difference $N_G(x) \setminus N_G(y)$. If $N_G(x) \subseteq N_G(y)$, then $\deg_G(x) \le \deg_G(y)$, contradicting $\deg_G(x) \ge \deg_G(y)$ unless $\deg_G(x) = \deg_G(y)$ and $N_G(x) = N_G(y)$. But even in that equality case we still get a contradiction as follows: because $y \in N_G(w)$ and $x \notin N_G(w)$, the vertex w is in $N_G(y)$ but not in $N_G(x)$, hence

$$w \in N_G(y) \setminus N_G(x),$$

so $N_G(x) \neq N_G(y)$ and therefore $N_G(x) \not\subseteq N_G(y)$. Thus in all cases,

$$N_G(x) \setminus N_G(y) \neq \varnothing.$$

Choose $z \in N_G(x) \setminus N_G(y)$. Then

$$xz \in E(G), \qquad yz \notin E(G).$$

Also by construction,

$$wy \in E(G), \qquad wx \notin E(G).$$

The four vertices w, x, y, z are distinct: $z \neq x$ (since xz is an edge), $z \neq y$ (since yz is a non-edge), and $z \neq w$ because $w \in N_G(y)$ but $z \notin N_G(y)$.

Therefore we may perform the 2-switch that deletes wy and xz and adds wx and yz:

$$G' := G - \{wy, xz\} + \{wx, yz\}.$$

This is a valid 2-switch because the added edges were non-edges in G. Moreover G' has the same degree sequence as G. Moreover, w loses neighbor y and gains neighbor x, so

$$|N_{G'}(w) \cap S| = |N_G(w) \cap S| + 1 = t + 1,$$

because $x \in S$ and $y \notin S$. This contradicts the maximality of t.

Hence our assumption $t < d_1$ was false, so $t = d_1$ and $N_G(w) = S$. Set $\widetilde{G} := G$. □

Proof of Havel-Hakimi Theorem

Proof. (⇐) Suppose d' is graphic. Realize the unsorted sequence $\widehat{d}$ by a simple graph H on vertices $v_2, \dots, v_n$ with

$$\deg_H(v_i) = \begin{cases} d_i - 1, & 2 \le i \le d_1 + 1, \\ d_i, & d_1 + 2 \le i \le n. \end{cases}$$

Add a new vertex v_1 and connect it to $v_2, v_3, \dots, v_{d_1+1}$. Then $\deg(v_1) = d_1$, and for $2 \le i \le d_1 + 1$ we increase $\deg_H(v_i)$ by 1, restoring $\deg(v_i) = d_i$, while other degrees stay d_i. Hence the resulting graph realizes d.

(⇒) Suppose d is graphic and let G be a realization on vertices $v_1, \dots, v_n$ with $\deg_G(v_i) = d_i$. Apply the previous lemma with $w = v_1$ and $S = \{v_2, \dots, v_{d_1+1}\}$ to obtain a graph $\widetilde{G}$ (with the same degree sequence) such that

$$N_{\widetilde{G}}(v_1) = \{v_2, \dots, v_{d_1+1}\}.$$

Now delete v_1 to form $\widetilde{G} - v_1$. In $\widetilde{G} - v_1$, each vertex v_i for $2 \le i \le d_1 + 1$ loses exactly the edge $v_1 v_i$, so its degree becomes $d_i - 1$, and every vertex v_i for $i \ge d_1 + 2$ keeps degree d_i. Therefore $\widetilde{G} - v_1$ realizes $\widehat{d}$ (hence also realizes its sorted version d'). Thus d' is graphic. □

The 2-switch operation lets us change adjacencies without changing the degree sequence. A natural question is whether *all* realizations of a fixed degree sequence (on the same labeled vertex set) can be connected by a sequence of 2-switches. The answer is *yes*.

Theorem 14.4. Fix a degree sequence on labeled vertices $v_1, \ldots, v_n$:

$$\deg(v_i) = d_i \quad (1 \le i \le n).$$

If G and H are two simple graphs on $\{v_1, \ldots, v_n\}$ with these degrees, then G can be transformed into H by a finite sequence of 2-switches.

Proof. By induction on n. The cases $n \le 3$ can be verified easily.

Assume $n \ge 4$. Let $w = v_1$ and $S = \{v_2, \ldots, v_{d_1+1}\}$. By the lemma, there exist graphs G^* and H^* obtainable from G and H (respectively) by 2-switches such that

$$N_{G^*}(w) = S \qquad \text{and} \qquad N_{H^*}(w) = S,$$

and all degrees remain d_i.

Now delete w from both graphs. The resulting graphs $G^* - w$ and $H^* - w$ are simple graphs on $\{v_2, \ldots, v_n\}$ with the same degree sequence

$$(d_2 - 1, \ldots, d_{d_1+1} - 1, d_{d_1+2}, \ldots, d_n)$$

(up to sorting). By the induction hypothesis, there is a sequence of 2-switches that transforms $G^* - w$ into $H^* - w$. Perform the same 2-switches in G^* (none of them needs to involve w, since w is absent from the smaller graphs). This transforms G^* into H^* while keeping $N(w) = S$ throughout.

Finally, concatenate the 2-switch sequence from G to G^*, then from G^* to H^*, then reverse the sequence from H to H^* to go from H^* to H. This yields a 2-switch sequence from G to H. □

The theorem is constructive: it gives a simple decision procedure (and, when it succeeds, a way to build a realization).

Havel–Hakimi algorithm. Given a nonincreasing sequence $d = (d_1, \ldots, d_n)$:

1. If all entries are 0, return GRAPHIC.
2. If $d_1 > n - 1$ or some entry is negative, return NOT GRAPHIC.
3. Subtract 1 from each of $d_2, \ldots, d_{d_1+1}$ and leave $d_{d_1+2}, \ldots, d_n$ unchanged. Sort the resulting sequence and set it to d in nonincreasing order.
4. Go back to Step 1.

If the process halts in Step 1, the original sequence is graphic; if it halts in Step 2, it is not.

14.3 Extremal problems

In extremal graph theory, you fix some *parameter* (like $|E(G)|$, $\delta(G)$, $\alpha(G)$, $\chi(G)$, etc.) and you ask for the *minimum* or *maximum* possible value among all graphs satisfying some constraint (e.g. “triangle-free”, “bipartite”, “no K_r”, “given order n”, etc.).

Example 14.2. Among all triangle-free graphs on n vertices, what is the maximum possible number of edges?

Mantel's theorem answers this: any triangle-free graph on n vertices has at most

$$\left\lfloor \frac{n^2}{4} \right\rfloor$$

edges, with equality for the complete bipartite graph $K_{\lfloor n/2\rfloor,\lceil n/2\rceil}$.

14.4 Existence of large bipartite subgraph

Theorem 14.5. Every graph G with m edges contains a bipartite subgraph with at least $m/2$ edges.

Proof 1 (probabilistic). Choose a random partition (A, B) of $V(G)$ by putting each vertex independently into A with probability $1/2$ (and into B otherwise). For an edge $xy \in E(G)$, let i_{xy} be the indicator of the event "xy is a cross-edge":

$$i_{xy} = \begin{cases} 1, & \text{if } x \in A,\ y \in B \text{ or } x \in B,\ y \in A, \\ 0, & \text{otherwise.} \end{cases}$$

For any fixed edge xy, exactly two of the four equally-likely placements of (x, y) make it a cross-edge, so

$$\mathbb{P}(i_{xy} = 1) = \frac{1}{2} \qquad \Longrightarrow \qquad \mathbb{E}[i_{xy}] = \frac{1}{2}.$$

Let X be the number of cross-edges under the random partition:

$$X = \sum_{xy \in E(G)} i_{xy}.$$

By linearity of expectation,

$$\mathbb{E}[X] = \sum_{xy \in E(G)} \mathbb{E}[i_{xy}] = \sum_{xy \in E(G)} \frac{1}{2} = \frac{m}{2}.$$

Therefore there exists *some* partition (A, B) for which $X \geq m/2$. The subgraph consisting of all cross-edges for that partition is bipartite (with parts A, B) and has $\geq m/2$ edges. □

Remark 14.2. This proof is pure existence: it guarantees a good partition exists, but it does not explicitly tell you how to find it. (We fix that next.)

Proof 2 (algorithmic). Start with any partition (A, B) of $V(G)$. For a vertex x, write

$$d_A(x) := \big|\{xy \in E(G) : y \in A\}\big|, \qquad d_B(x) := \big|\{xy \in E(G) : y \in B\}\big|.$$

So $d_A(x)$ counts neighbors of x on the A-side, and $d_B(x)$ counts neighbors on the B-side.

Improvement rule.

- If $x \in A$ and $d_A(x) > d_B(x)$, move x from A to B.
- If $x \in B$ and $d_B(x) > d_A(x)$, move x from B to A.

Why this increases cross-edges. Assume $x \in A$ and we move it to B. Before the move, edges from x to B were cross-edges (count $d_B(x)$) and edges from x to A were inside-edges (count $d_A(x)$). After the move, these roles swap: x has $d_A(x)$ cross-edges and $d_B(x)$ inside-edges. So the number of cross-edges changes by

$$\Delta = d_A(x) - d_B(x) > 0.$$

Thus every move strictly increases the number of cross-edges.

Termination. The number of cross-edges is an integer between 0 and m, and it strictly increases each move, so the process must stop.

At a local optimum. When the process stops, we have for every $x \in A$ that $d_A(x) \le d_B(x)$, and for every $x \in B$ that $d_B(x) \le d_A(x)$.

Sum these inequalities over each side:

$$\sum_{x \in A} d_A(x) \le \sum_{x \in A} d_B(x), \qquad \sum_{x \in B} d_B(x) \le \sum_{x \in B} d_A(x).$$

Add them:

$$\sum_{x \in A} d_A(x) + \sum_{x \in B} d_B(x) \le \sum_{x \in A} d_B(x) + \sum_{x \in B} d_A(x).$$

Now interpret each side:

- $\sum_{x \in A} d_A(x) = 2e(A)$ where $e(A)$ is the number of edges inside A.
- $\sum_{x \in B} d_B(x) = 2e(B)$ where $e(B)$ is the number of edges inside B.
- $\sum_{x \in A} d_B(x) = \sum_{x \in B} d_A(x) = e(A, B)$ where $e(A, B)$ is the number of cross-edges.

So the inequality becomes

$$2e(A) + 2e(B) \le 2e(A, B) \quad \Longleftrightarrow \quad e(A) + e(B) \le e(A, B).$$

But

$$m = e(A) + e(B) + e(A, B),$$

hence

$$m \le 2e(A, B) \quad \Longleftrightarrow \quad e(A, B) \ge \frac{m}{2}.$$

Therefore the final partition (A, B) produced by the algorithm has at least $m/2$ cross-edges, so the bipartite subgraph induced by these cross-edges has at least $m/2$ edges. □

14.5 Turan's Theorem

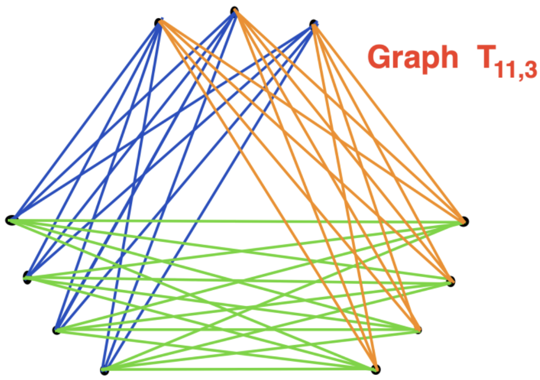
Turán's theorem is the natural generalization of Mantel's theorem: instead of forbidding triangles, we forbid K_{r+1} subgraphs for some fixed $r \ge 2$. The question is the same:

What is the maximum number of edges in an n-vertex graph that does not contain K_{r+1} as a subgraph?

Turán's theorem says that the extremal graphs are exactly the complete r-partite graphs with parts as equal in size as possible (the *Turán graphs*), and it gives an explicit formula for the maximum number of edges.

Definition 14.4 (Turán graph $T_{n,r}$). Fix integers $n \geq 1$ and $r \geq 1$. The *Turán graph* $T_{n,r}$ is the complete r-partite graph on n vertices whose part sizes differ by at most 1. Equivalently, write $n = qr + s$ with $0 \leq s < r$; then $T_{n,r}$ has s parts of size $q+1$ and $r-s$ parts of size q.

Definition 14.5. Let $f(n,r)$ denote the maximum number of edges in a simple n-vertex graph with no K_{r+1} subgraph.



Theorem 14.6 (Turán's Theorem). For all $n, r \geq 1$,

$$f(n,r) = |E(T_{n,r})| = \frac{1}{2}\left(1 - \frac{1}{r}\right)n^2 - \frac{s(r-s)}{2r} \qquad \text{where } s = n - r\left\lfloor \frac{n}{r} \right\rfloor$$

In particular, among all K_{r+1}-free graphs on n vertices, $T_{n,r}$ has the most edges.

Lemma 14.7. If $n \geq r+1$, then

$$|E(T_{n,r})| - |E(T_{n-r,r})| = (r-1)(n-r) + \binom{r}{2}.$$

Proof of lemma. We want to obtain $T_{n,r}$ from $T_{n-r,r}$ by adding one new vertex to each part. Then the difference in edge counts is exactly the number of edges incident to these new vertices.

Start with a copy G of $T_{n-r,r}$ whose parts are $V_1, \ldots, V_r$ (so $\sum_{i=1}^{r} |V_i| = n - r$). Form a graph H on n vertices by adding, for each $i \in [r]$, one new vertex v_i and placing v_i into part V_i. Join v_i to every vertex outside V_i (i.e. to all $n - r - |V_i|$ old vertices in $G - V_i$) and to all other new vertices v_j for $j \neq i$. Then H is exactly $T_{n,r}$.

(a) Edges among the new vertices. The vertex set $\{v_1, \ldots, v_r\}$ spans a complete graph K_r, so this contributes $\binom{r}{2}$ new edges.

(b) Edges from new vertices to old vertices. Fix $i \in [r]$. The new vertex v_i is adjacent to every old vertex not in V_i, so it has $n - r - |V_i|$ neighbors among the old vertices. Summing over all i gives

$$\sum_{i=1}^{r}(n - r - |V_i|) = \sum_{i=1}^{r}(n-r) - \sum_{i=1}^{r}|V_i| = r(n-r) - (n-r) = (r-1)(n-r),$$

using $\sum_{i=1}^{r} |V_i| = n - r$.
Hence

$$|E(T_{n,r})| - |E(T_{n-r,r})| = \binom{r}{2} + (r-1)(n-r),$$

as claimed. □

Proof of Turán's Theorem

Proof. First note that $T_{n,r}$ is K_{r+1}-free (an $(r+1)$-clique would require two vertices from the same part, but there are no edges inside a part). Therefore $f(n,r) \geq |E(T_{n,r})|$.

It remains to prove the reverse inequality

$$f(n,r) \leq |E(T_{n,r})|. \tag{5}$$

Fixing r, we prove (5) by induction on n.

Base case: $n \leq r$. Any n-vertex graph is automatically K_{r+1}-free, because a copy of K_{r+1} would need $r+1$ distinct vertices, but we only have $n \leq r$. So there is no restriction at all, and the maximum number of edges is attained by the complete graph:

$$f(n,r) = \binom{n}{2} = |E(K_n)|.$$

Then every n-vertex graph is automatically K_{r+1}-free, so $f(n,r) = \binom{n}{2}$, attained by K_n. On the other hand, when $n \leq r$, the Turán graph $T_{n,r}$ is also just K_n: we have r parts and at most one vertex per part, so every pair of vertices lies in different parts and hence every edge is present. Thus $|E(T_{n,r})| = \binom{n}{2}$, and (5) holds in this case.

Induction step: $n \geq r+1$. Assume (5) holds for all smaller values $n' < n$. Let G be a K_{r+1}-free simple graph on n vertices with

$$|E(G)| = f(n,r).$$

Let Q be the vertex set of a *largest clique* in G; then $|Q| \leq r$. Choose any set $Q' \subseteq V(G) \setminus Q$ of size $r - |Q|$ and define

$$F := Q \cup Q'.$$

So $|F| = r$ and $Q \subseteq F$.

Because Q is a maximum clique, every vertex $z \in V(G) \setminus Q$ is adjacent to at most $|Q| - 1$ vertices of Q; otherwise $Q \cup \{z\}$ would be a larger clique. Now for any $z \in V(G) \setminus F$ we have

$$\deg_F(z) = |N_G(z) \cap F| = |N_G(z) \cap Q| + |N_G(z) \cap Q'|.$$

The first term is at most $|Q| - 1$ by maximality of Q, and the second term is at most $|Q'| = r - |Q|$. Thus

$$\deg_F(z) \leq (|Q| - 1) + (r - |Q|) = r - 1 \qquad \text{for all } z \in V(G) \setminus F.$$

Summing over all $n - r$ vertices in $V(G) \setminus F$ gives

$$e(F,\ V(G) \setminus F) \leq (r-1)(n-r). \tag{a}$$

Inside F there are at most $\binom{r}{2}$ edges, since $|F| = r$:

$$e(F) \le \binom{r}{2}. \tag{b}$$

Let

$$G' := G - F$$

be the induced subgraph on $V(G) \setminus F$. Then G' has $n - r$ vertices and is still K_{r+1}-free, so by the induction hypothesis,

$$|E(G')| \le f(n-r, r) = |E(T_{n-r,r})|. \tag{c}$$

Every edge of G lies either inside G', between F and $V(G) \setminus F$, or inside F, so

$$|E(G)| = |E(G')| + e(F, V(G) \setminus F) + e(F).$$

Using (a), (b), and (c), we get

$$|E(G)| \le |E(T_{n-r,r})| + (r-1)(n-r) + \binom{r}{2}.$$

By the lemma,

$$|E(T_{n-r,r})| + (r-1)(n-r) + \binom{r}{2} = |E(T_{n,r})|.$$

Hence

$$|E(G)| \le |E(T_{n,r})|,$$

which proves (5). Therefore $f(n,r) = |E(T_{n,r})|$. □

15 Directed Graphs

Definition 15.1 (Directed graph). A *directed graph* (digraph) is a pair $G = (V, E)$ where $V = V(G)$ is a vertex set and

$$E = E(G) \subseteq V \times V$$

is a set of *ordered pairs*. An edge $(x, y) \in E$ is written $x \longrightarrow y$. We say x is the *tail* and y is the *head* for the edge $x \longrightarrow y$.

Definition 15.2 (In-/out-degree). Let G be a digraph and let $x \in V(G)$. The *out-degree* is the number of outgoing edges from x.

$$d_G^+(x) = d^+(x) := \left|\{y \in V(G) : x \longrightarrow y \in E(G)\}\right|,$$

The indegree is the number of incoming edges to x.

$$d_G^-(x) = d^-(x) := \left|\{y \in V(G) : y \longrightarrow x \in E(G)\}\right|.$$

Definition 15.3 (In-/out-neighborhood). Let G be a digraph and $x \in V(G)$. The *out-neighborhood* of x is

$$N_G^+(x) = N^+(x) := \{y \in V(G) : x \longrightarrow y \in E(G)\},$$

and the *in-neighborhood* of x is

$$N_G^-(x) = N^-(x) := \{y \in V(G) : y \longrightarrow x \in E(G)\}.$$

Thus

$$d^+(x) = |N^+(x)|, \qquad d^-(x) = |N^-(x)|.$$

Definition 15.4 (Adjacency matrix of a digraph). If $V(G) = \{v_1, \ldots, v_n\}$, the *adjacency matrix* $A(G) = (a_{ij}) \in \{0, 1\}^{n \times n}$ is defined by

$$a_{ij} = \begin{cases} 1, & \text{if } v_i \longrightarrow v_j \in E(G), \\ 0, & \text{otherwise.} \end{cases}$$

Equivalently, $a_{ij} = 1$ iff there is an edge from row-vertex v_i to column-vertex v_j.

Definition 15.5 (Oriented incidence matrix). Let G be a digraph with $V(G) = \{v_1, \ldots, v_n\}$ and $E(G) = \{e_1, \ldots, e_m\}$. The *(oriented) incidence matrix* $M(G) = (m_{i\ell}) \in \{-1, 0, 1\}^{n \times m}$ is defined by

$$m_{i\ell} = \begin{cases} -1, & \text{if } v_i \text{ is the tail of } e_\ell, \\ 1, & \text{if } v_i \text{ is the head of } e_\ell, \\ 0, & \text{otherwise.} \end{cases}$$

So each column (one directed edge) has exactly one -1 at its tail and one $+1$ at its head.

Definition 15.6 (Directed path). A *directed path* in a digraph is a sequence of vertices

$$v_1 \to v_2 \to \cdots \to v_k$$

such that $v_i \to v_{i+1} \in E(G)$ for each $i = 1, \ldots, k-1$. The *length* of this path is $k-1$.

Definition 15.7 (Symmetric digraph). A digraph G is *symmetric* if whenever $x \to y$ is an edge, the reverse edge $y \to x$ is also an edge. Equivalently,

$$(x, y) \in E(G) \implies (y, x) \in E(G).$$

In matrix terms (with respect to any ordering of $V(G)$), this is equivalent to

$$A(G) = A(G)^{\mathsf{T}}.$$

Definition 15.8 (Antisymmetric digraph). A digraph G is *antisymmetric* if it has no two-way pairs of edges: for all distinct vertices $x \neq y$,

$$(x, y) \in E(G) \implies (y, x) \notin E(G).$$

Equivalently, between any unordered pair $\{x, y\}$ there is *at most one* directed edge.

Definition 15.9 (Underlying graph). Given a digraph $G = (V, E)$, its *underlying (undirected) graph* $U(G)$ is obtained by forgetting directions:

$$V\big(U(G)\big) = V(G),$$

and for distinct $x, y \in V(G)$,

$$\{x, y\} \in E\big(U(G)\big) \iff (x, y) \in E(G) \text{ or } (y, x) \in E(G).$$

Definition 15.10 (Orientation / oriented graph). Let H be a simple undirected graph. An *orientation* of H is a digraph G obtained by replacing each undirected edge $\{x, y\} \in E(H)$ by exactly one of the two directed edges $x \to y$ or $y \to x$. A digraph obtained this way is called an *oriented graph* or an *orientation* of H.

Remark 15.1. An oriented graph is antisymmetric and has no loops, and its underlying graph is exactly the original H.

15.1 Tournaments and Landau's Theorem

Motivation: Imagine a round-robin competition with n teams: every pair of teams plays exactly one game. We build a directed graph T to record the outcome:

- *Vertices* are the teams.
- For distinct teams x and y, we draw a directed edge $x \to y$ if x beats y.

Between any two teams, exactly one of them wins, so between any two vertices we get exactly one directed edge. This directed complete graph is what we call a tournament.

Definition 15.11 (Tournament)**.** A *tournament* is an orientation of a complete graph. Equivalently, a digraph T is a tournament if for every pair of distinct vertices $x \neq y$, *exactly one* of $x \to y$ or $y \to x$ is an edge.

Definition 15.12 (King in a tournament)**.** Let T be a tournament. A vertex v is a *king* if every other vertex can be reached from v by a directed path of length at most 2; i.e. for all $u \neq v$,

$$v \to u \quad \text{or} \quad \exists w \text{ with } v \to w \to u.$$

Remark 15.2 (Historical motivation for kings)**.** Landau introduced kings in the 1950s while modeling dominance in animal societies: vertices are animals, and an edge $x \to y$ means "x defeats (or pecks) y" in a round-robin dominance graph. A "perfect boss" would be a vertex that beats everyone directly (a source), but tournaments do not always have such a vertex. Landau's observation was that you can still guarantee a weaker leader: a vertex k such that for every other v either $k \to v$ or there is some u with $k \to u \to v$, i.e. k reaches everyone within two steps. This is exactly the definition of a *king*, and it matches the idea of an individual that may not dominate everyone personally, but dominates the whole group through its allies.

Theorem 15.1 (Landau, 1953)**.** Every tournament has a king. Moreover, every vertex of maximum out-degree in a tournament is a king.

Proof. Let x be a vertex of maximum out-degree in a tournament T. Write

$$V(T) = \{x\} \dot{\cup} N^+(x) \dot{\cup} N^-(x).$$

Assume for contradiction that x is *not* a king. Then there exists a vertex y that is not reachable from x by a directed path of length ≤ 2. In particular $x \to y \notin E(T)$, hence $y \to x \in E(T)$, i.e. $y \in N^-(x)$.

Now fix any $z \in N^+(x)$, i.e. $x \to z \in E(T)$. If we had $z \to y \in E(T)$, then $x \to z \to y$ would be a directed path of length 2, contradicting the choice of y. Therefore for every $z \in N^+(x)$ we must have $y \to z \in E(T)$, so $z \in N^+(y)$. Thus

$$N^+(x) \subseteq N^+(y),$$

and moreover $x \in N^+(y)$ (since $y \in N^-(x)$ means $y \to x \in E(T)$). Hence

$$d^+(y) \;\geq\; |N^+(x)| + 1 \;=\; d^+(x) + 1,$$

contradicting that x has maximum out-degree. Therefore x is a king, and the theorem follows. □

Definition 15.13 (Transmitter). A vertex t in a tournament is a *transmitter* if it has in-degree 0 (equivalently, t beats everyone: $d^+(t) = |V(T)| - 1$).

Lemma 15.2. Let T be a tournament with no transmitter (i.e. $d^-(v) \geq 1$ for all v). Then for every vertex $v \in V(T)$ there exists a king u such that $u \to v$.

Proof. Fix $v \in V(T)$. Since T has no transmitter, $N^-(v) \neq \varnothing$. Consider the subtournament $T[N^-(v)]$ induced by $N^-(v)$, and choose $u \in N^-(v)$ of maximum out-degree *within* $T[N^-(v)]$. By Landau's theorem applied to $T[N^-(v)]$, the vertex u is a king of $T[N^-(v)]$.

We claim u is a king of the whole tournament T. Let $w \in V(T)$.

- If $w \in N^-(v)$, then w is reachable from u within 2 steps inside $T[N^-(v)]$, hence also inside T.
- If $w \notin N^-(v)$, then $v \to w$ (because in a tournament exactly one of wv or vw holds). Since $u \in N^-(v)$ we have $u \to v$, and therefore $u \to v \to w$ is a directed path of length 2 from u to w.

Thus $\text{dist}_T(u, w) \leq 2$ for all w, so u is a king. Finally, $u \to v$ holds by $u \in N^-(v)$. □

Proposition 15.3 (At least three kings when there is no transmitter). If a tournament T has no transmitter, then T has at least 3 kings.

Proof. By Landau's theorem, T has at least one king.

Assume for contradiction that T has at most two kings.

Not exactly one king: If T had exactly one king k, apply the previous lemma to the vertex $v = k$. It yields a king u with $u \to k$. Since k is the *only* king, we must have $u = k$, impossible (no loops).

Not exactly two kings: If T had exactly two kings a, b, apply the lemma to $v = a$ to get a king u with $u \to a$. Then $u \in \{a, b\}$, but $u \neq a$ (again no loops), so $u = b$ and hence $b \to a$. Similarly, applying the lemma to $v = b$ forces $a \to b$. This contradicts that exactly one of ab or ba can be an arc in a tournament.

Therefore T cannot have 1 or 2 kings, and hence it has at least 3 kings. □

16 Connection and Decomposition

Definition 16.1 (u,v-path). A *path* in a graph G is a sequence of vertices

$$P = (v_0, v_1, \dots, v_k)$$

such that $v_i v_{i+1} \in E(G)$ for all $i = 0, 1, \dots, k-1$, and no vertex is repeated. Its *length* is k (the number of edges).

If $v_0 = u$ and $v_k = v$, we call P a (u,v)-*path*. The vertices u and v are the *endpoints*, and the vertices $v_1, \dots, v_{k-1}$ (if any) are the *internal vertices*.

Definition 16.2 (Connected graph). An undirected graph G is *connected* if for every pair of vertices $u, v \in V(G)$, there exists a (u,v)-path in G.

Definition 16.3 (Reachability relation). For an undirected graph G, define a relation $\sim$ on $V(G)$ by

$$u \sim v \quad \Longleftrightarrow \quad \text{there exists a } (u,v)\text{-path in } G.$$

The relation $\sim$ is an equivalence relation on $V(G)$.

Definition 16.4 (Connected component). A *connected component* of an undirected graph G is an equivalence class of $V(G)$ under $\sim$. Equivalently, a component is a *maximal connected subgraph* of G.

Remark 16.1. Thus G is connected iff it has exactly one connected component.

Strong connectivity in digraphs.

Definition 16.5 (Strongly connected digraph). A digraph D is *strongly connected* if for every pair $u, v \in V(D)$ there exists a directed path from u to v *and* a directed path from v to u.

Definition 16.6 (Strong component). Define a relation $\approx$ on $V(D)$ by

$$u \approx v \quad \Longleftrightarrow \quad u \text{ reaches } v \text{ and } v \text{ reaches } u \text{ by directed paths.}$$

The equivalence classes of $\approx$ are called the *strongly connected components* (or *strong components*) of D.

16.1 Walks and Paths

Definition 16.7 (Walk). A *walk* in a graph G is a sequence of vertices

$$W = (v_0, v_1, \dots, v_k)$$

such that $v_i v_{i+1} \in E(G)$ for all $i = 0, 1, \dots, k-1$. Unlike a path, vertices (and edges) are allowed to repeat. The *length* of the walk is k (the number of edges).

Definition 16.8 (Closed walk). A walk $(v_0, \dots, v_k)$ is *closed* if $v_0 = v_k$.

Definition 16.9 (Odd/even walk). A walk is *odd* (resp. *even*) if its length is odd (resp. even).

Lemma 16.1. Every (u, v)-walk contains a (u, v)-path.

Idea: Given any (u, v)-walk, whenever a vertex is visited twice, the segment between the two visits forms a cycle that can be deleted without breaking the connection between u and v. Repeating this process until no vertex is repeated yields a (u, v)-path.

Proof. We induct on the length ℓ of the (u, v)-walk $W = (v_0, v_1, \dots, v_\ell)$, where $v_0 = u$ and $v_\ell = v$.

Base case: $\ell = 0$. Then W has no edges, so $u = v$ and the walk repeats no vertices. Hence W itself is a (u, v)-path.

Induction step. Assume $\ell > 0$ and that every (u, v)-walk of length $< \ell$ contains a (u, v)-path. Let W be any (u, v)-walk of length ℓ.

If W has no repeated vertices, then it is already a path and we are done. Otherwise, some vertex appears at least twice along the walk. Choose indices $0 \le k_1 < k_2 \le \ell$ such that

$$v_{k_1} = v_{k_2} =: w.$$

Now delete the closed "detour" from the first occurrence of w to the next occurrence of w. Define the shorter vertex sequence

$$W' := (v_0, v_1, \dots, v_{k_1}, v_{k_2+1}, \dots, v_\ell).$$

This is still a (u, v)-walk: consecutive vertices in W' are consecutive in W except at the splice, where we go directly from $v_{k_1} = w$ to v_{k_2+1}, and $w v_{k_2+1}$ is an edge because it occurs in W.

Moreover the length strictly decreases:

$$\ell' = \ell - (k_2 - k_1) < \ell.$$

By the induction hypothesis, W' contains a (u, v)-path. Since W' is obtained from W by deleting vertices, any (u, v)-path contained in W' is also contained in W.

Therefore W contains a (u, v)-path. This completes the induction. □

Lemma 16.2. Let G be a graph and let $u, v, w \in V(G)$. If G contains a (u, v)-path and a (v, w)-path, then G contains a (u, w)-path. In particular, the relation

$$u \sim w \iff \text{there exists a } (u, w)\text{-path in } G$$

is transitive.

Proof. Let

$$P = (u = v_0, v_1, \dots, v_k = v) \quad \text{and} \quad Q = (v = w_0, w_1, \dots, w_\ell = w)$$

be a (u, v)-path and a (v, w)-path, respectively.

Concatenate them to get the (u, w)-*walk*

$$W = (u = v_0, v_1, \dots, v_k = v = w_0, w_1, \dots, w_\ell = w).$$

This is a walk because consecutive vertices along P and along Q are adjacent, and we glue at the common vertex v.

By the lemma, the walk W contains a (u, w)-path. Hence G contains a (u, w)-path, as claimed.

Therefore, if $u \sim v$ and $v \sim w$, then $u \sim w$, so $\sim$ is transitive. □

Lemma 16.3. Every odd closed walk contains an odd cycle.

Idea: Traverse the closed walk and whenever you first revisit a vertex, you've found a cycle; remove that cycle and continue. This decomposes the walk into edge-disjoint cycles whose lengths add up to the (odd) length of the walk. A sum of integers is odd only if at least one of them is odd, so one of these cycles must have odd length.

Proof. Let $W = (v_0, v_1, \dots, v_\ell)$ be a closed walk with $v_0 = v_\ell$ and odd length ℓ. Among all odd closed subwalks of W, choose one of minimum length and call it

$$W^* = (u_0, u_1, \dots, u_m), \qquad u_0 = u_m, \qquad m \text{ odd}.$$

We claim that W^* has no repeated vertices other than $u_0 = u_m$; hence it is a cycle, and since m is odd, it is an odd cycle.

Suppose for contradiction that some vertex repeats inside W^*. Then there exist indices $0 \le i < j < m$ with $u_i = u_j$. Consider the two closed walks obtained by splitting at this repetition:

$$C_1 = (u_i, u_{i+1}, \dots, u_j) \quad (\text{closed since } u_i = u_j),$$

and

$$C_2 = (u_j, u_{j+1}, \dots, u_m = u_0, u_1, \dots, u_i) \quad (\text{also closed}).$$

Their lengths are

$$|C_1| = j - i, \qquad |C_2| = m - (j - i).$$

Because m is odd, exactly one of $j - i$ and $m - (j - i)$ is odd. Moreover $0 < j - i < m$, so both $|C_1|$ and $|C_2|$ are strictly smaller than m. Therefore whichever of C_1 or C_2 has odd length is a *shorter* odd closed walk than W^*, contradicting the minimality of W^*.

Hence W^* has no repeated vertices except its start/end, so it is a cycle. Since its length m is odd, W^* is an odd cycle contained in W. □

16.2 Kőnig's Theorem characterizing bipartite graphs

It is obvious that an odd cycle prevents a graph from being bipartite: you cannot 2-color its vertices so that every edge goes between the two colors. Kőnig's theorem says this is the *only* obstruction: if a graph fails to be bipartite, it must already contain an odd cycle.

Theorem 16.4 (Kőnig)**.** A graph G is bipartite if and only if G contains no odd cycle.

Idea: Start from any vertex, color it red, its neighbors blue, their uncolored neighbors red, and so on. This forces all vertices at even distance to be one color and odd distance the other. An odd cycle would force some vertex to be both colors at once, which is impossible. Conversely, if no odd cycle exists, this coloring never breaks, so the graph is bipartite.

Proof. (⇒) Suppose G is bipartite with bipartition (X, Y), so every edge has one endpoint in X and the other in Y. Let $C = (v_0, v_1, \ldots, v_k = v_0)$ be any cycle in G. Starting at $v_0 \in X$ (wlog), each step across an edge forces us to alternate sides:

$$v_0 \in X \Rightarrow v_1 \in Y \Rightarrow v_2 \in X \Rightarrow \cdots$$

After k steps we return to v_0, which is in X. Thus k must be even (otherwise we would land in Y). So every cycle has even length, hence there is no odd cycle.

(⇐) Suppose G has no odd cycle. It suffices to show each connected component is bipartite. So assume G is connected and fix a root vertex r.

For any vertex v, let $\operatorname{dist}(r, v)$ be the length of a shortest (r, v)-path. Define

$$X := \{v \in V(G) : \operatorname{dist}(r, v) \text{ is even}\}, \qquad Y := \{v \in V(G) : \operatorname{dist}(r, v) \text{ is odd}\}.$$

Clearly $X \cup Y = V(G)$ and $X \cap Y = \varnothing$.

We claim there is no edge with both endpoints in X (and similarly none with both endpoints in Y). Assume for contradiction that $uv \in E(G)$ with $u, v \in X$. Let P_u and P_v be shortest paths from r to u and from r to v. Then $|P_u|$ and $|P_v|$ are both even.

Let z be the last common vertex of P_u and P_v (their paths from r coincide up to z and then diverge). Write $P_u = z \rightsquigarrow u$ and $P_v = z \rightsquigarrow v$ for the suffixes beyond z. The closed walk formed by

$$(z \rightsquigarrow u) \cup (u \to v) \cup (v \rightsquigarrow z)$$

has length

$$|z \rightsquigarrow u| + 1 + |z \rightsquigarrow v|.$$

Now $\operatorname{dist}(r, u) = \operatorname{dist}(r, z) + |z \rightsquigarrow u|$ and $\operatorname{dist}(r, v) = \operatorname{dist}(r, z) + |z \rightsquigarrow v|$. Since both $\operatorname{dist}(r, u)$ and $\operatorname{dist}(r, v)$ are even, $|z \rightsquigarrow u|$ and $|z \rightsquigarrow v|$ have the same parity, so $|z \rightsquigarrow u| + |z \rightsquigarrow v|$ is even, and therefore

$$|z \rightsquigarrow u| + 1 + |z \rightsquigarrow v|$$

is odd. Thus we have an odd closed walk, which contains an odd cycle, contradiction.

Hence no edge lies inside X or inside Y, so every edge goes between X and Y. Therefore G is bipartite.

Applying the same construction to each connected component finishes the proof for general G. □

16.3 Cut vertices and edges

Definition 16.10 (Cut-vertex, cut-edge). Let G be a graph and let $c(G)$ denote the number of connected components of G.

- A vertex $v \in V(G)$ is a *cut-vertex* if $c(G - v) > c(G)$.
- An edge $e \in E(G)$ is a *cut-edge* (or *bridge*) if $c(G - e) > c(G)$.

Proposition 16.5. Assume G is connected and let $e \in E(G)$. Then

$$G - e \text{ is connected} \iff e \text{ lies on a cycle of } G.$$

Equivalently, e is a cut-edge iff e is not contained in any cycle.

Proof. Write $e = xy$.
($\Rightarrow$) If $G - e$ is connected, then there is an (x, y)-path P in $G - e$. Adding the edge xy to P creates a cycle in G containing e.
($\Leftarrow$) If e lies on a cycle C, then $C - e$ contains an (x, y)-path P in $G - e$. Now take any vertices $u, v \in V(G)$. Since G is connected, there is a (u, v)-path in G. If that path does not use e we are done; if it uses $e = xy$, replace the subedge xy by the x–y path P in $G - e$. Thus u and v are still connected in $G - e$, so $G - e$ is connected. □

Remark 16.2. Adding an edge $e = xy$ to a graph merges two components iff x and y were in different components. In that case the new edge e is a bridge in the new graph, so it lies on no cycle.

Proposition 16.6 (Few edges ⇒ many components). Every n-vertex graph G with $e(G) \le k$ has at least $n - k$ connected components. Moreover this is best possible: for every $n > k$ there exists an n-vertex graph with k edges and exactly $n - k$ components.

Proof. Start from the empty graph E_n on n vertices, which has $c(E_n) = n$ components. Adding one edge can reduce the number of components by *at most* 1 (it either joins two components, or stays inside one). After adding $e(G) \le k$ edges, we therefore have

$$c(G) \ \ge\ n - k.$$

For sharpness, take a path on $k + 1$ vertices (which has k edges and 1 component) and add $n - (k + 1)$ isolated vertices. The resulting graph has k edges and $1 + (n - k - 1) = n - k$ components. □

Proposition 16.7 (At least two non-cut vertices). Every graph G with $|V(G)| \ge 2$ has at least two vertices that are not cut-vertices.

Proof. It suffices to prove this for a connected component of G, so assume G is connected. Let

$$P = (v_0, v_1, \ldots, v_\ell)$$

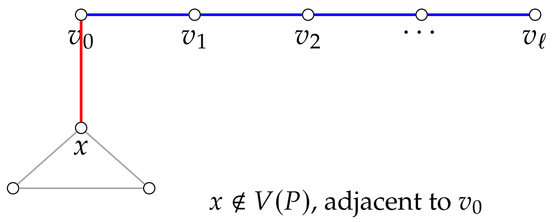
be a longest path in G (maximal length ℓ). We claim that v_0 and v_ℓ are not cut-vertices.

Suppose for contradiction that v_0 is a cut-vertex. Then $G - v_0$ has at least two components. All vertices $v_1, \ldots, v_\ell$ lie in a single component of $G - v_0$ (because they are connected by the subpath $v_1 - \cdots - v_\ell$). So there exists some vertex $x \notin \{v_1, \ldots, v_\ell\}$ adjacent to v_0 in a different component of $G - v_0$. Thus $v_0 x \in E(G)$ and $x \notin V(P)$, so

$$(x, v_0, v_1, \ldots, v_\ell)$$

is a path longer than P, contradiction. Hence v_0 is not a cut-vertex. The same argument applies to v_ℓ. □

Path $P = (v_0, \ldots, v_\ell)$ colored in blue
v_0 is a cut-vertex $\Rightarrow P$ is not a longest path



16.4 Eulerian circuits

Lemma 16.8. If $\delta(G) \geq 2$, then G contains a cycle.

Proof. Let $P = (v_0, v_1, \ldots, v_\ell)$ be a longest path in G. Because $\deg(v_0) \geq 2$, the vertex v_0 has a neighbor $x \neq v_1$. By maximality of P, this neighbor x must already lie on the path, say $x = v_i$ for some $i \geq 2$. Otherwise, $(x, v_0, v_1, \ldots, v_\ell)$ is a path: the edge xv_0 exists by choice of x, and no vertex repeats because x is new. This path has length $\ell + 1$, contradicting that P was chosen to be a longest path. Hence

$$v_0 v_1 \cdots v_i v_0$$

is a cycle. □

Definition 16.11 (Edge-decomposition). An *(edge-)decomposition* of a graph G is a partition of the edge set:

$$E(G) = E_1 \,\dot\cup\, E_2 \,\dot\cup\, \cdots \,\dot\cup\, E_t,$$

often with the intent that each $(V(G), E_i)$ has some nice structure (cycles, paths, stars, ...).

Definition 16.12 (Even graph / Eulerian graph). A graph G is *even* if every vertex has even degree.

Lemma 16.9. If G is even, then $E(G)$ can be partitioned into edge-disjoint cycles.

Proof. If $E(G) = \varnothing$ there is nothing to prove. Otherwise, consider the subgraph H spanned by the non-isolated vertices of G. Since G is even, every non-isolated vertex has degree at least 2, so $\delta(H) \geq 2$. By Lemma 16.8, H contains a cycle C.

Remove the edges of C to form $G_1 := G - E(C)$. Every vertex on C loses exactly 2 incident edges, so degrees remain even in G_1. Repeat the argument on G_1: if it still has edges, it contains a cycle, remove its edges, and so on. This process terminates because each step removes at least one edge. The removed cycles are pairwise edge-disjoint and their edges cover $E(G)$, giving the desired decomposition. □

Conjecture 16.10 (Hajós)**.** Every even (Eulerian) graph on n vertices can be decomposed into at most $\lfloor n/2 \rfloor$ cycles.

Conjecture 16.11 (Gallai)**.** Every n-vertex graph can be decomposed into at most $\left\lceil \frac{n}{2} \right\rceil$ paths.

Proposition 16.12 (A $K_{1,k}$-decomposition characterizes bipartite regular graphs)**.** Let G be a k-regular graph. Then G has an edge-decomposition into copies of $K_{1,k}$ (stars with k leaves) if and only if G is bipartite.

Proof. ($\Leftarrow$) If G is bipartite with bipartition (A, B), then for each $a \in A$ the set of all k edges incident to a forms a star $K_{1,k}$ centered at a. Because every edge has exactly one endpoint in A, these stars are edge-disjoint and their union is $E(G)$. So they give a $K_{1,k}$-decomposition.

($\Rightarrow$) Suppose $E(G)$ is partitioned into stars $S_1, \ldots, S_t$, each isomorphic to $K_{1,k}$. Let A be the set of star-centers (the unique degree-k vertex in each star), and let $B := V(G) \setminus A$. We claim (A, B) is a bipartition.

First, no vertex can be both a center and a leaf: if v is a leaf in some star, then the incident edge used there is already assigned to that star; but if v were also a center, *all* k edges incident to v would have to lie in the star centered at v, contradicting that at least one of those edges was assigned elsewhere. Hence $A \cap B = \varnothing$.

Now take any edge $uv \in E(G)$. It lies in exactly one star, and in that star exactly one endpoint is the center. Thus exactly one of $\{u, v\}$ lies in A, and the other lies in B. Therefore every edge goes between A and B, so G is bipartite. □

Definition 16.13 (Multigraph, simple graph)**.** A *multigraph* is a graph in which edges may have multiplicity (i.e. multiple edges between the same pair of vertices are allowed), and loops are also allowed. A *simple graph* is a graph with no loops and no multiple edges.

Definition 16.14 (Trail, circuit, Eulerian circuit)**.** Let G be a (multi)graph.

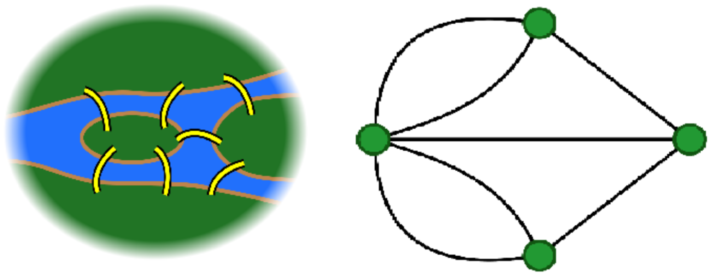
- A *trail* is a walk that uses each edge at most once.
- A *circuit* is a closed trail.
- An *Eulerian circuit* is a circuit that uses every edge of G exactly once.

Seven Bridges of Königsberg: In 1736, the city of Königsberg (now Kaliningrad) presented a well-known walking puzzle. The Pregel River split the city into four landmasses connected by seven bridges. The question was whether one can take a walk that crosses each bridge *exactly once* and returns to the starting point.

Abstract the situation by forming a multigraph G whose vertices are the landmasses and whose edges are the bridges (allowing multiple edges if two landmasses are connected by several bridges). The question is then:

Does G contain a closed walk that traverses each edge exactly once?

That is exactly what we now call an *Eulerian circuit*.



The Seven Bridges of Königsberg: can one cross each bridge exactly once and return to the start?

Suppose you are walking an Eulerian circuit. Every time you enter a vertex along some edge, you must also leave along a *different* unused edge (except that you start and end at the same vertex, which still matches up in pairs because you return). So the incident edges at each vertex get used in *enter/exit pairs*. That forces the degree of every vertex to be *even*.

A single walk cannot jump between disconnected pieces of the graph. So if two different components both contain edges, there is no way one closed walk can cover them all. Equivalently, among the components of G, at most one can be *nontrivial* (contain an edge).

In the Königsberg graph, vertices have odd degree, so an Eulerian circuit cannot exist. Euler's theorem below states that these two obstructions are not merely necessary but also sufficient: if G has at most one nontrivial component and every vertex has even degree, then G does have an Eulerian circuit.

Theorem 16.13 (Euler's Theorem (for multigraphs))**.** A multigraph G has an Eulerian circuit if and only if

1. G has at most one nontrivial component (i.e. among its connected components, at most one contains an edge), and
2. every vertex has even degree.

Proof. (⇒) If G has an Eulerian circuit, then every vertex has even degree: each time the circuit enters a vertex along some edge, it must leave along a distinct unused edge, so incident edges are paired. Also, an Eulerian circuit lives inside a single connected component containing edges, so there can be at most one nontrivial component.

(⇐) Assume G has at most one nontrivial component and every vertex has even degree. Discard isolated vertices; we may assume G is connected and has at least one edge.

Because all degrees are even, G is an even graph. Hence $E(G)$ decomposes into edge-disjoint cycles

$$E(G) = E(C_1) \dot{\cup} \cdots \dot{\cup} E(C_t).$$

If two circuits share a vertex, they can be spliced into one bigger circuit that uses exactly the union of their edges: start walking along the first circuit; upon first reaching a shared vertex, traverse the entire second circuit and return to the same vertex; then continue along the first circuit. The result is still a closed trail (no edge is repeated) and it uses all edges of both circuits.

As long as $t \geq 2$, we claim there exist $i \neq j$ with C_i and C_j sharing a vertex. Indeed, let H be the subgraph formed by the union of the cycles. Then H is connected (because H contains all edges of G, and G is connected). If all cycles were vertex-disjoint, then H would be a disjoint union of those cycles and hence disconnected, contradiction. So some two cycles intersect, and we can merge them.

Repeatedly merge intersecting circuits. Each merge reduces the number of circuits by 1, and the process must stop. When it stops, we have a single circuit using all edges of G, i.e. an Eulerian circuit. □

Remark 16.3 (Why Königsberg is often called the "birth" of graph theory)**.** The Königsberg bridges puzzle is remembered as the birth of graph theory because Euler solved it by *discarding almost all geometric information*. He did not use distances, angles, or coordinates; he kept only which landmasses are connected by which bridges, i.e. the adjacency structure of a graph. That shift created a new kind of mathematics: studying properties that depend only on connectivity (and are invariant under any redrawings of the picture). Euler's parity argument (odd vs. even degrees) is a first example of a purely graph-theoretic argument, and the resulting theorem is not about Königsberg in particular but about a general class of graphs.

17 Trees

17.1 Basic properties of trees

Definition 17.1. A graph is *acyclic* if it contains no cycle. An acyclic graph is also called a *forest*. A *tree* is a connected forest (equivalently: a connected acyclic graph).

Definition 17.2 (Spanning tree). Let G be a graph. A subgraph $H \subseteq G$ is a *spanning tree* of G if

$$H \text{ is a tree} \qquad \text{and} \qquad V(H) = V(G).$$

Definition 17.3 (Leaf). A *leaf* of a forest is a vertex of degree 1 (in that forest).

Proposition 17.1. Let T be a tree.

(i) If $|V(T)| \geq 2$, then T has at least two leaves.

(ii) If v is a leaf of T, then $T - v$ is a tree.

(iii) If $|V(T)| = n$, then $|E(T)| = n - 1$.

Proof. *(i)* Let P be a longest path in T, with endpoints a, b. We claim a is a leaf. If $\deg_T(a) \geq 2$, then a has a neighbor $x \neq$ the next vertex of P. Because T is acyclic, x cannot lie on P (otherwise we would create a cycle by going from a into P and back to a via x), so we could extend P to a longer path starting $x - a - \cdots$, contradicting maximality. Thus $\deg_T(a) = 1$. Similarly $\deg_T(b) = 1$. Hence T has at least two leaves.

(ii) Let v be a leaf of T, and let u be its unique neighbor. Set $T' := T - v$.

T' is connected: Take any two vertices $x, y \in V(T')$. Since T is connected, there is an x–y path P in T. If P contains v, then P must use the edge uv to enter v; but to reach $y \neq v$ it would have to leave v again, and the only way out is along uv once more, repeating an edge. This is impossible for a path, so P avoids v and is therefore contained in T'. Hence T' is connected.

T' is acyclic: If T' had a cycle, that same cycle must have been in T (we only deleted v and its incident edge, and deleting vertices and edges does not introduce a cycle). But T is acyclic, a contradiction.

Since T' is connected and acyclic, so T' is a tree.

(iii) We induct on $n = |V(T)|$. If $n = 1$, then T has 0 edges and the statement $|E(T)| = n - 1$ holds. Assume $n \geq 2$ and the statement holds for all smaller trees. By (i), T has a leaf v. Removing v produces a tree $T' := T - v$ with $n - 1$ vertices by (ii). By induction,

$$|E(T')| = (n-1) - 1 = n - 2.$$

Since v was a leaf, deleting v removed exactly one edge, so $|E(T)| = |E(T')| + 1 = (n-2) + 1 = n - 1$. □

17.2 Characterization of trees

Proposition 17.2 (Characterizations of trees)**.** Let $n \geq 1$ and let G be a graph on n vertices. The following are equivalent:

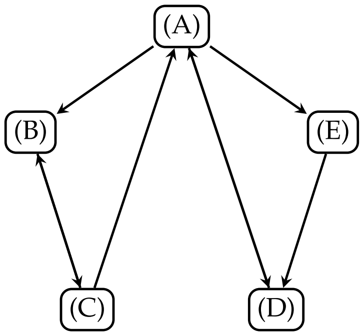
(A) G is a tree (connected and has no cycles).

(B) G is connected and has $n-1$ edges.

(C) G has no cycles and has $n-1$ edges.

(D) For any $u, v \in V(G)$, there is exactly one u–v path in G.

(E) Adding any edge $e \notin E(G)$ creates a graph with exactly one cycle.



Proof. $(A) \Rightarrow (D)$**.** Fix $u \neq v$. Since G is connected, there exists a u–v path. For uniqueness, suppose P, Q are two distinct u–v paths. Let x be the first vertex where they diverge and let y be the next vertex where they meet again. Then the subpath of P from x to y together with the subpath of Q from x to y forms a cycle, contradicting that G is acyclic. Hence the u–v path is unique.

$(D) \Rightarrow (A)$**.** Condition (D) implies G is connected (there is a path between every pair). Now if G had a cycle, pick two distinct vertices u, v on that cycle; going around the cycle in the two directions gives two different u–v paths, contradicting (D). Thus G has no cycles.

$(A) \Rightarrow (B)$**.** Follows from (iii) of previous proposition

$(B) \Rightarrow (C)$**.** Assume G is connected and $|E(G)| = n-1$. If G contains a cycle, delete an edge from that cycle. This does not disconnect the graph, since the remaining edges of the cycle still give a route between the endpoints. Repeat this until no cycles remain, obtaining a graph G' that is connected and acyclic.

Thus G' satisfies (A), and since $(A) \Rightarrow (B)$ we have

$$|E(G')| = n-1.$$

But each deletion reduces the number of edges by 1, so starting from $|E(G)| = n-1$ we can only end with $|E(G')| = n-1$ if we deleted *zero* edges. Therefore no cycle edge was ever available to delete, i.e. G had no cycles to begin with. Hence G is acyclic, which is (C).

$(C) \Rightarrow (B)$**.** Assume G has no cycles and $|E(G)| = n-1$. Let $G_1, \ldots, G_k$ be the connected components of G, and write $n_i := |V(G_i)|$ and $e_i := |E(G_i)|$. Each G_i is connected and acyclic, hence satisfies (A), so by $(A) \Rightarrow (B)$ applied to G_i we have $e_i = n_i - 1$ for all i. Summing over components gives

$$n-1 = |E(G)| = \sum_{i=1}^{k} e_i = \sum_{i=1}^{k} (n_i - 1) = \Big(\sum_{i=1}^{k} n_i \Big) - k = n - k,$$

so $k = 1$. Hence G is connected, i.e. (B) holds.

$(C) \Rightarrow (A)$. By $(C) \Rightarrow (B)$ the graph G is connected, and (C) also says G is acyclic; hence G is a tree.

$(A) \Rightarrow (E)$. Assume G is a tree, and let $e = uv \notin E(G)$. By $(A) \Rightarrow (D)$, there is a unique u–v path P in G. In $G + e$, the subgraph $P \cup \{e\}$ is a cycle. Moreover, since G had no cycles, any cycle in $G + e$ must use the new edge e, and then the rest of that cycle is a u–v path in G, which must be P by uniqueness. Thus $G + e$ has exactly one cycle.

$(E) \Rightarrow (D)$. Assume (E). Fix distinct $u, v \in V(G)$ and add the edge $e = uv$. By (E), the graph $G + e$ contains exactly one cycle, and this cycle must use e (otherwise it would already be a cycle in G). Removing e from that cycle leaves a u–v path in G, so at least one such path exists.

For uniqueness, if G had two distinct u–v paths $P \neq Q$, then $P \cup \{e\}$ and $Q \cup \{e\}$ would be two distinct cycles in $G + e$, contradicting that $G + e$ has exactly one cycle. Hence G has exactly one u–v path.

□

Corollary 17.3. Let T be a tree.

(i) Every edge of T is a cut-edge (bridge).

(iii) Every connected graph has a spanning tree.

Proof. (i) Fix $e = uv \in E(T)$. In a tree there is a unique (u, v)-path, namely the single edge uv. Deleting e destroys the only (u, v)-path, so $T - e$ is disconnected. Hence e is a cut-edge.

(iii) Let G be connected. If G is already acyclic, it is a tree and we are done. Otherwise, repeatedly delete an edge that lies on a cycle. This keeps the graph connected and strictly decreases the number of edges. The process must stop, and it stops exactly when no cycle remains, i.e. at a tree. Because we only deleted edges and never removed vertices, the resulting tree spans $V(G)$. □

Theorem 17.4 (Spanning tree exchange)**.** Let G be a connected graph and let T, T' be spanning trees of G. If $e \in E(T) \setminus E(T')$, then there exist edges $e', e'' \in E(T') \setminus E(T)$ such that

$$T_1 := T - e + e' \quad \text{and} \quad T_2 := T' + e - e''$$

are both spanning trees of G. (In fact one can take $e'' = e'$, so a single edge swap works in both directions.)

Proof. Let $e = uv \in E(T) \setminus E(T')$. Since T is a tree, e is a cut-edge of T. Let the two components of $T - e$ have vertex sets U and U' with $u \in U$ and $v \in U'$.

Because T' is connected, there is a (u, v)-path in T'. That path starts in U and ends in U', so at some point it must cross from U to U'. Therefore there exists an edge

$$e' = xy \in E(T') \quad \text{with} \quad x \in U, \; y \in U'.$$

Necessarily $e' \notin E(T)$, because T has no edges between U and U' after removing e. So $e' \in E(T') \setminus E(T)$.

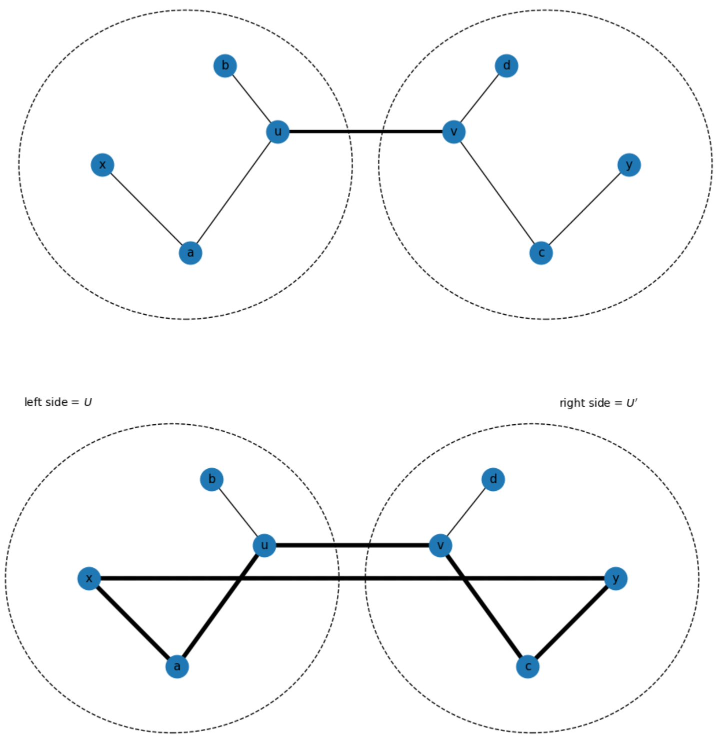
Now add e' to $T - e$. It reconnects the two components, and since $T - e$ is acyclic and we added exactly one edge between components, $T_1 := T - e + e'$ is connected and acyclic, hence a spanning tree.

Next consider $T' + e$. Since T' is a tree, adding e creates a unique cycle C. This cycle must cross the cut (U, U') at least once (it uses e itself), so it contains some edge e'' of T' that crosses between U and U'. Choose such an edge e'' on C with $e'' \neq e$. Then $e'' \in E(T') \setminus E(T)$ (same reason as above), and removing e'' breaks the unique cycle while keeping the graph connected. Thus

$$T_2 := T' + e - e''$$

is a spanning tree.

Finally, note we may simply take $e'' = e'$, because e' lies on the (u, v)-path in T', hence lies on the unique cycle in $T' + e$. □



Top: ***Tree*** *T. Removing $e = uv$ splits the vertices into U and U'*
Bottom: ***Tree*** *$T' + uv$. In T', the u–v path crosses via $e' = xy$; in $T' + e$ this creates one cycle*

17.3 Distance in graphs

Definition 17.4 (Distance). For vertices $x, y \in V(G)$, define the *distance*

$$d_G(x, y) := \text{length of a shortest } (x, y)\text{-path}.$$

If x and y lie in different components, set $d_G(x, y) = \infty$.

Remark 17.1 (Basic properties). For all $x, y, z \in V(G)$:

- $d(x, x) = 0$.
- $d(x, y) = d(y, x)$ (undirected graphs).
- (Triangle inequality) $d(x, z) \le d(x, y) + d(y, z)$.

On a connected graph (so distances are finite), d is a metric on $V(G)$.

Definition 17.5 (Eccentricity, diameter, radius). Assume G is connected.

- The *eccentricity* of a vertex u is
$$\varepsilon(u) := \max_{v \in V(G)} d(u, v).$$
- The *diameter* of G is
$$\operatorname{diam}(G) := \max_{u,v \in V(G)} d(u, v) = \max_{u \in V(G)} \varepsilon(u).$$
- The *radius* of G is
$$\operatorname{rad}(G) := \min_{u \in V(G)} \varepsilon(u).$$

Definition 17.6 (Center of a graph). Assume G is connected. The *center* of G is the set of vertices with minimum eccentricity:
$$\operatorname{Center}(G) := \{u \in V(G) : \varepsilon(u) = \operatorname{rad}(G)\}.$$
(So $\operatorname{Center}(G)$ is the set of vertices that are "as close as possible" to everyone else.)

Theorem 17.5 (Jordan, 1869). If T is a tree, then $\operatorname{Center}(T)$ is either

- a single vertex, or
- two adjacent vertices.

Proof. If $|V(T)| \le 2$ the statement is immediate, so assume $|V(T)| \ge 3$.

Let L be the set of leaves of T, and let
$$T' := T - L$$
be the graph obtained by deleting *all* leaves (simultaneously). By Proposition 17.1(ii), deleting a leaf from a tree yields a tree; repeating over all leaves shows T' is either empty or still a tree.

Key claim: For every vertex $x \in V(T')$,
$$\varepsilon_T(x) = \varepsilon_{T'}(x) + 1.$$

Why? Take a farthest vertex y from x in T. In a tree, any farthest vertex must be a leaf: if y were not a leaf, it has a neighbor further away from x along the unique x–y path, contradicting maximality. Thus a farthest vertex is in L, so when we delete all leaves, every farthest vertex disappears and the maximum distance drops by 1. Conversely, any farthest vertex in T' is at distance one less than a farthest leaf in T (just extend the path one step to a leaf), so the drop is exactly 1.

In particular, *all* eccentricities in the surviving tree drop by exactly 1, so the set of vertices minimizing eccentricity does not change when passing from T to T':

$$\text{Center}(T) = \text{Center}(T').$$

Now iterate the leaf-pruning process:

$$T = T_0 \supset T_1 \supset T_2 \supset \cdots, \qquad T_{i+1} := T_i - \{\text{leaves of } T_i\}.$$

Each step removes at least two vertices as long as $|V(T_i)| \geq 3$ (trees have ≥ 2 leaves), so the process must stop, and it stops exactly when T_m has either 1 vertex or 2 vertices. If T_m has 2 vertices, they must be adjacent (otherwise it would not be connected).

Since centers are preserved at every pruning step, we get

$$\text{Center}(T) = \text{Center}(T_m),$$

and $\text{Center}(T_m)$ is either a single vertex or two adjacent vertices. □

18 Matchings in bipartite graphs

Definition 18.1 (Matching). A *matching* in a graph G is a set of edges no two of which share an endpoint. Equivalently, each vertex is incident to at most one edge of the matching.

A matching *saturates* a vertex set $S \subseteq V(G)$ if every vertex in S is incident to some edge of the matching.

Motivation: SDR / jobs and applicants. Suppose there are jobs labeled $1, 2, \ldots, n$, and for each job i we are given a set A_i of applicants who are qualified for that job. We ask:

Can we assign a distinct qualified applicant to every job?

Model this question as a bipartite graph $G = (X, Y)$:

- $X = \{1, 2, \ldots, n\}$ represents the jobs.
- $Y = \bigcup_{i=1}^{n} A_i$ represents all applicants who appear in at least one list.
- Put an edge i–a if applicant a is qualified for job i (i.e., $a \in A_i$).

Then for each job $i \in X$, its neighborhood is exactly

$$N(i) = A_i.$$

An X-*saturating matching* is a matching that touches every job vertex $i \in X$ exactly once.

- "touches i" means we pick some edge i–a, i.e. we assign job i to applicant $a_i \in A_i$;
- "matching edges are disjoint" means no applicant vertex $a \in Y$ is used twice for two jobs.

Thus an X-saturating matching is the same thing as a *set of distinct representatives* (SDR) for $\{A_i\}$:

$$\text{choose } a_i \in A_i \text{ for each } i, \text{ with all } a_i \text{ distinct.}$$

Necessary condition: Fix a subset of jobs $J \subseteq X$. The only applicants who could possibly fill jobs in J are those adjacent to at least one job in J, namely the pool

$$N(J) = \bigcup_{i \in J} A_i.$$

Any X-saturating matching assigns *distinct* applicants to the jobs in J. So it would have to choose $|J|$ different vertices from the set $N(J)$. That is an injection

$$J \hookrightarrow N(J),$$

which is impossible if $|N(J)| < |J|$. Equivalently: if a group of $|J|$ jobs has access to fewer than $|J|$ applicants in total, then some job in that group must be left unmatched no matter what you do.

Therefore a necessary condition for an X-saturating matching is

$$\forall J \subseteq X : \qquad |N(J)| \geq |J|.$$

No subset of jobs is competing for too small a common applicant pool

It turns out that there are *no other obstructions*. In 1935, Philip Hall proved that this same inequality condition already guarantees an SDR exists. In graph language:

18.1 Hall's Marriage Theorem

Theorem 18.1 (Hall's Marriage Theorem)**.** In an (X, Y)-bigraph G, there exists an X-saturating matching iff for every $S \subseteq X$,

$$|S| \le |N(S)|.$$

(This is *Hall's condition*.)

Proof. Necessity of Hall's Condition was shown previously. To prove that Hall's Condition is sufficient, we induct on the size of X.

Base cases. $|X| = 0$ is trivial. If $|X| = 1$, Hall's condition implies the unique vertex of X has a neighbor, giving a matching of size 1.

Induction step. Let $|X| \ge 2$ and assume the theorem holds for all smaller X.

Case 1: Every nonempty proper subset $S \subsetneq X$ satisfies $|N(S)| > |S|$.

Every proper set of jobs has at least one *extra* available applicant:

$$|N(S)| \ge |S| + 1.$$

So we can safely commit to one job–applicant pair without creating a shortage for the rest.

Pick any job $x \in X$ and any neighbor $y \in N(x)$, and delete them:

$$X' := X \setminus \{x\}, \qquad Y' := Y \setminus \{y\},$$

letting G' be the induced bipartite graph on (X', Y').

Hall still holds in G': For any $S \subseteq X'$, we have $S \subsetneq X$, hence $|N_G(S)| \ge |S| + 1$. Removing y can delete at most one neighbor, so

$$|N_{G'}(S)| \ge |N_G(S)| - 1 \ge |S|.$$

By induction, G' has an X'-saturating matching M'. Since x and y are gone from G', adding xy causes no conflicts, and $M := M' \cup \{xy\}$ saturates all of X.

Case 2: There exists a nonempty proper $S \subsetneq X$ with $|S| = |N(S)|$.

The jobs in S have exactly $|S|$ available applicants total, so in any full assignment they must use *all* of $N(S)$. Hence jobs outside S cannot rely on applicants in $N(S)$, so we solve S and the remainder separately.

Let

$$G_1 := G[S \cup N(S)] \qquad \text{and} \qquad G_2 := G[(X \setminus S) \cup (Y \setminus N(S))].$$

Step 1: Match S inside G_1. For any $R \subseteq S$, Hall in G gives $|R| \le |N_G(R)|$, and $N_G(R) \subseteq N(S)$, so $N_{G_1}(R) = N_G(R)$ and Hall holds on the left side S in G_1. By induction (since $|S| < |X|$), G_1 has an S-saturating matching M_1.

Step 2: Match $X \setminus S$ inside G_2. Take $T \subseteq X \setminus S$. In G_2 we removed the applicants $N(S)$, so

$$N_{G_2}(T) = N_G(T) \setminus N(S).$$

Also $N_G(T \cup S) = N_G(T) \cup N(S)$, hence

$$|N_{G_2}(T)| = |N_G(T \cup S)| - |N(S)|.$$

By Hall in G, $|T \cup S| \le |N_G(T \cup S)|$, and using $|N(S)| = |S|$ we get

$$|N_{G_2}(T)| \ge |T \cup S| - |S| = |T|.$$

So Hall holds on $X \setminus S$ in G_2. By induction, G_2 has an $(X \setminus S)$-saturating matching M_2.

Step 3: Glue. M_1 uses only vertices in $S \cup N(S)$, while M_2 uses only vertices in $(X \setminus S) \cup (Y \setminus N(S))$. These sets are disjoint, so

$$M := M_1 \cup M_2$$

is a matching that saturates all of X.

□

18.2 Hakimi's Theorem on orientations with given outdegrees

Theorem 18.2 (Hakimi (1965)). Let $e(G) = m$ and let d_i be nonnegative integers assigned to vertices v_i. Then G has an orientation with outdegree d_i at each v_i iff

$$\sum_i d_i = m \quad \text{and} \quad \forall\, U \subseteq V(G) : \sum_{v_i \in U} d_i \ge e(G[U]).$$

Proof. *(⇒)* In any orientation, each edge contributes 1 to exactly one outdegree, hence $\sum_i d_i = m$. If some U violated $\sum_{v_i \in U} d_i \ge e(G[U])$, then applying the handshake lemma in the induced subgraph $G[U]$ would contradict that all its edges must be oriented out of some endpoint in U.

(⇐) Build an auxiliary bipartite graph $H = (X, Y)$:

$$X = \{x_e : e \in E(G)\}, \qquad Y = \bigcup_i \{d_i \text{ copies of } v_i\}.$$

If $e = v_i v_j$, connect x_e to every copy of v_i and every copy of v_j.

A perfect matching in H chooses for each edge e exactly one of its endpoints; orient e *out of* that chosen endpoint. Then the outdegree of v_i equals the number of its copies matched, namely d_i.

So it suffices to prove Hall for H with respect to X. Let $S \subseteq X$ be an edge set; let F be the set of vertices incident to edges in S. Then $N(S)$ consists of all copies of vertices in F, and thus

$$|N(S)| = \sum_{v_i \in F} d_i \;\ge\; e(G[F]) \;\ge\; |S|.$$

Hence Hall holds and H has a perfect matching, giving the desired orientation. □

Corollary 18.3. 1. Every k-regular bipartite multigraph has a perfect matching.

2. Every k-regular bipartite multigraph can be decomposed into k edge-disjoint perfect matchings.

Proof. (i) Let $G = (X, Y)$ be k-regular. For any $S \subseteq X$, count edges between S and $N(S)$:

$$|S| \cdot k = e(S, N(S)) \le |N(S)| \cdot k,$$

so $|S| \le |N(S)|$. Hall gives a perfect matching.

(ii) Remove one perfect matching; the remaining graph is $(k-1)$-regular. Apply (i) repeatedly. □

18.3 Birkhoff–von Neumann Theorem

Definition 18.2 (Doubly stochastic matrix). An $n \times n$ matrix M is *doubly stochastic* if all entries are nonnegative and every row sum and every column sum equals 1.

Definition 18.3 (Convex combination). A *convex combination* of matrices is a linear combination with nonnegative coefficients that sum to 1.

Theorem 18.4 (Birkhoff (1946), von Neumann (1953)). Every doubly stochastic matrix is a convex combination of permutation matrices.

Proof. Induct on the number of nonzero entries of M.

Base case ($m = n$). If M has exactly n positive entries, then each of the n rows has row-sum 1 so each row contains at least one positive entry, hence exactly one; likewise each column contains exactly one positive entry. Thus every row and column contains a single 1, so M is a permutation matrix.

Induction step. Assume $m > n$ and the claim holds for all doubly stochastic matrices with fewer than m positive entries.

Step 1: Build the support graph and verify Hall. Let $G = (X, Y; E)$ be the bipartite graph with

$$X = \{1, \dots, n\} \text{ (columns)}, \qquad Y = \{1, \dots, n\} \text{ (rows)},$$

and

$$(j, i) \in E \quad \Longleftrightarrow \quad M_{ij} > 0.$$

(So a row-vertex $i \in Y$ is adjacent exactly to those columns $j \in X$ where M_{ij} is positive.)

We claim G satisfies Hall's condition on the left side Y: for every $S \subseteq Y$,

$$|N(S)| \ge |S|.$$

Indeed, if $N(S) \subseteq X$ is the set of columns that have a positive entry in some row of S, then rows in S have *no* positive entries outside $N(S)$, hence

$$\sum_{i \in S} \sum_{j \in N(S)} M_{ij} = \sum_{i \in S} \sum_{j=1}^{n} M_{ij} = \sum_{i \in S} 1 = |S|.$$

On the other hand, for each fixed column j, the column sum is 1, so

$$\sum_{i \in S} M_{ij} \le \sum_{i=1}^{n} M_{ij} = 1.$$

Summing this over $j \in N(S)$ gives

$$\sum_{i \in S} \sum_{j \in N(S)} M_{ij} \;\le\; \sum_{j \in N(S)} 1 \;=\; |N(S)|.$$

Combining with the previous equality yields $|S| \le |N(S)|$, proving Hall.

Therefore G has a perfect matching P.

Step 2: Subtract a scaled permutation matrix and keep nonnegativity. Let π be the permutation corresponding to the matching P, i.e. $(\pi(i), i) \in E$ for all rows i, and let P also denote the associated permutation matrix:

$$P_{ij} = \begin{cases} 1 & \text{if } j = \pi(i), \\ 0 & \text{otherwise.} \end{cases}$$

Since $(\pi(i), i) \in E$, we have $M_{i,\pi(i)} > 0$ for all i. Define

$$\varepsilon := \min_{1 \le i \le n} M_{i,\pi(i)} \;>\; 0.$$

Now set

$$M' := M - \varepsilon P.$$

Then $M' \ge 0$ entrywise, because we subtract ε only from the matched entries $M_{i,\pi(i)}$, and ε was chosen to be at most each of them.

Also, in every row and every column, exactly one entry of P equals 1, hence

$$\sum_{j=1}^{n} M'_{ij} = \sum_{j=1}^{n} M_{ij} - \varepsilon \sum_{j=1}^{n} P_{ij} = 1 - \varepsilon, \qquad \sum_{i=1}^{n} M'_{ij} = \sum_{i=1}^{n} M_{ij} - \varepsilon \sum_{i=1}^{n} P_{ij} = 1 - \varepsilon.$$

Finally, M' has strictly fewer positive entries than M: at least one matched entry achieves the minimum ε, so for some i_0 we have $M'_{i_0,\pi(i_0)} = M_{i_0,\pi(i_0)} - \varepsilon = 0$, and no previously-zero entry becomes positive.

Step 3: Renormalize and apply induction. Define

$$M'' := \frac{1}{1-\varepsilon} M'.$$

Then M'' is doubly stochastic and has fewer than m positive entries. By the induction hypothesis, M'' is a convex combination of permutation matrices:

$$M'' = \sum_{k=1}^{r} \lambda_k Q^{(k)}, \qquad \lambda_k \ge 0, \;\; \sum_{k=1}^{r} \lambda_k = 1,$$

where each $Q^{(k)}$ is a permutation matrix.

Multiply by $1 - \varepsilon$ and substitute $M' = M - \varepsilon P$:

$$M - \varepsilon P = (1 - \varepsilon) \sum_{k=1}^{r} \lambda_k Q^{(k)}.$$

Hence

$$M = \varepsilon P + \sum_{k=1}^{r} \big((1-\varepsilon)\lambda_k\big) Q^{(k)}.$$

The coefficients are nonnegative and sum to

$$\varepsilon + \sum_{k=1}^{r} (1-\varepsilon)\lambda_k = \varepsilon + (1-\varepsilon)\cdot 1 = 1,$$

so this is a convex combination of permutation matrices. This completes the induction. □

18.4 Defect formula in bipartite graphs

Definition 18.4 (Matching number). Let $\alpha'(G)$ denote the maximum size of a matching in G.

Assume G is bipartite with bipartition (X, Y).

Definition 18.5 (Defect). For $S \subseteq X$, the *defect* of S is

$$d(S) := |S| - |N(S)|.$$

(So $d(S) = 0$ is exactly Hall's equality case.)

Theorem 18.5 (Defect Formula). For a bipartite graph $G = (X, Y)$,

$$\alpha'(G) = \min_{S \subseteq X} \Big(|X| - d(S)\Big) = |X| - \max_{S \subseteq X} d(S).$$

Proof. Let $S \subseteq X$ achieve the minimum in the formula. Form G' by adding $d(S)$ new vertices to Y, each adjacent to every vertex of X.

We claim G' satisfies Hall's condition. Let $S' \subseteq X$. Then in G'

$$|N_{G'}(S')| = |N_G(S')| + d(S) \geq |S'| - d(S') + d(S) \geq |S'|,$$

since $d(S) \geq d(S')$ by maximality of S among defects.

Hence G' has an X-saturating matching. Restricting to G yields a matching covering $|X| - d(S)$ vertices of X, so $\alpha'(G) \geq |X| - d(S)$.

Conversely, any matching in G covers at most $|N(S)|$ vertices of S, so leaves at least $d(S)$ vertices of X unmatched. Thus $\alpha'(G) \leq |X| - d(S)$. □

18.5 Vertex covers and König–Egerváry

Definition 18.6 (Vertex cover). A set $S \subseteq V(G)$ is a *vertex cover* if every edge has at least one endpoint in S. Let $\beta(G)$ be the minimum size of a vertex cover.

Lemma 18.6 (Vertex covers vs. independent sets)**.** A set $S \subseteq V(G)$ is a vertex cover of G if and only if $V(G) \setminus S$ is an independent set.

Proof. ($\Rightarrow$) Assume S is a vertex cover and let $I := V(G) \setminus S$. If I were not independent, there would be an edge uv with $u, v \in I$. But then $u, v \notin S$, so the edge uv has no endpoint in S, contradicting that S is a vertex cover. Hence I is independent.

($\Leftarrow$) Assume $I := V(G) \setminus S$ is independent. Let uv be any edge of G. If neither endpoint were in S, then both endpoints would lie in I, contradicting that I is independent. Thus every edge has at least one endpoint in S, so S is a vertex cover. □

Lemma 18.7. For every n-vertex graph G,

$$\alpha(G) + \beta(G) = n,$$

where $\alpha(G)$ is the maximum size of an independent set and $\beta(G)$ is the minimum size of a vertex cover.

Proof. By Lemma 18.6, S is a vertex cover iff $V(G) \setminus S$ is independent. So for any vertex cover S we have

$$|V(G) \setminus S| \le \alpha(G) \qquad \Longrightarrow \qquad n - |S| \le \alpha(G) \qquad \Longrightarrow \qquad |S| \ge n - \alpha(G).$$

Taking the minimum over all vertex covers gives $\beta(G) \ge n - \alpha(G)$.

Conversely, let I be a maximum independent set with $|I| = \alpha(G)$. Then $V(G) \setminus I$ is a vertex cover by Lemma 18.6, hence

$$\beta(G) \le |V(G) \setminus I| = n - \alpha(G).$$

Combining both inequalities yields $\beta(G) = n - \alpha(G)$, i.e. $\alpha(G) + \beta(G) = n$. □

Lemma 18.8. For every graph G,

$$\alpha'(G) \le \beta(G) \le 2\alpha'(G),$$

where $\alpha'(G)$ is the size of a maximum matching and $\beta(G)$ is the size of a minimum vertex cover.

Proof. **Lower bound $\alpha'(G) \le \beta(G)$.** Let M be a matching. Any vertex cover must contain at least one endpoint of each edge in M, and the edges in M are disjoint, so covering them requires at least $|M|$ vertices. Thus $\beta(G) \ge |M|$ for every matching M, hence $\beta(G) \ge \alpha'(G)$.

Upper bound $\beta(G) \le 2\alpha'(G)$. Let M be a *maximal* matching (cannot be extended by adding an edge). Let S be the set of endpoints of edges in M; then $|S| = 2|M|$. We claim S is a vertex cover: if there were an edge uv with $u, v \notin S$, then u and v are unmatched by M, so the edge uv could be added to M, contradicting maximality. Hence every edge meets S.

Therefore $\beta(G) \le |S| = 2|M|$. Finally, since $|M| \le \alpha'(G)$ (a maximum matching is at least as large as any matching), we get $\beta(G) \le 2\alpha'(G)$. □

Warning: maximal vs. maximum.

A *maximum* matching is one with the largest possible size (globally optimal). A *maximal* matching is one that cannot be extended by adding another edge (locally stuck everywhere). Maximal ≠ maximum: a maximal matching can be far from optimal.

Proposition 18.9. Let M be a maximal matching in a graph G, and let M^* be a maximum matching. Then

$$|M| \;\geq\; \tfrac{1}{2}\,|M^*|.$$

Equivalently, any maximal matching is a 2-approximation to a maximum matching.

Proof. Let S be the set of endpoints of edges in M, so $|S| = 2|M|$. Since M is maximal, S is a vertex cover: if an edge had both endpoints outside S, we could add it to M, contradicting maximality.

Now every edge of the maximum matching M^* must meet this vertex cover S. Because edges in a matching are disjoint, each vertex of S can cover *at most one* edge of M^*. Hence

$$|M^*| \;\leq\; |S| \;=\; 2|M|.$$

Rearranging gives $|M| \geq \frac{1}{2}|M^*|$. □

Theorem 18.10 (König–Egerváry, 1931). In every bipartite graph G,

$$\alpha'(G) = \beta(G).$$

Equivalently: maximum matching size equals minimum vertex cover size.

Proof. First, every edge of a matching must be covered by a *different* vertex, hence $\beta(G) \geq \alpha'(G)$.

For the reverse inequality, apply the Defect Formula. Choose $T \subseteq X$ such that

$$\alpha'(G) = |X| - |T| + |N(T)|.$$

Consider the set

$$C := (X \setminus T) \;\cup\; N(T).$$

We claim C is a vertex cover. Indeed, any edge not incident to $X \setminus T$ must meet T, hence meets $N(T)$ on the other side. So no edge is uncovered.

Therefore

$$\beta(G) \leq |C| = |X| - |T| + |N(T)| = \alpha'(G).$$

□

18.6 Edge covers and Gallai's Theorem

Definition 18.7 (Edge cover). An *edge cover* is a set of edges that covers all vertices. Let $\beta'(G)$ be the minimum size of an edge cover.

Theorem 18.11 (Gallai)**.** If G has no isolated vertices and $|V(G)| = n$, then

$$\alpha'(G) + \beta'(G) = n.$$

Proof. Let M be a maximum matching.

Upper bound. Build an edge cover by taking all edges of M and for each vertex not covered by M, add one incident edge. This yields an edge cover of size $n - |M|$, so

$$\beta'(G) \leq n - \alpha'(G) \quad \Rightarrow \quad \alpha'(G) + \beta'(G) \leq n.$$

Lower bound. Let L be a minimum edge cover. In L, each edge has an endpoint not covered by any other edge of L; otherwise we could delete it and still cover all vertices. Hence L is a forest with no P_4, so every component is a star. Say L has k star components. Then

$$|L| = n - k.$$

Taking one edge from each star gives a matching of size k, so $\alpha'(G) \geq k$. Therefore

$$\alpha'(G) + \beta'(G) \geq k + (n - k) = n.$$

□

Theorem 18.12 (König, 1916)**.** If G is bipartite with no isolated vertices, then

$$\alpha(G) = \beta'(G).$$

Proof. Using the identities

$$\alpha(G) + \beta(G) = n, \qquad \alpha'(G) + \beta'(G) = n,$$

and König–Egerváry $\alpha'(G) = \beta(G)$, we get

$$\alpha(G) = n - \beta(G) = n - \alpha'(G) = \beta'(G).$$

□

Legend.

1. $\alpha(G)$ = maximum independent set size
2. $\alpha'(G)$ = maximum matching size
3. $\beta(G)$ = minimum vertex cover size
4. $\beta'(G)$ = minimum edge cover size

Statement	General graphs	Bipartite graphs	Attribution / note
$\alpha(G) + \beta(G) = n$	equality	equality	Complement of a vertex cover is an independent set.
$\alpha'(G) \le \beta(G)$	inequality	equality	$\le$ is trivial; $=$ is Kőnig–Egerváry (1931).
$\beta(G) \le 2\alpha'(G)$	inequality	inequality	Endpoints of a *maximal* matching form a vertex cover.
$\alpha'(G) + \beta'(G) = n$ (no isolated vertices)	equality	equality	Gallai (1959); assumes no isolated vertices.
$\alpha(G) = \beta'(G)$ (no isolated vertices)	false in general	equality	Kőnig (1916), derived from Kőnig–Egerváry + the identities above.

19 Matchings in general graphs

Definition 19.1 (k-factor)**.** A *k-factor* of G is a k-regular spanning subgraph. A 1-factor is called a *perfect matching*.

Definition 19.2 (Odd components)**.** An *odd component* is a connected component with an odd number of vertices. Let $o(G)$ be the number of odd components of G.

19.1 Tutte's 1-factor theorem

Motivation: Hall's theorem completely settles perfect matchings in *bipartite* graphs by looking at neighborhoods. How about all graphs in general?

We want a clean description of *all* graphs that fail to have a perfect matching, so we start by listing the most obvious obstructions and then generalize.

The first obstruction is parity of the number of vertices: if $|V(G)|$ is odd, then no matching can cover all vertices, since edges cover vertices two at a time.

The parity also shows up more subtly in disconnected graphs. Even when $|V(G)|$ is even, if G has a connected component with an odd number of vertices (an *odd component*), then a perfect matching is still impossible: clearly any matching pairs vertices *within* components, so an odd component must leave at least one vertex unmatched.

Now imagine G is connected, but removing a vertex u causes $G - u$ to split into (say) two odd components. In each odd component of $G - u$, some vertex must remain unmatched internally, and the only way to match it is to use an edge to the deleted vertex u. But u can be matched to at most one vertex, so if $G - u$ has two odd components, at least one leftover vertex cannot be rescued. No perfect matching exists.

Removing vertices can create many odd components, and each odd component needs an "escape" to the removed set. Formally, let $U \subseteq V(G)$ and consider $G - U$. In any matching, every odd component of $G - U$ must contribute at least one vertex that is unmatched *inside that component* (because it has odd order). The only way to cover that leftover vertex is to match it across an edge into U. Thus each odd component of $G - U$ demands at least one distinct vertex of U, so a perfect matching can exist only if

$$o(G - U) \leq |U| \qquad \text{for every } U \subseteq V(G),$$

where $o(G - U)$ denotes the number of odd components of $G - U$.

This is why Tutte's condition is stated in terms of odd components. Tutte's theorem says this necessary parity obstruction is also sufficient: *the only reason a general graph fails to have a perfect matching is that, after removing some set U, too many odd components are created.*

If G has a 1-factor, then for every $S \subseteq V(G)$,

$$o(G - S) \leq |S|.$$

Theorem 19.1 (Tutte, 1947)**.** A graph G has a 1-factor iff

$$o(G - S) \leq |S| \quad \text{for all } S \subseteq V(G).$$

(This is called *Tutte's condition*.)

The following proof of Tutte's theorem is due to László Lovász.

Proof. Applying Tutte's condition with $X = \varnothing$ yields $o(G) = o(G - \varnothing) \leq |\varnothing| = 0$, so G has no odd components. In particular, every component of G has even order, hence $|V(G)|$ is even.

Suppose for contradiction, that there exists a graph on n vertices satisfying Tutte's condition but having no perfect matching. Among all such n-vertex graphs, choose one G with the *maximum possible number of edges*. Thus:

- G satisfies $o(G - X) \leq |X|$ for every $X \subseteq V(G)$,
- G has no perfect matching, and
- adding any missing edge to G produces a graph *with* a 1-factor (indeed, $G \neq K_n$ else it has a perfect matching)

Define

$$U := \{v \in V(G) : v \text{ is adjacent to every other vertex of } G\},$$

and

$$W := V(G) \setminus U.$$

We call vertices in U as universal vertices, and vertices in W as non-universal vertices.

Let G'' be the subgraph of G induced by W. Note that $G'' = G - U$.

Claim: G'' is a disjoint union of cliques

Equivalently, adjacency is an equivalence relation on W. Reflexivity and symmetry are automatic; we must show transitivity.

Lemma 19.2. If $a, b, c \in W$ with $ab, bc \in E(G)$, then $ac \in E(G)$.

Proof. Suppose for contradiction, that $a, b, c \in W$ with

$$ab, bc \in E(G) \quad \text{but} \quad ac \notin E(G).$$

Since $b \in W$ is not universal, there exists some vertex d with $bd \notin E(G)$.

By maximality of G with respect to edges:

- In the graph $G_1 := G + ac$, there is a 1-factor F_1.
- In the graph $G_2 := G + bd$, there is a 1-factor F_2.

Moreover, each of the added edges *must* lie in the corresponding 1-factor: if $ac \notin F_1$, then F_1 would be a 1-factor of G; similarly for bd and F_2. Thus

$$ac \in F_1, \qquad bd \in F_2,$$

and clearly $ac \notin F_2$, $bd \notin F_1$.

Consider the union $F_1 \cup F_2$ as a graph on $V(G)$, with edge set contained in $E(G) \cup \{ac, bd\}$.

Lemma 19.3. $F_1 \cup F_2$ decomposes into a disjoint union of cycles and isolated edges; on each cycle, edges alternate between F_1 and F_2.

Proof. Each vertex has degree 1 in F_1 and in F_2, hence degree ≤ 2 in $H := F_1 \cup F_2$. Thus every component of H is a path, a cycle, or a single edge.

If v has $d_H(v) = 1$ and $e = vw$ is its unique incident edge in H, then both matchings must use e at v, so $e \in F_1 \cap F_2$. At w the same argument shows $d_H(w) = 1$; otherwise some matching would give w degree 2. Hence this component is just the isolated edge vw and not a longer path.

If a component has no vertex of degree 1, then all its vertices have degree 2, so it is a cycle. At each vertex on this cycle one incident edge is from F_1 and one from F_2, so the edges alternate between F_1 and F_2.

Thus components are alternating cycles or isolated edges, and in particular there are no nontrivial paths. □

Let C be the unique cycle in $F_1 \cup F_2$ that contains the edge ac. We distinguish two cases depending on whether bd lies on C.

Case 1: $bd \notin C$**.**

Along C the edges alternate between F_1 and F_2; in particular ac is an F_1-edge on C. Define

$$F_1' := \big(F_1 \setminus E(C)\big) \cup \big(F_2 \cap E(C)\big).$$

On C we swap the roles of F_1 and F_2; outside C we leave F_1 unchanged.

Because C is an alternating cycle, every vertex on C still has degree 1 in F_1', and vertices outside C are unchanged; thus F_1' is a 1-factor of G_1.

Note that:

- $ac \in E(C) \cap F_1$ and $ac \notin F_2$, so ac is removed and not reinserted; hence $ac \notin F_1'$.
- All edges of C other than ac and bd belong to $E(G)$. Since $bd \notin C$ in this case, every edge of C other than ac actually lies in $E(G)$.

Therefore every edge of F_1' belongs to $E(G)$, so F_1' is a 1-factor of G itself, contradicting that G has no 1-factor.

Thus Case 1 is impossible.

Case 2: $bd \in C$.

Now C contains both ac and bd. Deleting ac and bd from C splits it into two vertex-disjoint paths. Exactly one of these paths has d as an endpoint; call that path P.

We may relabel a and c (if necessary) so that P has endpoints a and d: we assume P goes from a to d.

Consider the cycle

$$C' := P \cup \{bd, ab\}.$$

Here $ab \in E(G)$ by assumption, and bd is the added edge used in G_2.

Lemma 19.4. C' is an alternating cycle with respect to F_2.

Proof of claim. On C, edges alternate between F_1 and F_2. The path $P \subseteq C$ therefore alternates between F_1 and F_2. The edge bd is in F_2 by construction, whereas $ab \notin F_2$ because b is already matched to d in F_2. Tracing around C' we encounter edges alternately in F_2 and outside F_2, so C' is an F_2-alternating cycle. □

Define

$$F_2' := \big(F_2 \setminus E(C')\big) \cup \big(E(C') \setminus F_2\big),$$

i.e., swap membership of edges along C'. Again, since C' is alternating, F_2' is a 1-factor of G_2.

Moreover:

- bd is an F_2-edge of C', so it is removed and not reinserted; thus $bd \notin F_2'$.
- All other edges of C' lie in $E(G)$.

Hence every edge of F_2' belongs to $E(G)$, so F_2' is a 1-factor of G, contradicting again that G has no 1-factor.

Both cases lead to contradictions. Therefore our assumption that $ab, bc \in E(G)$ but $ac \notin E(G)$ is false, and adjacency is transitive on W.

Thus $G'' = G[W]$ is a disjoint union of complete graphs (each component is a clique). □

Let the connected components of G'' be

$$H_1, \dots, H_t,\ K_1, \dots, K_s,$$

where each H_i has odd order and each K_j has even order (so there are t odd components and s even components in G'').

Lemma 19.5. If $t \le |U|$, then G has a 1-factor.

Proof. Assume $t \le |U|$. We explicitly construct a perfect matching M of G.

Step 1: Match inside even components. Each K_j is a complete graph of even order, hence admits a 1-factor. Fix one 1-factor M_j in each K_j.

Step 2: Use U to fix parity in odd components. For each odd component H_i (a clique of odd order), choose a vertex $u_i \in V(H_i)$. Every vertex in U is adjacent to all vertices of G, so in particular each u_i is adjacent to every vertex of U.

Since $t \leq |U|$, we can choose t distinct vertices $v_1, \ldots, v_t \in U$ and match them with $u_1, \ldots, u_t$ via edges

$$v_1u_1, \ldots, v_tu_t \in E(G).$$

Step 3: Match remaining vertices in odd components. For each i, the induced subgraph $H_i - \{u_i\}$ is a clique on $|H_i| - 1$ vertices, which is even. Thus it has a 1-factor; call it N_i.

Step 4: Match remaining vertices in U. Let

$$R := U \setminus \{v_1, \ldots, v_t\}$$

be the set of vertices in U not yet matched.

We claim that $|R|$ is even. Since G has no odd components, we already know $|V(G)|$ is even. Count the vertices already matched in Steps 1–3:

- For each even component K_j, M_j matches all vertices of K_j.
- For each odd component H_i, the matching N_i covers $V(H_i) \setminus \{u_i\}$, and the edge v_iu_i covers u_i and v_i.

Thus all vertices outside R are matched in pairs; their total number is even, so $|R|$ is also even.

The induced subgraph $G[R]$ is complete (as a subset of U), so $G[R]$, being a clique on an even number of vertices, has a 1-factor L.

Step 5: Combine all edges. Define

$$M := \Big(\bigcup_j M_j\Big) \;\cup\; \Big(\bigcup_i N_i\Big) \;\cup\; \{v_1u_1, \ldots, v_tu_t\} \;\cup\; L.$$

By construction:

- Each vertex of each K_j is incident with exactly one edge of M_j.
- Each vertex of each H_i is incident with exactly one edge in $N_i \cup \{v_iu_i\}$.
- Each vertex of U is incident with exactly one edge in $\{v_1u_1, \ldots, v_tu_t\} \cup L$.

These sets of edges are pairwise vertex-disjoint, so M is a 1-factor of G.

This contradicts the assumption that G has no 1-factor. Therefore our assumption $t \leq |U|$ is false. □

Hence we must have

$$t > |U|.$$

But the components $H_1, \ldots, H_t$ are exactly the odd components of G'', and $G'' = G[W] = G - U$. Thus

$$o(G - U) = t > |U|.$$

This contradicts Tutte's condition in G. Therefore no such counterexample graph G exists. Hence every graph satisfying Tutte's condition has a 1-factor, completing the proof of the sufficiency direction of Tutte's theorem. □

Definition 19.3 (Deficiency). For $S \subseteq V(G)$ define the *deficiency*

$$\operatorname{def}(S) := o(G-S) - |S|.$$

Define

$$\operatorname{def}(G) := \max_{S \subseteq V(G)} \operatorname{def}(S).$$

Lemma 19.6 (Parity). For every $S \subseteq V(G)$,

$$\operatorname{def}(S) = o(G-S) - |S| \equiv n \pmod 2,$$

where $n = |V(G)|$.

Idea. Count $\big(o(G-S) + |S|\big)$ modulo 2 by partitioning $V(G)$ into components of $G-S$ plus S. □

Fix $T \subseteq V(G)$ with $\operatorname{def}(T) = \operatorname{def}(G)$, and among such sets choose T maximal under inclusion.

Lemma 19.7. Let $T \subseteq V(G)$ be such that $\operatorname{def}_G(T) = \operatorname{def}(G)$ and, among all such sets, T is maximal with respect to inclusion. Then:

1. If u lies in an odd component C of $G-T$, then $C-u$ satisfies Tutte's condition (i.e. $C-u$ has a 1-factor).
2. $G-T$ has no even component.

Proof. Recall the definitions:

$$\operatorname{def}_G(S) := o(G-S) - |S|, \qquad \operatorname{def}(G) := \max_{S \subseteq V(G)} \operatorname{def}_G(S),$$

where $o(H)$ denotes the number of odd components of a graph H.

(i) $C-u$ **satisfies Tutte's condition.**

Let C be an odd component of $G-T$, and fix $u \in V(C)$. Consider the graph $C-u$, and for each $S \subseteq V(C-u)$ define the *local deficiency*

$$\operatorname{def}_{C-u}(S) := o\big((C-u) - S\big) - |S|.$$

We want to show $\operatorname{def}_{C-u}(S) \le 0$ for all S, which is exactly Tutte's condition on $C-u$.

Key idea. If for some S the local deficiency $\operatorname{def}_{C-u}(S)$ were positive, then by combining S with T and u we would build a larger set $T' := T \cup S \cup \{u\}$ whose global deficiency $\operatorname{def}_G(T')$ is at least as large as $\operatorname{def}_G(T)$, contradicting the maximality of T.

Fix $S \subseteq V(C-u)$. Define

$$T' := T \cup S \cup \{u\}.$$

We compare $\operatorname{def}_G(T')$ with $\operatorname{def}_G(T)$.

First, analyze the odd components:

- In $G - T$, the component C is odd. When we pass from $G - T$ to $G - T'$, we are removing the vertices u and S from C, and we do *nothing* to any other component of $G - T$, because $S \subseteq V(C - u)$ lies entirely inside C.
- The other components of $G - T$ (besides C) remain exactly as they are when we go to $G - T'$, so they keep their odd/even status.

Inside C, after removing u and S, the remaining part is $(C - u) - S$, whose odd components are exactly the ones counted by $o((C - u) - S)$.

Therefore the total number of odd components of $G - T'$ is

$$o(G - T') = \big(o(G - T) - 1\big) + o\big((C - u) - S\big).$$

(The "−1" accounts for the fact that C itself disappears and is replaced by the components of $(C - u) - S$.)

Now compute the deficiency:

$$\begin{aligned}\mathrm{def}_G(T') &= o(G - T') - |T'| \\ &= \big(o(G - T) - 1 + o((C - u) - S)\big) \;-\; \big(|T| + |S| + 1\big) \\ &= \big(o(G - T) - |T|\big) \;+\; \big(o((C - u) - S) - |S| - 2\big) \\ &= \mathrm{def}_G(T) \;+\; \mathrm{def}_{C-u}(S) - 2.\end{aligned}$$

Thus we have the exact relation

$$\mathrm{def}_G(T') = \mathrm{def}_G(T) + \mathrm{def}_{C-u}(S) - 2.$$

Now recall the parity lemma (applied to $C - u$): since C is odd, $C - u$ has an even number of vertices, and for any $S \subseteq V(C - u)$ we have

$$\mathrm{def}_{C-u}(S) \equiv |V(C - u)| \equiv 0 \pmod 2,$$

so $\mathrm{def}_{C-u}(S)$ is an even integer.

Suppose, for contradiction, that there exists S with $\mathrm{def}_{C-u}(S) > 0$. Then $\mathrm{def}_{C-u}(S) \geq 2$, and so

$$\mathrm{def}_G(T') = \mathrm{def}_G(T) + \mathrm{def}_{C-u}(S) - 2 \geq \mathrm{def}_G(T).$$

Now two things happen:

- Since $\mathrm{def}(G)$ is the *maximum* deficiency, we must have $\mathrm{def}_G(T') \leq \mathrm{def}(G) = \mathrm{def}_G(T)$, so in fact $\mathrm{def}_G(T') = \mathrm{def}_G(T)$.
- But $T' \supsetneq T$ (we added at least u), so we have found a strictly larger set with the same maximal deficiency, contradicting the assumption that T was maximal by inclusion among those sets.

Therefore our assumption was impossible, and we conclude that

$$\mathrm{def}_{C-u}(S) \leq 0 \quad \text{for all } S \subseteq V(C - u),$$

which means $C - u$ satisfies Tutte's condition.

(ii) $G - T$ **has no even component.**

Assume, for contradiction, that $G - T$ has an even component D. We will enlarge T by one vertex in D without changing its deficiency, again contradicting the maximality of T.

Because D is connected and finite, we can choose some vertex $v \in D$ that is not a cut vertex of D (for instance, a leaf of a spanning tree of D). Then $D - v$ remains connected, and since $|D|$ is even, $|D - v|$ is odd. So $D - v$ is a *single odd component*.

Now consider $T' := T \cup \{v\}$. Then

$$G - T' = (G - T) - v.$$

In $G - T$ the component D was even, so it did not contribute to $o(G - T)$. In $G - T'$, that same vertex set becomes $D - v$, which is an odd component. All other components are unaffected.

Thus the number of odd components increases by 1:

$$o(G - T') = o(G - T) + 1.$$

At the same time, the size of T increases by 1:

$$|T'| = |T| + 1.$$

Hence

$$\mathrm{def}_G(T') = o(G - T') - |T'| = (o(G - T) + 1) - (|T| + 1) = o(G - T) - |T| = \mathrm{def}_G(T).$$

So T' has the same (maximal) deficiency as T, but $T' \supsetneq T$, again contradicting the maximality of T with respect to inclusion.

Therefore no such even component D can exist, and $G - T$ has only odd components. □

19.2 Berge-Tutte formula

Theorem 19.8 (Berge–Tutte Formula)**.** For every graph G on n vertices,

$$\alpha'(G) = \frac{1}{2}\big(n - \mathrm{def}(G)\big).$$

Proof. Let $d := \mathrm{def}(G) = \max_S \mathrm{def}(S)$. For any S, a matching misses at least $\mathrm{def}(S)$ vertices, hence

$$\alpha'(G) \le \frac{1}{2}\big(n - \mathrm{def}(S)\big),$$

so $\alpha'(G) \le \frac{1}{2}(n - d)$.

It remains to find a matching covering all but d vertices. Proceed by induction on n.

Let T be maximal with deficiency d. By the previous lemma, $G - T$ has only odd components, and for every odd component C and every $u \in V(C)$, the graph $C - u$ has a 1-factor; by induction, each $C - u$ has a perfect matching.

There are $|T| + d$ odd components of $G - T$. It suffices to match the vertices of T into distinct odd components.

Form an auxiliary bipartite graph H as follows:

- left class: the vertices of T,
- right class: the odd components of $G - T$,
- we put an edge tC in H if the vertex $t \in T$ has at least one neighbor in the component C (in the original graph G).

Goal. We want to use Hall's theorem on H to match every vertex $t \in T$ to a *distinct* odd component of $G - T$. Intuitively, this means: each vertex of T will be "assigned" to a different odd component where it has a neighbor, so we can later attach t into that component and use the perfect matching of $C - u$ inside.

Thus we must show that H satisfies Hall's condition:

$$\forall S \subseteq T: \qquad |S| \le |N_H(S)|.$$

Fix some subset $S \subseteq T$. We will prove $|S| \le |N_H(S)|$ by comparing deficiencies.

First recall that

$$\mathrm{def}_G(T) = d \quad \text{and} \quad \mathrm{def}_G(T) = o(G - T) - |T|$$

so

$$o(G - T) = |T| + d.$$

In words: the graph $G - T$ has exactly $|T| + d$ odd components.

Now consider the smaller set $T \setminus S$. The graph $G - (T \setminus S)$ is obtained from $G - T$ by *adding back* the vertices of S (together with their incident edges). We want a lower bound on the number of odd components of $G - (T \setminus S)$.

Key observation. An odd component C of $G - T$ is adjacent to S in H *iff* some vertex of C has a neighbor in S (in G).

Now look at $G - (T \setminus S)$:

- The vertices of S are present again, so any odd component of $G - T$ that has a neighbor in S might merge with some other components or change its structure. These components are exactly those in $N_H(S)$.
- On the other hand, any odd component C of $G - T$ that is *not* adjacent to S in H (i.e. has no edges to S) is completely untouched by adding back S: no new edges connect C to anything, so C remains an isolated component of $G - (T \setminus S)$, and it stays odd.

Therefore, when passing from $G - T$ to $G - (T \setminus S)$, *at least* all odd components of $G - T$ that are not in $N_H(S)$ remain odd components. Hence

$$o\big(G - (T \setminus S)\big) \ \ge\ \big(\text{all odd components of } G - T\big) - \big(\text{those adjacent to } S\big) = (|T| + d) - |N_H(S)|.$$

Now compute the deficiency of $T \setminus S$:

$$\begin{aligned}
\mathrm{def}_G(T \setminus S) &= o\big(G - (T \setminus S)\big) - |T \setminus S| \\
&\ge (|T| + d) - |N_H(S)| - (|T| - |S|) \\
&= d - |N_H(S)| + |S|.
\end{aligned}$$

By the definition of $d = \mathrm{def}(G)$ as the *maximum* deficiency over all vertex sets, we know

$$\mathrm{def}_G(T \setminus S) \le d.$$

Combining this with the inequality above gives

$$d - |N_H(S)| + |S| \ \le\ \mathrm{def}_G(T \setminus S) \ \le\ d,$$

so

$$|S| \le |N_H(S)|.$$

Since this is true for every $S \subseteq T$, the bipartite graph H satisfies Hall's condition and therefore has a matching that saturates T.

Finishing the construction. Thus we can assign to each $t \in T$ a distinct odd component C of $G - T$ in which t has a neighbor. In each such C, we already know (from part (i)) that for any chosen vertex $u \in C$, the graph $C - u$ has a perfect matching. We choose u to be the neighbor of t in C.

So for each matched pair (t, C):

- take the perfect matching of $C - u$ inside C,
- add the edge tu.

This covers every vertex of T and all vertices in those components except for exactly one leftover vertex per component. Counting carefully (using that $G - T$ has $|T| + d$ odd components) shows that in total we miss exactly d vertices of G.

Hence we obtain a matching of size

$$\frac{n-d}{2},$$

so

$$\alpha'(G) \geq \frac{n-d}{2}.$$

Together with the upper bound $\alpha'(G) \leq (n-d)/2$ proved earlier, this yields

$$\alpha'(G) = \frac{1}{2}(n-d),$$

which is the Berge–Tutte formula.

□

Theorem 19.9. Let G be a k-regular multigraph on an even number n of vertices. Assume G is $(k-1)$-edge-connected (i.e., removing any $k-2$ edges keeps G connected, equivalently every nontrivial edge cut has size at least $k-1$). Let $G' = G - F$ where F is any set of $k-1$ edges. Then G' has a 1-factor (a perfect matching).

Proof. We argue by contradiction. Suppose G' has no perfect matching. By Tutte's 1-factor theorem, there exists a set $S \subseteq V(G')$ such that

$$o(G' - S) - |S| \geq 2. \tag{$\star$}$$

So the deficiency of S in G' is at least 2.

Let m be the number of edges of G' that join S to the odd components of $G' - S$ (i.e. edges with one endpoint in S and the other in an odd component of $G' - S$).

Upper bound on m. Every vertex in G has degree k, and G' is obtained from G by deleting $k-1$ edges, so in G' every vertex has degree at most k. Each edge counted in m is incident with exactly one vertex of S, and each vertex of S is incident with at most k edges in G'. Hence

$$m \leq k|S|. \tag{6}$$

Now we derive a *lower* bound on m using regularity and edge-connectivity.

Fix an odd component H of $G' - S$. Let ℓ_H be the number of edges of G with exactly one endpoint in H:

$$\ell_H := |\{uv \in E(G) : u \in V(H),\ v \notin V(H)\}|.$$

In other words, ℓ_H is the size of the edge cut $[V(H), \overline{V(H)}]$ in G.

Because G is k-regular,

$$\sum_{v \in V(H)} d_G(v) = k|H|.$$

Counting degrees inside H another way: each edge of H contributes 2, and each edge leaving H contributes 1. Thus

$$\sum_{v \in V(H)} d_G(v) = 2|E(H)| + \ell_H,$$

hence

$$k|H| = 2|E(H)| + \ell_H \quad \Rightarrow \quad \ell_H = k|H| - 2|E(H)|.$$

Parity of ℓ_H. The term $2|E(H)|$ is even, so ℓ_H has the same parity as $k|H|$. Since H is an odd component, $|H|$ is odd, hence

$$\ell_H \equiv k|H| \equiv k \pmod 2.$$

Using edge-connectivity. G is $(k-1)$-edge-connected, so every nontrivial edge cut has size at least $k-1$; in particular $\ell_H \geq k-1$. Combining with $\ell_H \equiv k \pmod 2$ forces

$$\ell_H \geq k. \tag{7}$$

(Indeed, the smallest integer $\geq k-1$ with the same parity as k is k.)

Relating ℓ_H to m and the deleted edges.

For each odd component H of $G' - S$, define:

- a_H = number of edges of G' with one endpoint in H and the other in S. These are exactly the edges of G' from H to S. By definition,

$$m = \sum_H a_H.$$

- b_H = number of edges in the deleted set F with exactly one endpoint in H.

Claim: For each odd component H,

$$\ell_H = a_H + b_H.$$

Reason: any edge of G with exactly one endpoint in H either:

- survives in G' and then must go from H to S (it cannot go to another component of $G' - S$, or those components would be connected), and is counted in a_H;
- or it is one of the deleted edges in F and is counted in b_H.

There are no other possibilities. Thus $\ell_H = a_H + b_H$.

Summing over all odd components of $G' - S$ gives

$$\sum_H \ell_H = \sum_H a_H + \sum_H b_H = m + \sum_H b_H.$$

From (7), summing over all odd components yields

$$\sum_H \ell_H \;\geq\; k \cdot o(G' - S).$$

So

$$m + \sum_H b_H \;\geq\; k \cdot o(G' - S).$$

Each deleted edge in F can be incident with at most two odd components of $G' - S$, so it contributes at most 2 to the sum $\sum_H b_H$. Therefore

$$\sum_H b_H \;\leq\; 2|F| = 2(k-1),$$

and hence

$$m + 2(k-1) \;\geq\; k \cdot o(G' - S),$$

i.e.

$$m \;\geq\; k \cdot o(G' - S) - 2(k-1). \tag{8}$$

Final contradiction. Combine the upper bound (6) and lower bound (8):

$$k|S| \;\geq\; m \;\geq\; k \cdot o(G' - S) - 2(k-1),$$

so

$$k\big(o(G' - S) - |S|\big) \;\leq\; 2(k-1).$$

Dividing by $k > 0$,

$$o(G' - S) - |S| \;\leq\; \frac{2(k-1)}{k}.$$

Since $\frac{2(k-1)}{k} < 2$ for all $k \geq 2$ and the left-hand side is an integer, we must have

$$o(G' - S) - |S| \leq 1.$$

But this contradicts ($\star$), which says $o(G' - S) - |S| \geq 2$. Therefore our assumption that G' has no perfect matching was false, and G' does in fact have a perfect matching. □

Remark 19.1 (Petersen (1891)). If G is 3-regular and has no cut-edge (i.e., is 2-edge-connected), then G has a 1-factor. This is the case $k = 3$ of the theorem.

Theorem 19.10 (Berge). Let G be a k-regular multigraph on an even number of vertices, and assume G is $(k-1)$-edge-connected. Then *every edge of G lies in some perfect matching*.

Proof. Fix an edge $uv \in E(G)$. Remove the other $k-1$ edges incident to u, obtaining a graph G'. By the previous theorem G' has a perfect matching, and since u has only the edge uv remaining, that matching must contain uv. □

Definition 19.4 (f-factor). Given G and a function $f : V(G) \to \mathbb{N}$, an *f-factor* of G is a spanning subgraph $H \subseteq G$ such that

$$d_H(v) = f(v) \quad \text{for all } v \in V(G).$$

A *k-factor* is an f-factor with $f \equiv k$. A 1-factor is a perfect matching.

Theorem 19.11 (Petersen 2-factor theorem). Every $2k$-regular multigraph has a 2-factor. In fact, every $2k$-regular multigraph decomposes into edge-disjoint 2-factors.

Proof. Assume G is connected (work componentwise otherwise). Since all degrees are even, G has an Eulerian circuit. Orient each edge in the direction traversed by the Eulerian circuit. Then each vertex has indegree k and outdegree k.

Build a bipartite graph $B = (U, W)$ with $U = W = V(G)$, and put an edge $u \in U$ to $w \in W$ for each directed edge $u \to w$ in G. Then B is k-regular bipartite, hence decomposes into k perfect matchings. Each perfect matching selects exactly one outgoing edge and one incoming edge at every vertex of G, which forms a 2-factor. Taking all k perfect matchings yields a decomposition of G into k 2-factors. □

19.3 Algorithmic aspects of matchings

This section is about *how you can actually find* matchings (not just prove they exist). The recurring theme is the same in every setting:

Find an augmenting path, flip it, and your matching gets bigger.

Definition 19.5 (M-alternating and M-augmenting paths). Let M be a matching in G. A path P is *M-alternating* if its edges alternate between belonging to M and not belonging to M. An M-alternating path whose endpoints are not covered by M is an *M-augmenting path*.

Theorem 19.12 (Berge). A matching M in G is maximum iff there is no M-augmenting path.

Proof. ($\Rightarrow$) If P is an M-augmenting path, then replacing the M-edges of P by the non-M edges of P increases the size of the matching:

$$M' := M \triangle E(P) \quad \text{satisfies} \quad |M'| = |M| + 1,$$

contradicting maximality.

($\Leftarrow$) Assume M has no augmenting path, but M' is a larger matching. Consider the symmetric difference

$$F := M \triangle M'.$$

Every vertex has degree at most 2 in F, so each component of F is an even cycle or a path alternating between edges of M and M'. Since $|M'| > |M|$, some component must be a path with more M'-edges than M-edges. Such a path begins and ends with M'-edges, so its endpoints are not covered by M, making it an M-augmenting path—contradiction. □

Algorithms for bipartite matchings: In bipartite graphs $G = (X, Y)$, augmenting paths are easy to compute since no odd cycles exist. There are efficient polynomial-time algorithms for maximum matching (e.g. Hopcroft–Karp).

The key structural bonus is Kőnig–Egerváry:

$$\alpha'(G) = \beta(G) \qquad \text{(bipartite } G\text{)}.$$

This is algorithmically important because *minimum vertex cover is NP-hard in general graphs*: outside the bipartite world, we should not expect an efficient exact algorithm.

In bipartite graphs, however, Kőnig–Egerváry says the "hard" problem (minimum vertex cover) *collapses* to the "easy" one (maximum matching):

> compute a maximum matching M (polynomial time), then *extract* from M a vertex cover C with $|C| = |M|$ (also polynomial time).

In general graphs, the augmenting-path search can be obstructed by *odd cycles*. Edmonds' blossom algorithm handles this by contracting odd cycles during the search.

Edmonds' idea is: treat the entire odd cycle as a single super-vertex. After contracting the blossom, the parity ambiguity disappears and the search can continue in the smaller graph. If an augmenting path is found in the contracted graph, you can *expand* the blossom and lift that augmenting path back to the original graph (there is a guaranteed way to route through the cycle so the matching increases by 1).

Approximation algorithms for maximum matchings and minimum vertex covers

From earlier we have:

(a) $\alpha'(G) \le \beta(G) \le 2\alpha'(G)$.

(b) If M is a *maximal* matching and M^* is maximum, then

$$|M| \ge \tfrac{1}{2}|M^*| = \tfrac{1}{2}\alpha'(G).$$

These two facts turn one greedy routine into several certified approximations.

> **Greedy routine.** Build a maximal matching M (scan edges; add an edge if both endpoints are free). This runs in $O(|E|)$ time.

(1) A 2-approximation for maximum matching size. By (b), we have $\alpha'(G) \le 2|M|$. So M is a factor-2 approximation to the maximum matching: it is guaranteed to be at least half-optimal.

(2) A certified factor-2 vertex cover. Let C be the set of endpoints of edges in M. Maximality implies C is a vertex cover, and $|C| = 2|M|$.

Moreover, any vertex cover must hit every edge in M, so $\beta(G) \ge |M|$. Hence we get the *certificate sandwich*

$$|M| \le \beta(G) \le |C| = 2|M|.$$

So in linear time we produce both

- a feasible cover C of size $2|M|$, and
- a matching lower bound $|M|$ proving no cover smaller than $|M|$ exists,

which immediately certifies that C is within factor 2 of optimal.

Bottom line. Even though minimum vertex cover is NP-hard in general graphs, the previously proved lemmas imply that a maximal matching computed in $O(|E|)$ time simultaneously yields:

a large matching (within factor 2 of optimal) and a small cover (within factor 2 of optimal),

20 Connectivity

20.1 Vertex connectivity

Recall: Let $G = (V, E)$ be a graph and let $S \subseteq V$. We write $G - S$ for the graph obtained from G by deleting all vertices in S and all edges incident to those vertices. Equivalently, $G - S$ is the induced subgraph on $V \setminus S$.

Definition 20.1 (Cut-set / vertex cut). A set $S \subseteq V(G)$ is a *cut-set* (or *vertex cut*) if $G - S$ is not connected.

Definition 20.2 (k-connected). A graph G is *k-connected* if

$$|V(G)| > k \quad \text{and} \quad \text{for every } S \subseteq V(G) \text{ with } |S| < k,\ G - S \text{ is connected.}$$

You must delete *at least* k vertices before you can disconnect the graph.

Definition 20.3 (Vertex connectivity). The *(vertex) connectivity* of G, denoted $\kappa(G)$, is the maximum integer k such that G is k-connected.

$$\kappa(G) \ := \ \min\big\{|S| : S \subseteq V(G) \text{ and } G - S \text{ is disconnected or has at most one vertex}\big\}.$$

If $|V(G)| = 1$, we set $\kappa(G) = 0$ by convention.

Remark 20.1. For vertex connectivity, multiedges do not matter. Therefore we may assume G is a *simple* graph.

Example 20.1. For $n \geq 2$, the complete graph K_n satisfies

$$\kappa(K_n) = n - 1.$$

Reason: if you remove fewer than $n - 1$ vertices, at least two vertices remain, and any two remaining vertices are adjacent, so the remaining graph is still connected. But if you remove $n - 1$ vertices, exactly one vertex remains, so you certainly cannot have connectivity $> n - 1$.

Example 20.2. The graph K_1 is connected, but

$$\kappa(K_1) = 0$$

by the convention built into the definition of k-connected (you cannot have $|V(G)| > k$ for any $k \geq 1$).

Example 20.3 (Cycles). For $n \geq 3$,

$$\kappa(C_n) = 2.$$

Reason: deleting one vertex from a cycle leaves a path, which is connected. Deleting another internal vertex of the path breaks it into two path components.

Example 20.4 (Complete bipartite graphs)**.** For $r, s \geq 1$,

$$\kappa(K_{r,s}) = \min\{r, s\}.$$

Reason: if you delete fewer than $\min\{r, s\}$ vertices, at least one vertex remains in *each* part, and then every remaining vertex in one part is adjacent to every remaining vertex in the other, so the graph stays connected. On the other hand, deleting all vertices from the smaller part (size $\min\{r, s\}$) leaves an independent set, hence disconnected (unless only one vertex remains, in which case you still cannot do better than $\min\{r, s\}$).

If a graph is k-connected, then in particular every vertex has degree at least k:

$$\kappa(G) \geq k \implies \delta(G) \geq k.$$

Hence, for an n-vertex k-connected graph,

$$2|E(G)| = \sum_{v \in V(G)} d(v) \geq nk \quad \implies \quad |E(G)| \geq \left\lceil \frac{nk}{2} \right\rceil.$$

The *Harary graph* $H_{k,n}$ is k-connected and has exactly $\lceil \frac{nk}{2} \rceil$ edges, proving this crude lower bound is actually *tight*.

Fix integers $2 \leq k < n$. Place vertices on a cycle and label them

$$V(H_{k,n}) = \{1, 2, \ldots, n\}$$

with arithmetic taken modulo n.

Let $t = \lfloor k/2 \rfloor$. Connect each vertex i to its t nearest neighbors on each side on the cycle:

$$i \sim i \pm 1,\ i \pm 2,\ \ldots,\ i \pm t.$$

After this step, every vertex has degree $2t$.

When k is odd: If k is even, we are done (since $2t = k$). If k is odd, then $2t = k - 1$, so we add one more "(almost) diagonal" per vertex in the following standard way.

- **Case A: k odd and n even.** Add the perfect matching of opposite vertices:

  $$i \sim i + \frac{n}{2} \qquad \text{for all } i \in \{1, \ldots, n\}.$$

  Now every vertex has degree $(k - 1) + 1 = k$.

- **Case B: k odd and n odd.** A k-regular simple graph on n odd vertices is impossible (since nk is odd), so the best you can do is "almost" regular while still hitting $\lceil \frac{nk}{2} \rceil$ edges. Add the $(n - 1)/2$ edges

  $$i \sim i + \frac{n-1}{2} \qquad \text{for } i = 1, 2, \ldots, \frac{n-1}{2}.$$

  This increases the degrees of $n - 1$ vertices by 1, leaving exactly one vertex with degree $k - 1$ after Step 1 (or equivalently: all but one vertex have degree k). One may also present an equivalent variant where exactly one vertex has degree $k + 1$ and the rest have degree k; either way the edge-count is optimal.

Definition 20.4 (Vertex k-split). Let G be a graph, let $x \in V(G)$, and let $k \geq 1$. A graph H is obtained from G by a *vertex k-split at x* if H is formed by:

- deleting x;
- introducing two new vertices x_1, x_2 with an edge x_1x_2;
- redistributing the edges incident to x among x_1 and x_2 so that

$$N_H(x_1) \cup N_H(x_2) = N_G(x) \cup \{x_1, x_2\},$$

  and

$$d_H(x_i) \geq k \quad \text{for each } i \in \{1, 2\}.$$

Intuitively: you "split" x into two adjacent vertices x_1, x_2 so that together they see all the old neighbors of x, and each new vertex still has degree at least k.

Lemma 20.1. If G is k-connected and H is obtained from G by a vertex k-split, then H is k-connected.

Proof. Suppose for contradiction that H has a vertex cut S with $|S| \leq k-1$. Let x be the vertex of G that is split into adjacent vertices x_1, x_2 in H, and define

$$T := \{x_1, x_2\} \cap S.$$

We distinguish cases according to $|T|$.

Case 1: $|T| = 0$ (so $x_1, x_2 \notin S$). We claim that $G - S$ is disconnected. Indeed, suppose instead that $G - S$ is connected. Take any two vertices $u, v \in V(H) \setminus S$. Note that $V(H) \setminus \{x_1, x_2\} = V(G) \setminus \{x\}$, and $S \subseteq V(H) \setminus \{x_1, x_2\}$ in this case, so u and v correspond to vertices of $G - S$ as well. Since $G - S$ is connected, there is a u–v path P in $G - S$.

If P avoids x, then P is also a path in $H - S$ (all edges away from the split are unchanged), so u and v are connected in $H - S$.

If P uses x, write P as a sequence of vertices and focus on every subwalk of the form $a - x - b$ where $a, b \in N_G(x)$. In H, the neighbors of x were redistributed so that each old neighbor $a \in N_G(x)$ is adjacent to at least one of x_1 or x_2. Now replace each occurrence of $a - x - b$ in P by:

$$a - x_i - b \quad \text{if } a, b \in N_H(x_i) \setminus \{x_{3-i}\},$$

and otherwise by $a - x_i - x_{3-i} - b$, where x_i is chosen so that $a \in N_H(x_i)$ and x_{3-i} is the other split vertex. This replacement is always possible because $x_1x_2 \in E(H)$ and a (resp. b) is adjacent in H to whichever split vertex it was assigned to. Moreover, since $x_1, x_2 \notin S$, none of the new internal vertices we introduce lies in S, so the modified walk lies in $H - S$. After suppressing repeated vertices if necessary, we obtain a u–v path in $H - S$.

Thus any two vertices of $H - S$ are connected, so $H - S$ is connected, contradicting that S is a vertex cut. Therefore $G - S$ must be disconnected, and S is a vertex cut of G with $|S| \leq k-1$, contradicting that G is k-connected.

Case 2: $|T| = 1$. Without loss of generality, assume $T = \{x_1\}$, i.e. $x_1 \in S$ and $x_2 \notin S$. Define

$$S_G := (S \setminus \{x_1\}) \cup \{x\} \subseteq V(G).$$

Then $|S_G| = |S| \leq k - 1$.

We claim that $G - S_G$ is disconnected. Suppose for contradiction that $G - S_G$ is connected. Consider the subgraph induced by $V(H) \setminus (S \cup \{x_2\})$ inside $H - S$. Its vertex set is exactly $V(G) \setminus S_G$, and all edges among these vertices are the same in G and H (because the only modification from G to H involves x and its incident edges). Hence, since $G - S_G$ is connected, all vertices of $H - S$ other than x_2 lie in a single component of $H - S$.

It remains to show that x_2 also lies in that component. Because $d_H(x_2) \geq k$, the vertex x_2 has at least $k - 1$ neighbors in H other than x_1. But $S \setminus \{x_1\}$ has size at most $k - 2$, so it cannot contain all those neighbors. Therefore there exists a vertex

$$y \in N_H(x_2) \setminus S.$$

In particular, $y \neq x_1$ and $y \neq x_2$, so y is one of the vertices already in the big component of $H - S$, and the edge x_2y survives in $H - S$. Thus x_2 is connected to that big component, so $H - S$ is connected, contradicting that S is a vertex cut.

Hence $G - S_G$ is disconnected, so S_G is a vertex cut of G with $|S_G| \leq k - 1$, again contradicting that G is k-connected.

Case 3: $|T| = 2$ **(so** $x_1, x_2 \in S$**).** Define

$$S_G := (S \setminus \{x_1, x_2\}) \cup \{x\} \subseteq V(G).$$

Then $|S_G| = |S| - 1 \leq k - 2$.

Now observe that $H - S$ is exactly the same graph as $G - S_G$: both have vertex set

$$V(G) \setminus \big((S \setminus \{x_1, x_2\}) \cup \{x\}\big) = V(G) \setminus S_G,$$

and the edge sets on this vertex set coincide (since the only altered edges were those incident to x, and x is removed in $G - S_G$, while x_1, x_2 are removed in $H - S$). Therefore $H - S$ disconnected implies $G - S_G$ disconnected, so S_G is a vertex cut of G of size at most $k - 2$, contradicting k-connectedness of G.

In all cases we obtain a vertex cut of G of size at most $k - 1$, contradicting that G is k-connected. Therefore no such set S exists, and H is k-connected. □

Lemma 20.2. For the hypercube Q_k,

$$\kappa(Q_k) = k.$$

Proof. Write

$$Q_k = Q_{k-1} \square Q_1,$$

which consists of two copies of Q_{k-1}, call them G_1 and G_2, joined by a perfect matching between corresponding vertices.

Since Q_{k-1} is $(k-1)$-connected by the induction hypothesis, each of G_1 and G_2 is $(k-1)$-connected. Let S be a separating set in Q_k. We show that $|S| \geq k$, which proves $\kappa(Q_k) \geq k$. Combined with $\kappa(Q_k) \leq \delta(Q_k) = k$, we obtain $\kappa(Q_k) = k$.

Case 1: S disconnects G_1 or G_2:

Assume S disconnects G_1. Since G_1 is $(k-1)$-connected, removing fewer than $k-1$ vertices cannot disconnect it. Thus S must contain at least $k-1$ vertices of G_1. Furthermore, S must contain at least one vertex from G_2; otherwise, all vertices of G_2–S remain connected inside G_2, and through the matching edges they reconnect the separated parts of $G_1 - S$. Hence in this case,

$$|S| \geq (k-1) + 1 = k.$$

Case 2: $G_1 - S$ and $G_2 - S$ are both connected:

In this case, the only way for $Q_k - S$ to be disconnected is for S to delete all matching edges between the two copies. Each matching edge has one endpoint in G_1 and one in G_2. Thus S must contain the endpoint of every matching edge, so

$$|S| \geq |V(G_1)| = 2^{k-1}.$$

Since $k \geq 2$, we have $2^{k-1} \geq k$, so again $|S| \geq k$.

In all cases, every separating set has size at least k. Thus $\kappa(Q_k) \geq k$, and since $\delta(Q_k) = k$ we obtain

$$\kappa(Q_k) = k.$$

□

Theorem 20.3 (Niu–Zhu (1994), Chiue–Shieh (1999)). If G and H are connected graphs, then

$$\kappa(G \square H) \geq \kappa(G) + \kappa(H).$$

The proof follows the same idea as the lemma and is omitted (can be found in the textbook). This theorem strengthens the lemma, which in particular recovers $\kappa(Q_k) \geq \kappa(Q_{k-1}) + \kappa(K_2) = k$ as a corollary.

20.2 Edge connectivity

Definition 20.5. An *edge cut* of G is a set of edges $F \subseteq E$ such that the graph $G - F$ (obtained by deleting all edges in F) is disconnected.

Definition 20.6. A connected graph G is *k-edge-connected* if it stays connected after the deletion of any set of fewer than k edges; that is, for every $F \subseteq E(G)$ with $|F| \leq k-1$, the graph $G - F$ is connected.

Definition 20.7. The *edge-connectivity* of a connected graph G is

$$\kappa'(G) \; := \; \min\{|F| : \; F \text{ is an edge cut of } G\}.$$

Equivalently, $\kappa'(G)$ is the largest integer k such that G is k-edge-connected.

Definition 20.8 (Cut edges across a partition). For a partition of the vertex set into two nonempty parts S and $\overline{S} := V(G) \setminus S$, write

$$[S, \overline{S}] \ := \ \{uv \in E(G) : \ u \in S, \ v \in \overline{S}\}.$$

These are exactly the edges *crossing* from S to its complement; deleting $[S, \overline{S}]$ isolates S from $\overline{S}$.

$$\delta(S) \ := \ [S, \overline{S}] \ = \ \{uv \in E : \ u \in S, \ v \in \overline{S} := V \setminus S\}.$$

A *bond* is a nonempty cut $\delta(S)$ that is minimal (by inclusion, not “minimum size") as a disconnecting set. This is weaker than being *minimum* (having smallest cardinality among all edge cuts).

Motivation: It is often useful to think of a graph as a communication network: vertices are routers/servers and edges are physical links like cables. A failed hub can wipe out *many* connections at once.

Vertex connectivity measures robustness to these *node failures*. If a graph is k-connected, then deleting any $k-1$ vertices (and all incident edges) still leaves the network connected, meaning there is still a route between every remaining pair of devices. Equivalently, $\kappa(G)$ is the minimum number of hubs you must lose (by failure) before the network splits into separate islands that cannot communicate. Vertex cuts therefore identify the critical choke points: the small set of nodes whose removal fragments the system. A larger $\kappa(G)$ means connectivity is spread out rather than concentrated in a few fragile hubs, so the network is genuinely harder to break.

Edge-connectivity quantifies *robustness to link failures*. If a network is k-edge-connected, then it still has a functioning route between every pair of nodes even after any $k-1$ links disappear. Equivalently, $\kappa'(G)$ is the minimum number of links an adversary (or bad luck) must remove to split the network into disconnected pieces. Cuts of the form $[S, \overline{S}]$ capture the bottlenecks: they are precisely the links that carry *all* communication between two regions of the network. So studying edge cuts and $\kappa'(G)$ is really studying where the network links are fragile, how many redundant links it has, and how hard it is to knock it offline.

Proposition 20.4.

1. Every minimal disconnecting set is an edge cut (i.e. equals $\delta(S)$ for some nontrivial $S \subset V$).
2. If G is connected and $\emptyset \neq S \subsetneq V(G)$, then $\delta(S)$ is a bond if and only if both induced subgraphs $G[S]$ and $G[\overline{S}]$ are connected.

(i) Let $F \subseteq E(G)$ be a minimal disconnecting set, so $G - F$ is disconnected. Let C be the vertex set of one connected component of $G - F$, and put $\overline{C} := V(G) \setminus C$. By definition of component, there are no edges of $G - F$ between C and $\overline{C}$; hence every edge of G with one endpoint in C and the other in $\overline{C}$ must have been removed, i.e.

$$\delta(C) = [C, \overline{C}] \subseteq F.$$

We claim equality holds. If not, choose $e \in F \setminus \delta(C)$. Then both ends of e lie in C or both lie in $\overline{C}$, so removing e is irrelevant to separating C from $\overline{C}$. In particular, $G - (F \setminus \{e\})$ is still disconnected (since C is still isolated from $\overline{C}$), contradicting the minimality of F. Thus $F = \delta(C)$, so F is an edge cut.

(ii) Assume G is connected and fix $\emptyset \neq S \subsetneq V$, writing $\overline{S} := V \setminus S$.

($\Rightarrow$) Suppose $\delta(S)$ is a bond. If $G[S]$ were disconnected, let C be the vertex set of a component of $G[S]$ (with $\emptyset \neq C \subsetneq S$). Because C is a component inside S, there are no edges between C and $S \setminus C$, so any edge leaving C must go to $\overline{S}$. Hence

$$\delta(C) = [C, V \setminus C] = [C, \overline{S}] \subseteq [S, \overline{S}] = \delta(S).$$

Moreover, $\delta(C) \neq \emptyset$: since G is connected, the set C cannot be isolated from $V \setminus C$, so at least one edge leaves C. Therefore $\delta(C)$ is a nonempty disconnecting set properly contained in $\delta(S)$, contradicting that $\delta(S)$ is minimal. Thus $G[S]$ must be connected; by symmetry, $G[\overline{S}]$ is connected as well.

($\Leftarrow$) Now assume $G[S]$ and $G[\overline{S}]$ are both connected. We show $\delta(S)$ is minimal. Let $F \subsetneq$ $attachment$ $\delta(S)$ be a proper subset. Then some crossing edge $uv \in \delta(S) \setminus F$ remains, with $u \in S$ and $v \in \overline{S}$. Since F removes only crossing edges, the induced subgraphs $G[S]$ and $G[\overline{S}]$ are unchanged and remain connected inside $G - F$. The surviving edge uv links these two connected pieces, so $G - F$ is connected. Thus removing any proper subset of $\delta(S)$ cannot disconnect the graph, i.e. $\delta(S)$ is a bond.

This proves both directions.

Theorem 20.5 (Whitney, 1932)**.**

$$\kappa(G) \;\leq\; \kappa'(G) \;\leq\; \delta(G).$$

Proof. Let v be a vertex of minimum degree $\delta(G)$. Delete all edges incident to v. Then v becomes isolated so these $\delta(G)$ edges form an edge cut. Hence

$$\kappa'(G) \leq \delta(G).$$

To prove $\kappa(G) \leq \kappa'(G)$, let $\lambda = \kappa'(G)$, and let $[S, \overline{S}]$ be a smallest edge cut of size λ. We will show that from this *edge* cut we can build a *vertex* cut of size at most $\kappa'(G)$. That will immediately imply $\kappa(G) \leq= \kappa'(G)$. If every vertex of S is adjacent to every vertex of $\overline{S}$, then the cut contains *all* possible edges between the two sides, hence

$$\lambda = |[S, \overline{S}]| = |S|\,|\overline{S}|.$$

In particular, since both S and $\overline{S}$ are nonempty, $|S|\,|\overline{S}| = |S|(|V| - |S|) \geq |V| - 1$ (the product is minimized when one side has size 1). Thus $\kappa'(G) \geq |V(G)| - 1$. But always $\kappa(G) \leq |V(G)| - 1$, so $\kappa(G) \leq \kappa'(G)$ and we are done.

Otherwise, there exist $x \in S$ and $y \in \overline{S}$ with $xy \notin E(G)$. We will construct a set of vertices T with $|T| \leq \lambda$ such that removing T separates x from y.

Define T by "picking one endpoint" from each cut edge, carefully arranged so that every x–y path hits T:

- For each cut edge incident to x (i.e. edges xv with $v \in \overline{S}$), put the *other* endpoint v into T.
- For every remaining cut edge uv with $u \in S$ and $v \in \overline{S}$ and $u \neq x$, put the endpoint in S (namely u) into T.

So each cut edge contributes *at most one* vertex to T, and therefore

$$|T| \le \|[S, \overline{S}]\| = \kappa'(G).$$

Why does T separate x and y? Consider any path P from x to y in G. Since $x \in S$ and $y \in \overline{S}$, the path must at some point cross from S to $\overline{S}$. Let uv be the *first* edge of P that has one endpoint in S and the next vertex in $\overline{S}$. Then $uv \in [S, \overline{S}]$ is a cut edge.

There are two possibilities:

- If $u = x$, then uv is a cut edge incident to x, and by construction we placed the other endpoint $v \in \overline{S}$ into T. So P hits T at the vertex v.
- If $u \ne x$, then uv is one of the "remaining" cut edges, and by construction we placed its S-endpoint u into T. So P hits T at the vertex u.

In either case, *every* x–y path meets T. Equivalently, in the graph $G - T$ there is no path from x to y, so T is a vertex separator for x and y. □

Proposition 20.6. If $S \subseteq V(G)$, then

$$\left|[S, \overline{S}]\right| = \sum_{v \in S} d(v) - 2\,|E(G[S])|.$$

Proof. The sum $\sum_{v \in S} d(v)$ counts every edge with both endpoints in S twice, and it counts each edge in $[S, \overline{S}]$ once. Subtracting $2|E(G[S])|$ leaves exactly the number of edges joining S and $\overline{S}$. □

Proposition 20.7. Let G be a simple graph and $S \ne \varnothing$. If $[S, \overline{S}] < \delta(G)$, then $|S| > \delta(G)$.

Proof. Since $\delta(G) > [S, \overline{S}]$, we have

$$\delta(G) > \sum_{v \in S} d(v) - 2e(G[S]).$$

Because $e(G[S]) \le |S|(|S| - 1)/2$, we obtain

$$\sum_{v \in S} d(v) - 2e(G[S]) \ge |S|\delta(G) - |S|(|S| - 1).$$

Hence

$$\delta(G) > |S|\big(\delta(G) - (|S| - 1)\big).$$

Thus

$$0 > (|S| - 1)\big(\delta(G) - |S|\big).$$

Since $|S| - 1 \ge 0$, the inequality implies

$$0 > \delta(G) - |S|,$$

and therefore $|S| > \delta(G)$. □

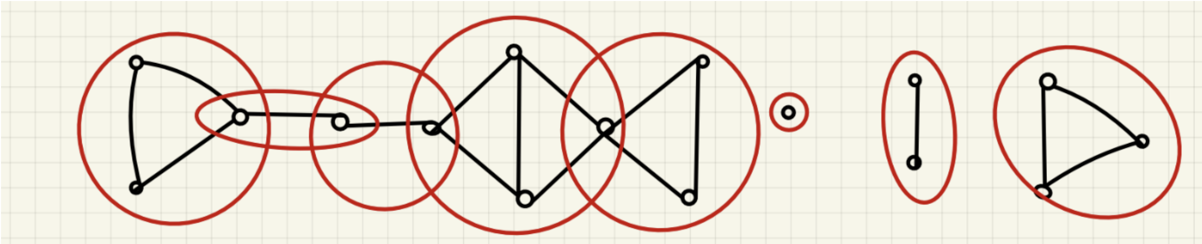
Theorem 20.8. If $\operatorname{diam}(G) = 2$, then $\kappa'(G) = \delta(G)$.

Proof. We want to show that $\kappa'(G) \geq \delta(G)$. Suppose for a contradiction that there is a smallest edge cut of size

$$[S, \overline{S}] \; < \; \delta(G).$$

Pick $x \in S$ and $y \in \overline{S}$ such that all edges joining S and $\overline{S}$ lie between these two sides. Since $d(x), d(y) \geq \delta(G)$ and fewer than $\delta(G)$ edges join S and $\overline{S}$, both x and y have neighbors on their own side. Because $\operatorname{diam}(G) = 2$, the distance between x and y is at most 2. But any x–y path must use an edge between S and $\overline{S}$, and there are fewer than $\delta(G)$ such edges, contradicting the assumption that $d(x), d(y) \geq \delta(G)$.

Hence $\kappa'(G) \geq \delta(G)$, and since always $\kappa'(G) \leq \delta(G)$, we obtain $\kappa'(G) = \delta(G)$. □



20.3 Block decomposition

When we study the *structure* of a graph, the first coarse decomposition is into connected components. Inside a connected component, however, connectivity can still be "fragile": there may be a single vertex whose removal breaks the component apart. Such vertices are the *cut-vertices*.

Blocks isolate the parts that have *no* such vulnerability. Informally, a block is a maximal region of the graph that cannot be separated by deleting a single vertex. Thus blocks are the natural "2-connected pieces" of a graph, and a connected graph can be viewed as blocks stitched together at cut-vertices (this leads to the block–cutvertex tree).

Definition 20.9 (Block)**.** A *block* of a graph G is a maximal connected subgraph of G that has no cut-vertex (with cut-vertices understood *inside that subgraph*).

Remark 20.2. With this definition, blocks include the "degenerate" cases: an isolated vertex forms a block, and a bridge edge uv forms a block. In general, distinct blocks intersect in at most one vertex, and any common vertex must be a cut-vertex of G.

Proposition 20.9 (Basic properties of blocks)**.** Let G be a graph, and let a *block* mean a maximal connected subgraph of G with no cut-vertex (with cut-vertices understood inside the subgraph). Then:

1. If an edge e lies on a cycle, then e is not a block. In fact, the unique block containing e contains the entire cycle.
2. An edge uv is a block (i.e. $\{u, v\}$ with the edge uv) if and only if uv is a cut-edge (bridge) of G.
3. If T is a tree with $|V(T)| \geq 2$, then the blocks of T are exactly its edges.
4. Every block with at least 3 vertices is 2-connected.
5. Any two distinct blocks intersect in at most one vertex.
6. Only cut-vertices can lie in more than one block.
7. Every edge of G lies in exactly one block.

Proof. **(i)** Let C be a cycle containing e. The cycle C is connected and has no cut-vertex (removing any one vertex from a cycle still leaves a path), so C is contained in some block B by maximality. In particular, $e \subseteq B$, so e cannot itself be a block unless C had only 2 vertices, which is impossible in a simple graph. Moreover, since B contains C, the block containing e contains the entire cycle.

(ii) If uv is a cut-edge, then in $G - uv$ the vertices u and v lie in different components. Any connected subgraph of G that contains uv and also contains any vertex besides u, v must include vertices from at least one side of the separation, and then removing u or v disconnects that subgraph (it isolates the other side through the bridge). Hence the only connected subgraph containing uv with no cut-vertex is the edge itself, so uv is a block.

Conversely, if uv is a block and uv were *not* a cut-edge, then uv lies on a cycle. By (i), an edge on a cycle cannot be a block. So uv must be a cut-edge.

(iii) In a tree every edge is a cut-edge. Thus by (ii) every edge is a block. There are no other blocks when $|V(T)| \geq 2$.

(iv) A connected graph on at least 3 vertices is 2-connected if and only if it has no cut-vertex. So a block with at least 3 vertices is 2-connected by definition.

(v) Let $B_1 \neq B_2$ be blocks. Suppose for contradiction that $B_1 \cap B_2$ contains two distinct vertices $a \neq b$. Then $B_1 \cup B_2$ is connected (they share vertices) and has no cut-vertex: removing any vertex w cannot disconnect $B_1 \cup B_2$ because $B_1 - w$ and $B_2 - w$ remain connected (each block has no cut-vertex), and even if $w = a$ (or $w = b$) the other common vertex b (or a) still links the two pieces. Thus $B_1 \cup B_2$ is a larger connected subgraph with no cut-vertex, contradicting maximality of B_1 and B_2. Hence $|B_1 \cap B_2| \leq 1$.

(vii) (Existence.) Fix an edge $e \in E(G)$. Consider the collection of connected subgraphs of G that contain e and have no cut-vertex. Since G is finite, there is one with a maximum number of vertices; call it B. By construction, B is connected, has no cut-vertex, contains e, and is maximal with these properties, hence B is a block. So every edge lies in *at least* one block.

(Uniqueness.) If an edge $e = uv$ lay in two distinct blocks B_1 and B_2, then $u, v \in B_1 \cap B_2$, so $|B_1 \cap B_2| \geq 2$, contradicting (v). Hence every edge lies in *exactly* one block.

(vi) Suppose a vertex v lies in two distinct blocks $B_1 \neq B_2$. Then B_1 and B_2 cannot both be the single vertex $\{v\}$, so in each B_i there is an edge incident to v. Choose $va \in E(B_1)$ and $vb \in E(B_2)$ with $a \neq v$ and $b \neq v$. If $G - v$ had a path P from a to b, then $v - a - P - b - v$ would be a cycle containing both edges va and vb. By (i), edges on a cycle belong to a block containing that cycle, and by (vii) an edge lies in a *unique* block; this would force va and vb to lie in the same block, contradicting $B_1 \neq B_2$. Therefore no such path exists, so a and b lie in different components of $G - v$, meaning $G - v$ is disconnected. Hence v is a cut-vertex. □

21 Properties of k-connected graphs

Definition 21.1. A digraph $\vec{G}$ is strongly connected if for all $x, y \in V(\vec{G})$ there exists a directed x, y-path.

Definition 21.2.

$$\kappa(\vec{G}) = \max\{k : \vec{G} \text{ is } k\text{-strongly connected}\},$$

that is, removing any $k - 1$ vertices leaves $\vec{G}$ strongly connected.

Remark 21.1. For a directed cut, $[S, \overline{S}]$ and $[\overline{S}, S]$ need not be equal.

Definition 21.3. A digraph $\vec{G}$ is k-edge-connected if every directed edge cut has size at least k.

21.1 Menger's Theorem

Definition 21.4. For vertices x, y in a graph G with $xy \notin E(G)$, an (x, y)-separating set is a set

$$S \subseteq V(G) \setminus \{x, y\}$$

such that $G - S$ has no (x, y)-path.

Definition 21.5. Let $\kappa(x, y)$ be the minimum size of an (x, y)-separating set in G.

Definition 21.6. A family of (x, y)-paths is independent if the paths share no internal vertices. Let $\lambda(x, y)$ be the maximum number of pairwise independent (x, y)-paths.

Remark 21.2. It is always true that

$$\kappa(x, y) \geq \lambda(x, y),$$

since removing one vertex from each path in an independent family separates x and y.

Theorem 21.1 (Menger, 1927)**.** For all vertices x, y in a graph G with $xy \notin E(G)$,

$$\kappa(x, y) = \lambda(x, y).$$

We will prove a stronger version and obtain this as a corollary.

Definition 21.7. Let $X, Y \subseteq V(G)$ be two disjoint vertex sets. An (X, Y)-path is a path in G that starts in X, ends in Y, and has all internal vertices in $V(G) \setminus (X \cup Y)$.

Definition 21.8. A strict (X, Y)-path is an (X, Y)-path whose only vertices in $X \cup Y$ are its endpoints.

Definition 21.9. An (X, Y)-cut is a set $Z \subseteq V(G)$ such that $G - Z$ contains no (X, Y)-path.

Definition 21.10. An (X, Y)-link is a collection of pairwise internally disjoint (X, Y)-paths.

Theorem 21.2 (Pym, 1969)**.** Let G be a graph and let $X, Y \subseteq V(G)$ be disjoint. Then the minimum size of an (X, Y)-cut equals the maximum size of an (X, Y)-link.

Proof. Let

$$\lambda := \min\{|Z| : Z \text{ is an } (X, Y)\text{-cut}\}, \qquad \nu := \max\{\text{size of an } (X, Y)\text{-link}\}.$$

Easy direction: $\nu \leq \lambda$**.** If $\mathcal{P}$ is an (X, Y)-link of size r, then its paths are vertex-disjoint. Any cut must meet each path of $\mathcal{P}$ in at least one vertex, and those vertices must be distinct. Hence every barrier has size at least r, hence $\nu \leq \lambda$.

Hard direction: $\lambda \leq \nu$**.** We prove $\nu \geq \lambda$ by induction on $|V(G)| + |E(G)|$. The statement is trivial for graphs with very few vertices/edges, so assume $|V(G)| + |E(G)|$ is large and the theorem holds for all smaller graphs/digraphs.

If G has no (X, Y)-path, then X itself is an (X, Y)-barrier, so $\lambda = |X|$ and $\nu = |X \cap Y|$ (because the only (X, Y)-paths have length 0). In particular $\nu \geq \lambda$ holds trivially in this degenerate case. So assume (X, Y)-paths exist, hence $\lambda \geq 1$.

Fix a minimum (X, Y)-barrier Z with $|Z| = \lambda$. We split into two cases.

Case 1: there is a minimum barrier Z with $Z \neq X$ and $Z \neq Y$. Because $G - Z$ has no (X, Y)-path, every (X, Y)-path must visit Z. Consider paths from X *to* Z and from Z *to* Y, avoiding Z until the endpoint:

- Let G_1 be the subgraph of G induced by all vertices that lie on some (X, Z)-path whose internal vertices avoid Z.
- Let G_2 be the subgraph of G induced by all vertices that lie on some (Z, Y)-path whose internal vertices avoid Z.

Claim 1: $V(G_1) \cap V(G_2) = Z$. Certainly $Z \subseteq V(G_1) \cap V(G_2)$ (every $z \in Z$ is on a trivial z–z path). Conversely, suppose $v \notin Z$ lies in both $V(G_1)$ and $V(G_2)$. Then there is an (X, Z)-path P_1 that reaches some $z \in Z$ and passes through v without meeting Z earlier, and there is a (Z, Y)-path P_2 starting at some $z' \in Z$ and passing through v without meeting Z again. By following P_1 from X to v and then P_2 from v to Y, we obtain an (X, Y)-path that avoids Z, contradicting that Z is a barrier. Thus $V(G_1) \cap V(G_2) = Z$.

In particular, G_1 and G_2 are both *smaller* than G (since $Z \neq X, Y$ and $\lambda \geq 1$), so the induction hypothesis applies to each.

Claim 2: the minimum size of an (X, Z)*-barrier in* G_1 *is exactly* λ*.* First, Z itself is an (X, Z)-barrier in G_1, so the minimum is at most λ. On the other hand, any (X, Z)-barrier in G_1 is automatically an (X, Y)-barrier in G: every (X, Y)-path hits Z, and its initial segment from X to the first vertex of Z is a Z-avoiding (X, Z)-path, hence lives in G_1. So blocking all such (X, Z)-paths blocks all (X, Y)-paths. Since λ is the minimum size of an (X, Y)-barrier, no (X, Z)-barrier in G_1 can have size $< \lambda$. Therefore the minimum size is λ.

By induction applied to G_1, there exists an (X, Z)-link in G_1 of size λ. Because the paths are vertex-disjoint and there are only $|Z| = \lambda$ possible endpoints in Z, this link uses *every* vertex of Z as an endpoint exactly once.

Similarly, applying the same argument to G_2 shows that G_2 contains a (Z, Y)-link of size λ, again using every $z \in Z$ exactly once as an endpoint.

Now for each $z \in Z$, concatenate the unique X–z path from the first link with the unique z–Y path from the second link. These concatenated paths are pairwise vertex-disjoint: within each link they are disjoint, and outside Z the two links live in $G_1 \setminus Z$ and $G_2 \setminus Z$, which are disjoint by Claim 1. Thus we obtain an (X, Y)-link of size λ, so $\nu \geq \lambda$.

Case 2: every minimum barrier is X and/or Y.

By symmetry we may assume X is a minimum barrier, so $|X| = \lambda$.

If $X \subseteq Y$, then for each $x \in X$ the length-0 path (x) is an (X, Y)-path. These $|X| = \lambda$ trivial paths are vertex-disjoint, giving an (X, Y)-link of size λ. Hence $\nu \geq \lambda$.

Otherwise pick $x \in X \setminus Y$. Since X is a *minimum* barrier, $X \setminus \{x\}$ is not a barrier. So $G - (X \setminus \{x\})$ contains an (X, Y)-path, and that path must start at x and immediately leave X along some edge (or arc) xw with $w \notin X$. Fix such an edge $e := xw$, and consider the smaller graph $G' := G - e$.

Let Z' be a minimum (X, Y)-barrier in G'. There are two subcases.

Subcase 2a: G' *has an* (X, Y)*-link of size* λ*.* Then the same link also exists in G (adding an edge cannot destroy existing paths), so $\nu \geq \lambda$.

Subcase 2b: G' *has no* (X, Y)*-link of size* λ*.* Then the maximum link size in G' is $< \lambda$, so by the induction hypothesis (applied to G'),

$$|Z'| = \min\{\text{barrier size in } G'\} = \max\{\text{link size in } G'\} < \lambda.$$

In particular, $|Z'| \leq \lambda - 1$.

Now Z' cannot be an (X, Y)-barrier in G (otherwise $\lambda \leq |Z'| < \lambda$), so $G - Z'$ contains an (X, Y)-path. But $G' - Z'$ contains no (X, Y)-path (since Z' *is* a barrier in G'). Therefore *every* (X, Y)-path in $G - Z'$ must use the deleted edge $e = xw$. Consequently, every (X, Y)-path meets $Z' \cup \{x\}$ and also meets $Z' \cup \{w\}$. Thus both sets are (X, Y)-barriers in G.

Since λ is the minimum barrier size in G,

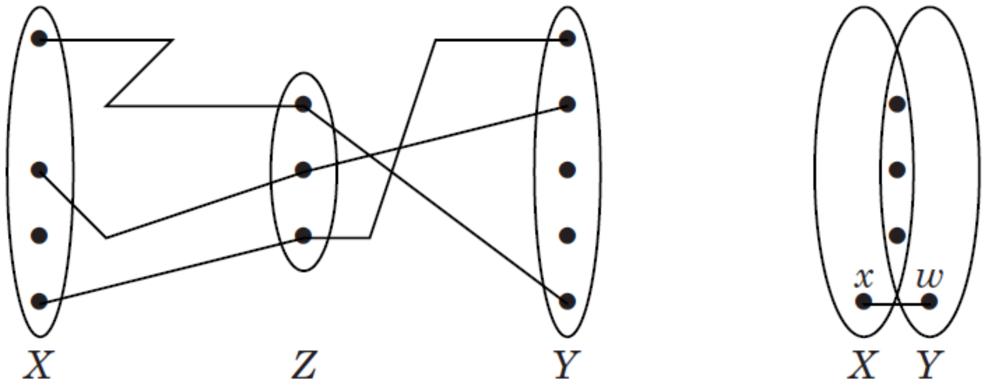
$$\lambda \leq |Z' \cup \{x\}| = |Z'| + 1 \leq \lambda,$$

so $|Z'| = \lambda - 1$ and both $Z' \cup \{x\}$ and $Z' \cup \{w\}$ are *minimum* barriers of size λ. By the hypothesis of Case 2, the only minimum barriers are X and/or Y. Because $x \in X \setminus Y$, the barrier $Z' \cup \{x\}$ cannot equal Y, hence

$$Z' \cup \{x\} = X.$$

Also $w \notin X$ by construction, so $Z' \cup \{w\} \neq X$, and therefore

$$Z' \cup \{w\} = Y.$$



It follows that

$$Z' = X \setminus \{x\} = Y \setminus \{w\},$$

so every vertex of Z' lies in $X \cap Y$.

Finally, we build an (X, Y)-link of size λ in G: take the $\lambda - 1$ trivial paths (z) for each $z \in Z'$ (these are (X, Y)-paths since $z \in X \cap Y$), together with the length-1 path xw. These λ paths are vertex-disjoint (their vertices are all distinct), so $\nu \geq \lambda$.

In all cases we have shown $\nu \geq \lambda$. Together with $\nu \leq \lambda$, we conclude $\nu = \lambda$. □

Corollary 21.3 (Menger's Theorem)**.** Let G be a graph and let $x \neq y$ be vertices. Then the minimum size of an x–y vertex separator equals the maximum number of pairwise internally vertex-disjoint x–y paths from Pym's Theorem with $X = \{x\}, Y = \{y\}$.

Definition 21.11. Let G be a graph (or digraph) and let $x, y \in V(G)$. Define $\kappa'(x, y)$ to be the minimum size of an x–y *edge cut*, i.e.

$$\kappa'(x, y) \; := \; \min\{|F| : \; F \subseteq E(G) \text{ and } y \text{ is not reachable from } x \text{ in } G - F\}.$$

Definition 21.12. For vertices x, y in a graph G, let $\lambda'(x, y)$ be the maximum size of a set of pairwise edge-disjoint (x, y)-paths.

Definition 21.13 (Line graph)**.** Let G be a (multi)graph. The *line graph* $L(G)$ is the graph with vertex set $E(G)$, where two distinct vertices $e, f \in E(G)$ are adjacent in $L(G)$ exactly when the corresponding edges share an endpoint in G.

Definition 21.14 (Line digraph)**.** Let D be a digraph. The *line digraph* $L(D)$ has vertex set $E(D)$, and two arcs $e = (u, v)$ and $f = (v, w)$ are adjacent (i.e. there is an arc from e to f) precisely when they are consecutive directed edges in D.

Theorem 21.4 (Menger's theorem, edge form)**.** Let G be a graph (or digraph) and let $x, y \in V(G)$ with $x \neq y$. Then
$$\kappa'(x, y) = \lambda'(x, y).$$

Proof. The inequality $\lambda'(x, y) \leq \kappa'(x, y)$ is immediate: if F is any x–y edge cut and $\mathcal{P}$ is a set of pairwise edge-disjoint x–y paths, then each path in $\mathcal{P}$ must use at least one edge of F, and since the paths are edge-disjoint those edges of F must be distinct. Hence $|F| \geq |\mathcal{P}|$, and taking minima/maxima gives $\lambda'(x, y) \leq \kappa'(x, y)$.

For the reverse inequality, reduce to the vertex version of Menger by subdividing edges. Form a new graph (or digraph) G^* by subdividing every edge $e = uv$ once: replace e by a length-2 path $u - s_e - v$, where s_e is a new vertex unique to e (and in the directed case, replace (u, v) by (u, s_e) and (s_e, v)). Then:

- An x–y path in G corresponds to an x–y path in G^* whose internal vertices among the new subdivision vertices are exactly the s_e for edges used by the original path.
- Two x–y paths in G are edge-disjoint if and only if the corresponding x–y paths in G^* are internally vertex-disjoint (because distinct edges correspond to distinct subdivision vertices).
- A set of edges F separates x from y in G if and only if the set of subdivision vertices $\{s_e : e \in F\}$ separates x from y in G^*.

Therefore
$$\lambda'(x, y) = \lambda_{G^*}(x, y) \qquad \text{and} \qquad \kappa'(x, y) = \kappa_{G^*}(x, y),$$
where the right-hand sides are the *vertex*-disjoint-path / vertex-separator parameters in G^*. Applying the vertex form of Menger to G^* yields $\lambda_{G^*}(x, y) = \kappa_{G^*}(x, y)$, and hence $\lambda'(x, y) = \kappa'(x, y)$ as desired. □

Corollary 21.5 (Global connectivity via local disjoint paths)**.** Let G be a graph (or digraph), and let $\lambda(x, y)$ denote the maximum number of pairwise internally vertex-disjoint x–y paths.

1. G is k-connected $\iff$ $\lambda(x, y) \geq k$ for all distinct $x, y \in V(G)$.
2. G is k-edge-connected $\iff$ $\lambda'(x, y) \geq k$ for all distinct $x, y \in V(G)$.

Proof. **(2)** Recall the standard identity
$$\kappa'(G) = \min_{\substack{x,y \in V(G) \\ x \neq y}} \kappa'(x, y).$$

Thus G is k-edge-connected iff $\kappa'(x, y) \geq k$ for all $x \neq y$. By Theorem 21.4, $\kappa'(x, y) = \lambda'(x, y)$ for all $x \neq y$, giving the equivalence.

(1) Similarly,
$$\kappa(G) = \min_{\substack{x,y \in V(G) \\ x \neq y}} \kappa(x, y),$$
where $\kappa(x, y)$ is the minimum size of an x–y vertex separator. By the vertex form of Menger, $\kappa(x, y) = \lambda(x, y)$ for all $x \neq y$. Hence $\kappa(G) \geq k$ iff $\lambda(x, y) \geq k$ for all $x \neq y$, i.e. G is k-connected iff the stated local path condition holds. □

21.2 Network flows and Max-Flow Min-Cut Theorem

Menger's theorems say, roughly, that "many disjoint routes" between two terminals exist if and only if you must delete many vertices/edges to separate the terminals. Network flows package this idea into an optimization problem: instead of asking for disjoint paths directly, we send *flow* through a capacitated network and compare it to the cheapest way to block the flow.

Definition 21.15 (Flow network). A *flow network* is a digraph $D = (V, A)$ together with a *capacity* function $c : A \to \mathbb{R}_{\geq 0}$ and two distinct vertices $s, t \in V$ called the *source* and *sink*.

Definition 21.16 (Feasible flow and its value). A *flow* is a function $f : A \to \mathbb{R}_{\geq 0}$ such that:

- **(Capacity constraints)** $0 \leq f(a) \leq c(a)$ for every arc $a \in A$.
- **(Flow conservation)** For every $v \in V \setminus \{s, t\}$,

$$\sum_{(u,v)\in A} f(u,v) = \sum_{(v,w)\in A} f(v,w).$$

The *value* of f is the net flow out of s,

$$|f| := \sum_{(s,w)\in A} f(s,w) - \sum_{(u,s)\in A} f(u,s),$$

which equals the net flow into t by conservation.

Definition 21.17 (s–t cut and its capacity). An s–t *cut* is a partition $(S, \overline{S})$ of V with $s \in S$ and $t \in \overline{S}$. Its *capacity* is

$$c(S, \overline{S}) := \sum_{\substack{(u,v)\in A \\ u\in S,\ v\in \overline{S}}} c(u,v),$$

the total capacity of arcs leaving S.

Theorem 21.6 (Max-Flow Min-Cut). In any flow network,

$$\max\{|f| : f \text{ is a feasible flow}\} = \min\{c(S, \overline{S}) : (S, \overline{S}) \text{ is an } s\text{–}t \text{ cut}\}.$$

How this recovers Menger: Let G be an undirected graph and fix distinct vertices x, y. Turn G into a flow network by replacing each undirected edge $\{u, v\}$ with two opposite arcs (u, v) and (v, u), and assign capacity 1 to every arc. Take $s = x$ and $t = y$.

- Any set of k edge-disjoint x–y paths gives a flow of value k: send one unit of flow along each path. (With unit capacities, edge-disjointness ensures no edge is asked to carry more than 1.)
- Any x–y edge cut F gives an s–t cut of capacity $|F|$: choose S to be the vertices reachable from x in $G - F$; then every arc from S to $\overline{S}$ corresponds to an edge of F, and conversely.

Thus the maximum flow value equals the maximum number of edge-disjoint x–y paths, and the minimum cut capacity equals the minimum size of an x–y edge cut. By Theorem 21.6 these are equal, which is exactly the edge form of Menger:

$$\lambda'(x, y) = \kappa'(x, y).$$

Vertex version: Replace each vertex $v \notin \{x, y\}$ by two vertices v^{in} and v^{out} joined by an arc $v^{\text{in}}v^{\text{out}}$ of capacity 1, and redirect every original arc entering v to v^{in} and every original arc leaving v from v^{out}. Give all redirected arcs capacity ∞ (or a sufficiently large number).

Then sending one unit of flow through $v^{\text{in}}v^{\text{out}}$ "uses up" the capacity-1 vertex v, so integral flows correspond to collections of internally vertex-disjoint x–y paths, and minimum cuts correspond to minimum x–y vertex separators. Applying Max-Flow Min-Cut in this transformed network yields Menger's theorem in its vertex form.

Menger's theorems are the unit-capacity, disjoint-path special cases of Max-Flow Min-Cut. Flows generalize them by allowing arbitrary capacities (not just 0/1), fractional routing, and weighted "cost to destroy" bottlenecks, which is exactly why they became the framework of modern network design.

21.3 The Ford–Fulkerson algorithm

The max-flow min-cut theorem tells us the optimum flow value equals the minimum cut capacity. Ford–Fulkerson is the basic method that actually *finds* a maximum flow by repeatedly routing more flow along an *augmenting path* in a residual network.

Definition 21.18 (Residual network). Let $(D = (V, A), c, s, t)$ be a flow network and let f be a feasible flow. The *residual capacity* of an arc $(u, v) \in A$ is $c(u, v) - f(u, v)$. The *residual network* D_f has vertex set V and contains:

- a *forward* arc (u, v) with capacity $c(u, v) - f(u, v)$ whenever $c(u, v) - f(u, v) > 0$;
- a *backward* arc (v, u) with capacity $f(u, v)$ whenever $f(u, v) > 0$.

Definition 21.19 (Augmenting path). An *augmenting path* (with respect to f) is a directed s–t path in the residual network D_f. Its *bottleneck* (or residual capacity) is

$$\Delta \;:=\; \min\{c_f(e) : e \text{ is an arc on the path}\},$$

where c_f denotes residual capacities.

The Ford-Fulkerson algorithm runs as follows:

1. Initialize $f \equiv 0$.
2. While there exists an augmenting path P from s to t in D_f:
   (a) Let Δ be the bottleneck residual capacity of P.
   (b) For each arc on P:
      - if the arc is a forward arc (u, v) (original direction), increase $f(u, v)$ by Δ;
      - if the arc is a backward arc (v, u) (undoing flow), decrease $f(u, v)$ by Δ.

(c) Update the residual network and repeat.

3. Output f.

Theorem 21.7 (Correctness of Ford–Fulkerson)**.** If all capacities are integers, Ford–Fulkerson terminates and outputs a maximum s–t flow. Moreover, when it terminates, the set S of vertices reachable from s in the residual network D_f defines a minimum cut $(S, \overline{S})$.

Proof. Let f be the current flow and let P be an augmenting path with bottleneck Δ. Augmenting by Δ preserves feasibility: on forward arcs we do not exceed capacity because $\Delta \le c(u,v) - f(u,v)$, and on backward arcs we do not make flow negative because $\Delta \le f(u,v)$. Flow conservation is preserved at internal vertices of P because we add Δ to exactly one incoming/outgoing arc in the residual sense, so net flow at each internal vertex remains 0. Thus each augmentation produces a feasible flow whose value increases by $\Delta > 0$.

If capacities are integers, then every residual capacity is an integer, so each Δ is a positive integer. Hence the flow value strictly increases by at least 1 each iteration. Since the flow value is always at most the total capacity leaving s, only finitely many augmentations are possible, so the algorithm terminates.

Now suppose the algorithm terminates at flow f, so there is *no* augmenting path in D_f. Let S be the set of vertices reachable from s in D_f. Then $t \notin S$ by assumption, so $(S, \overline{S})$ is an s–t cut.

We claim the value of f equals the capacity of this cut. Consider any arc (u,v) of the original network with $u \in S$ and $v \in \overline{S}$. If $f(u,v) < c(u,v)$, then the forward residual arc (u,v) would have positive residual capacity, so v would be reachable from s, contradicting $v \in \overline{S}$. Hence every such arc is *saturated*: $f(u,v) = c(u,v)$.

Similarly, for any original arc (u,v) with $u \in \overline{S}$ and $v \in S$, if $f(u,v) > 0$ then the backward residual arc (v,u) would exist with positive residual capacity, making u reachable from s, again a contradiction. Thus every arc entering S from $\overline{S}$ carries *zero* flow: $f(u,v) = 0$.

Therefore the net flow crossing from S to $\overline{S}$ equals

$$\sum_{\substack{(u,v)\in A \\ u\in S,\ v\in \overline{S}}} f(u,v) \; - \sum_{\substack{(u,v)\in A \\ u\in \overline{S},\ v\in S}} f(u,v) \; = \sum_{\substack{(u,v)\in A \\ u\in S,\ v\in \overline{S}}} c(u,v) \; = \; c(S,\overline{S}).$$

But the left-hand side is exactly $|f|$ (the flow value), by flow conservation inside S. Hence $|f| = c(S, \overline{S})$.

Finally, for any feasible flow g and any cut $(S, \overline{S})$ we always have $|g| \le c(S, \overline{S})$ (the cut is an upper bound on flow). Thus $|f| = c(S, \overline{S})$ implies f is a maximum flow and $(S, \overline{S})$ is a minimum cut. □

Remark 21.3 (Complexity and a standard refinement). Ford–Fulkerson depends on how augmenting paths are chosen. With irrational capacities it may not terminate. With integer capacities it terminates, but the number of iterations can be large. Choosing augmenting paths by BFS in the residual network gives the Edmonds–Karp algorithm, which runs in polynomial time $O(|V||E|^2)$.

Bipartite matching via max flow Let $G = (L \cup R, E)$ be bipartite. Build a flow network by adding a source s and sink t, directing edges left-to-right, and giving unit capacities:

$$s \to \ell \ (1) \quad (\ell \in L), \qquad \ell \to r \ (1) \quad (\ell r \in E), \qquad r \to t \ (1) \quad (r \in R).$$

Because all capacities are 1, any integral s–t flow is a collection of edge-disjoint paths $s \to \ell \to r \to t$. Reading off the middle edges gives a matching

$$M_f := \{\ell r \in E : \ f(\ell, r) = 1\}, \qquad |M_f| = |f|.$$

Conversely, any matching M yields a flow of value $|M|$ by sending one unit along $s \to \ell \to r \to t$ for each $\ell r \in M$. Hence maximum matching size equals maximum flow value, and Ford–Fulkerson becomes exactly the usual augmenting-path algorithm for matchings.

21.4 Expansion and Fan Lemma

Lemma 21.8 (Expansion Lemma)**.** Let $k \geq 1$ and let G be a k-connected graph. Let G' be obtained from G by adding a new vertex y adjacent to at least k vertices of G. Then G' is k-connected.

Proof. We must show that deleting any set of at most $k-1$ vertices leaves G' connected.

Let $S \subseteq V(G')$ with $|S| \leq k-1$. We prove that $G' - S$ is connected.

Case 1: $y \in S$**.** Then

$$G' - S \ = \ G - (S \setminus \{y\}).$$

Since $|S \setminus \{y\}| \leq k-2 < k$ and G is k-connected, the graph $G - (S \setminus \{y\})$ is connected. Hence $G' - S$ is connected.

Case 2: $y \notin S$**.** Because y has at least k neighbors in G and S contains at most $k-1$ vertices, S cannot contain all neighbors of y. Thus we may choose a neighbor

$$u \in N_{G'}(y) \setminus S.$$

Consider the subgraph induced by the original vertices $V(G) \setminus S$. This is exactly the graph $G - (S \cap V(G))$, and since $|S \cap V(G)| \leq |S| \leq k-1$, k-connectivity of G implies that $G - (S \cap V(G))$ is connected.

Therefore, for every vertex $v \in V(G) \setminus S$, there is a path from u to v in $G - (S \cap V(G)) \subseteq G' - S$. Appending the edge yu shows that y is connected to every vertex of $V(G) \setminus S$ inside $G' - S$. Hence $G' - S$ is connected.

In all cases, deleting at most $k-1$ vertices does not disconnect G'. Therefore G' is k-connected. □

Definition 21.20 (x, U-fan)**.** Let G be a graph, let $x \in V(G)$, and let $U \subseteq V(G)$ with $x \notin U$. An *x, U-fan of size k* is a family of paths $P_1, \ldots, P_k$ such that:

1. each P_i is an x–u_i path for some (not necessarily distinct) $u_i \in U$;
2. the paths are *internally vertex-disjoint*: for $i \neq j$,

$$\big(V(P_i) \setminus \{x, u_i\}\big) \ \cap \ \big(V(P_j) \setminus \{x, u_j\}\big) = \emptyset;$$

3. each P_i meets U only at its endpoint u_i (equivalently, $V(P_i) \cap U = \{u_i\}$).

Equivalently: the paths all start at x, end in U, and are pairwise disjoint except for the common start vertex x.

Lemma 21.9 (Fan Lemma; Dirac (1960)). Let $k \geq 1$ and let G be a graph with $|V(G)| \geq k+1$. Then G is k-connected if and only if for every vertex $x \in V(G)$ and every set $U \subseteq V(G)$ with $x \notin U$ and $|U| \geq k$, there exists an x, U-fan of size k.

Proof. **(⇒)** Assume G is k-connected. Fix $x \in V(G)$ and $U \subseteq V(G)$ with $x \notin U$ and $|U| \geq k$. Choose any subset $U_0 \subseteq U$ with $|U_0| = k$.

Form a new graph G' from G by adding a new vertex y adjacent to every vertex of U_0. By Lemma 21.8 (Expansion Lemma), G' is k-connected.

Apply Menger's theorem (vertex form) in G' to the vertices x and y. Since G' is k-connected, every x–y separator has size at least k, hence Menger yields k pairwise internally vertex-disjoint x–y paths $Q_1, \ldots, Q_k$ in G'.

Each Q_i must enter y through one of its neighbors, and $N_{G'}(y) = U_0$. Let $u_i \in U_0$ be the neighbor of y used by Q_i, and let P_i be the subpath of Q_i from x to the *first* vertex of U_0 encountered along Q_i (which is necessarily u_i).

Then P_i is an x–u_i path in the original graph G (it does not use y), and by construction it meets U_0 only at its endpoint u_i. Moreover, because the Q_i are internally disjoint, the truncated paths P_i are still internally disjoint and share only the start vertex x. Thus $P_1, \ldots, P_k$ form an x, U_0-fan of size k, and since $U_0 \subseteq U$, they are also an x, U-fan of size k.

(⇐) Assume the fan condition holds for all x and U with $|U| \geq k$. Suppose for contradiction that G is not k-connected. Then there exists a vertex cut $S \subseteq V(G)$ with $|S| \leq k-1$ such that $G - S$ is disconnected. Choose vertices x and z in distinct components of $G - S$.

Let $U := S \cup \{z\}$. Then $x \notin U$ and

$$|U| = |S| + 1 \leq k.$$

If $|U| < k$, enlarge U by adding arbitrary vertices of $V(G) \setminus (\{x\} \cup U)$ until $|U| = k$; this is possible since $|V(G)| \geq k+1$. Call the resulting set still U. Then $x \notin U$ and $|U| = k$.

Now consider any x, U-fan of size k. Since the paths in a fan are internally vertex-disjoint, and all of them must reach U, at most one of the fan paths can end at z. The remaining $k-1$ paths must end in $U \setminus \{z\}$.

But every x–z path in G meets S (because x and z are in different components of $G - S$). In particular, every path from x to *any* vertex of U that lies in the component of z in $G - S$ must pass through S. Since $|S| \leq k-1$, pigeonhole says k internally disjoint x–U paths cannot all avoid sharing an internal vertex in S: with k paths, at least two would have to pass through the same vertex of S.

Thus no x, U-fan of size k can exist, contradicting the hypothesis. Therefore G must be k-connected. □

21.5 Dirac's theorem on k vertices on common cycle

Theorem 21.10 (Dirac (1960)). Let $k \geq 2$. Every set of k vertices in a k-connected graph G lies on a common cycle of G.

Proof. We argue by induction on k.

Base case $k = 2$. Let $S = \{x, y\}$. Since G is 2-connected, Menger's theorem gives two internally vertex-disjoint x–y paths. Their union is a cycle containing x and y.

Induction step. Assume $k \geq 3$ and the statement holds for $k-1$. Let S be any set of k vertices in G, and fix $x \in S$. Because G is k-connected it is also $(k-1)$-connected, so the induction hypothesis applies to

$$S_0 := S \setminus \{x\}, \qquad |S_0| = k-1.$$

Hence there exists a cycle C in G containing all vertices of S_0.

If $x \in V(C)$, then C already contains S and we are done. So assume $x \notin V(C)$. Write the vertices of S_0 in their cyclic order along C as

$$s_1, s_2, \ldots, s_{k-1},$$

and let A_i denote the s_i–s_{i+1} segment of C (indices taken modulo $k-1$). Thus the segments $A_1, \ldots, A_{k-1}$ partition the edges of C, and by construction each A_i contains no vertex of S_0 internally.

We now insert x into the cycle using the Fan Lemma (Lemma 21.9).

Case 1: $|V(C)| \geq k$. Apply the Fan Lemma in the k-connected graph G to the vertex x and the set $U := V(C)$ (note that $|U| \geq k$). We obtain an x, U-fan of size k, i.e. internally disjoint paths

$$P_1, \ldots, P_k$$

from x to *distinct* vertices $u_1, \ldots, u_k \in V(C)$, each meeting C only at its endpoint.

Since the k endpoints $u_1, \ldots, u_k$ lie on C and the cycle C is partitioned into only $k-1$ segments $A_1, \ldots, A_{k-1}$, by the pigeonhole principle there exist two endpoints, say u and v, that lie on the same segment A_i. Let Q be the u–v subpath of C contained in A_i. By definition of A_i, the interior of Q contains no vertex of S_0.

Let R be the complementary u–v subpath of C (so $C = Q \cup R$ and $Q \cap R = \{u, v\}$). Then R contains every vertex of S_0. Now consider the subgraph

$$C' \ := \ R \ \cup \ P_u \ \cup \ P_v,$$

where P_u, P_v are the two fan paths ending at u and v. Because P_u and P_v are internally disjoint and meet C only at their endpoints, the union $R \cup P_u \cup P_v$ is a simple cycle: it goes from u along P_u to x, back along P_v to v, and then along R to return to u.

This cycle C' contains x and all of $S_0 \subseteq V(R)$, hence contains S.

Case 2: $|V(C)| = k-1$. Then $V(C) = S_0$ (a cycle on $k-1$ vertices cannot contain more than $k-1$ distinct vertices). Apply the Fan Lemma with parameter $k-1$ (valid since G is $(k-1)$-connected) to x and $U := V(C)$. We obtain an x, U-fan of size $k-1$, whose endpoints must therefore be *all* vertices of C. In particular, choose two adjacent vertices u, v on C, and let uv denote the corresponding edge of C. Let R be the u–v path on C that avoids the edge uv; then R contains all other vertices of C.

Let P_u, P_v be the fan paths from x to u and to v. As above, P_u and P_v meet C only at their endpoints and are internally disjoint, so

$$C' \ := \ R \ \cup \ P_u \ \cup \ P_v$$

is a cycle. This cycle contains x and all vertices of $C = V(C)$, hence it contains S.

In both cases we found a cycle containing all k vertices of S. This completes the induction. □

21.6 Ford-Fulkerson CSDR

Definition 21.21 (CSDR). Let $A = \{A_1, \dots, A_m\}$ be a family of sets. Recall a *system of distinct representatives* (SDR) for A is an injective map

$$\varphi : [m] \to \bigcup_{i=1}^{m} A_i \quad \text{such that} \quad \varphi(i) \in A_i \text{ for all } i.$$

Equivalently, it is a set $R = \{\varphi(1), \dots, \varphi(m)\}$ of m *distinct* elements with $\varphi(i) \in A_i$.

Now let $A = \{A_1, \dots, A_m\}$ and $B = \{B_1, \dots, B_m\}$ be two families of m sets. A *common system of distinct representatives* (CSDR) for A and B is a set R of m elements such that R is an SDR for A and also an SDR for B (possibly with different assignments). Equivalently, there exist bijections

$$\varphi_A : [m] \to R, \qquad \varphi_B : [m] \to R$$

with $\varphi_A(i) \in A_i$ and $\varphi_B(j) \in B_j$ for all i, j.

Theorem 21.11 (Ford–Fulkerson (1958)). Let $A = \{A_1, \dots, A_m\}$ and $B = \{B_1, \dots, B_m\}$ be families of m sets. For $I, J \subseteq [m]$ write

$$A(I) := \bigcup_{i \in I} A_i, \qquad B(J) := \bigcup_{j \in J} B_j.$$

Then A and B have a CSDR if and only if

$$|A(I) \cap B(J)| \;\geq\; |I| + |J| - m \qquad \text{for all } I, J \subseteq [m].$$

Proof. It is convenient to rewrite the condition as

$$|A(I) \cap B(J)| + (m - |I|) + (m - |J|) \;\geq\; m \qquad \text{for all } I, J \subseteq [m], \tag{$*$}$$

which is equivalent by rearranging terms.

Step 1: Build a layered digraph. Let $X := \left(\bigcup_i A_i\right) \cup \left(\bigcup_j B_j\right)$ be the ground set of elements. Construct a digraph D with vertex set

$$V(D) = \{s, t\} \;\cup\; A' \;\cup\; X \;\cup\; B', \quad \text{where} \quad A' := \{a_1, \dots, a_m\},\ B' := \{b_1, \dots, b_m\}.$$

Add arcs

$$s \to a_i \ \ (1 \leq i \leq m), \qquad a_i \to x \ \ (x \in A_i),$$
$$x \to b_j \ \ (x \in B_j), \qquad b_j \to t \ \ (1 \leq j \leq m).$$

Every directed s–t path in D has the form

$$s \to a_i \to x \to b_j \to t \quad \text{with} \quad x \in A_i \cap B_j.$$

Step 2: CSDR $\Longleftrightarrow$ m internally disjoint s–t paths. We claim: A and B have a CSDR if and only if D contains m pairwise internally vertex-disjoint directed s–t paths.

($\Rightarrow$) Suppose R is a CSDR. Choose bijections φ_A, φ_B as in Definition 21.21. For each $i \in [m]$ consider the path

$$P_i : \ s \to a_i \to \varphi_A(i) \to b_{\varphi_B^{-1}(\varphi_A(i))} \to t.$$

All these paths are internally vertex-disjoint because: distinct i give distinct a_i; the representatives $\varphi_A(i)$ are distinct elements of R; and distinct representatives also force distinct b-vertices (since φ_B is a bijection onto R). Hence we obtain m internally disjoint s–t paths.

(⇐) Conversely, suppose D has m internally vertex-disjoint s–t paths. Each such path uses exactly one vertex of A' and one of B'. Since there are only m vertices in each of A' and B', disjointness forces the paths to use *all* vertices $a_1, \ldots, a_m$ and *all* vertices $b_1, \ldots, b_m$ exactly once. Let R be the set of the m distinct element-vertices $x \in X$ used by the paths. Assign to each A_i the unique $x \in R$ lying on the path through a_i, and to each B_j the unique $x \in R$ lying on the path through b_j. This makes R an SDR for both A and B, i.e. a CSDR.

So the claim holds.

Step 3: Apply Menger and identify the relevant separators. By the directed vertex version of Menger's theorem, D has m internally disjoint s–t paths if and only if every s–t separating set (vertex cut) has size at least m.

Fix any s–t separating set $R \subseteq V(D) \setminus \{s, t\}$, and define index sets

$$I := \{\, i \in [m] : a_i \notin R \,\}, \qquad J := \{\, j \in [m] : b_j \notin R \,\}.$$

Then R must contain every element-vertex in $A(I) \cap B(J)$: indeed, if $x \in A(I) \cap B(J)$ and $x \notin R$, then there exist $i \in I$ and $j \in J$ with $x \in A_i \cap B_j$, so the path $s \to a_i \to x \to b_j \to t$ avoids R, contradicting that R separates s from t. Therefore

$$A(I) \cap B(J) \ \subseteq\ R.$$

Also, R contains exactly the A'-vertices not in I and the B'-vertices not in J, so

$$|R| \ \geq\ |A(I) \cap B(J)| + (m - |I|) + (m - |J|).$$

Conversely, for every choice of $I, J \subseteq [m]$, the set

$$R_{I,J} := \big(\{a_i : i \notin I\}\big) \ \cup\ \big(A(I) \cap B(J)\big) \ \cup\ \big(\{b_j : j \notin J\}\big)$$

is an s–t separating set (it destroys every possible $s \to a_i \to x \to b_j \to t$ path), and it has size exactly

$$|R_{I,J}| = |A(I) \cap B(J)| + (m - |I|) + (m - |J|).$$

Hence the minimum size of an s–t separating set is

$$\min_{I,J \subseteq [m]} \Big(|A(I) \cap B(J)| + (m - |I|) + (m - |J|)\Big).$$

Therefore, every s–t separator has size at least m if and only if $(*)$ holds for all I, J. By Step 2 and Menger's theorem, this is equivalent to existence of a CSDR. □

Remark 21.4. The inequality $|A(I) \cap B(J)| \geq |I| + |J| - m$ says: no matter how many A-sets you insist on representing ($|I|$) and how many B-sets you insist on representing ($|J|$), the overlap pool $A(I) \cap B(J)$ must be large enough to supply the representatives that must serve *both* families simultaneously.

Theorem 21.12. If G is a 3-regular graph with $|V(G)| \geq 4$, then

$$\kappa(G) = \kappa'(G).$$

Proof. Since G is 3-regular, $\delta(G) = 3$. The general inequalities

$$\kappa(G) \leq \kappa'(G) \leq \delta(G)$$

imply $\kappa(G) \leq \kappa'(G) \leq 3$. Thus it suffices to prove the reverse inequality

$$\kappa'(G) \leq \kappa(G),$$

i.e. for each possible value of $\kappa(G) \in \{0, 1, 2, 3\}$ we will exhibit an edge cut of size $\kappa(G)$.

Case 0: $\kappa(G) = 0$**.** Then G is disconnected, so $\kappa'(G) = 0$ as well. Hence $\kappa'(G) = \kappa(G) = 0$.

Case 1: $\kappa(G) = 1$**.** Then G has a cut-vertex v. Let the components of $G - v$ have vertex sets $C_1, \ldots, C_r$ with $r \geq 2$. Every component C_i must contain at least one neighbor of v; otherwise it would already be a component of G. Since $d(v) = 3$, by the pigeonhole principle there is a component, say C_1, containing *exactly one* neighbor u of v. (Indeed, if every component contained at least two neighbors of v, then $d(v) \geq 2r \geq 4$.)

We claim that the edge uv is a cut-edge. After deleting uv, the vertex set C_1 has no remaining edges to $V(G) \setminus C_1$: it had no edges to the other C_i (different components of $G - v$), and by choice it had no other edge to v. Hence C_1 is isolated in $G - uv$, so uv is a bridge and $\kappa'(G) = 1 = \kappa(G)$.

Case 2: $\kappa(G) = 2$**.** Then G has no cut-vertex, but it does have a 2-vertex separator. Fix one, say $\{u, v\}$. Let the components of $G - \{u, v\}$ have vertex sets $C_1, \ldots, C_r$, where $r \geq 2$.

Observation: each component meets both u and v. That is, for every i there is at least one edge from u into C_i and at least one edge from v into C_i. Indeed, if some component C_i had no neighbor of u, then all its connections to the rest of the graph would go through v, and removing v alone would disconnect G, making v a cut-vertex, contradicting $\kappa(G) = 2$. Symmetrically for u.

For each i define

$$a_i := \|[\,\{u\}, C_i\,]\|, \qquad b_i := \|[\,\{v\}, C_i\,]\|.$$

By the observation, $a_i \geq 1$ and $b_i \geq 1$ for all i.

Now distinguish whether $uv \in E(G)$.

Subcase 2a: $uv \in E(G)$. Then u has exactly 2 edges to $G - \{u, v\}$, so $\sum_i a_i = 2$; similarly $\sum_i b_i = 2$. Since each $a_i, b_i \geq 1$, we must have $r = 2$ and

$$(a_1, a_2) = (1, 1), \qquad (b_1, b_2) = (1, 1).$$

Fix C_1. Then exactly two edges join C_1 to $\{u, v\}$, namely one from u and one from v. Deleting these two edges disconnects C_1 from the rest of the graph, so G has an edge cut of size 2. Hence $\kappa'(G) \leq 2 = \kappa(G)$.

Subcase 2b: $uv \notin E(G)$. Then u has 3 edges to $G - \{u, v\}$, so $\sum_i a_i = 3$; similarly $\sum_i b_i = 3$. Since each $a_i, b_i \geq 1$, we have $r \in \{2, 3\}$.

If $r = 3$, then necessarily $a_i = b_i = 1$ for all i, and again any component C_i is joined to $\{u, v\}$ by exactly two edges; deleting those two edges disconnects C_i.

If $r = 2$, then (a_1, a_2) is either $(1, 2)$ or $(2, 1)$, and the same holds for (b_1, b_2). Let e_u be the *unique* edge from u to the component to which u has only one neighbor (so e_u is the edge accounting for the "1" in (a_1, a_2)). Define e_v analogously for v.

- If e_u and e_v go to the *same* component, then that component has $a_i = b_i = 1$, so the two edges e_u, e_v form a 2-edge cut isolating it.

- If e_u and e_v go to *different* components, then deleting $\{e_u, e_v\}$ disconnects G as follows: after deletion, u has all remaining edges into one component and none into the other, while v has all remaining edges into the other component and none into the first. Since $uv \notin E(G)$ and there are no edges between C_1 and C_2, there is no path from u to v in $G - \{e_u, e_v\}$, so the graph is disconnected.

In all situations we obtain an edge cut of size 2, hence $\kappa'(G) \le 2 = \kappa(G)$.

Case 3: $\kappa(G) = 3$. Then $\kappa'(G) \ge \kappa(G) = 3$, but also $\kappa'(G) \le \delta(G) = 3$, so $\kappa'(G) = 3$.

Combining all cases yields $\kappa'(G) = \kappa(G)$ for every 3-regular graph with $|V(G)| \ge 4$. □

21.7 Characterization of 2-connected graphs

Definition 21.22 (Subdivision of an edge). Let G be a graph and let $e = uv \in E(G)$. *Subdividing* e means deleting e, introducing a new vertex w, and adding the edges

$$e_1 := uw, \qquad e_2 := wv.$$

The resulting graph is denoted G'.

Theorem 21.13 (Characterizations of 2-connected graphs). Let G be a graph with $|V(G)| \ge 3$. The following conditions are equivalent:

1. (A)G is connected and has no cut-vertices.
2. (B) For all $x, y \in V(G)$, there exist two internally vertex-disjoint x–y paths.
3. (C) For all $x, y \in V(G)$, there exists a cycle containing both x and y.
4. (D) $\delta(G) \ge 1$ and for all edges $e, e' \in E(G)$, there exists a cycle containing both e and e'.
5. (F) $\delta(G) \ge 2$ and for all edges $e, e' \in E(G)$, there exists a cycle containing both e and e'.

Proof. We prove a cycle of implications.

(A) $\Longleftrightarrow$ (B). This is exactly Menger's theorem in the case $k = 2$.

(B) $\Longleftrightarrow$ (C). There are two internally disjoint x–y paths iff their union contains a cycle through x and y.

(C)$\Rightarrow$(A). Condition (C) implies G is connected because any two vertices lie on a common cycle, hence are connected by a path. To see there is no cut-vertex, suppose (for contradiction) that v is a cut-vertex. Then $G - v$ has at least two components; pick x, y in different components of $G - v$. Any cycle containing x and y would give two x–y paths on the cycle, and at least one of them avoids v, contradicting that x and y are disconnected in $G - v$. Hence no cut-vertex exists.

(C)$\Rightarrow$(F). Assume (C). Clearly $\delta(G) \ge 2$.

First, $\delta(G) \ge 2$.CNow fix two edges $e = xy$ and $e' = uv$ of G (they may share endpoints, and may even coincide). Form a new graph H from G by adding two new vertices a and b such that

$$N_H(a) = \{x, y\}, \qquad N_H(b) = \{u, v\},$$

and no other new edges are added. In particular, $\deg_H(a) = \deg_H(b) = 2$.

By the Expansion Lemma, H is 2-connected. Hence H satisfies property (C): any two vertices lie on a common cycle. Applying this to a and b, there exists a cycle C in H containing both a and b.

Since a has degree 2 in H, the cycle C must use both edges incident to a, namely ax and ay. Thus C contains the length-2 subpath $x - a - y$ (in one direction around the cycle). Similarly, C must use both bu and bv, so it contains the subpath $u - b - v$.

Now delete the vertices a and b from the cycle C and replace the subpaths $x - a - y$ and $u - b - v$ by the edges xy and uv, respectively. Concretely, we obtain a closed walk in G by

$$x - a - y \rightsquigarrow xy, \qquad u - b - v \rightsquigarrow uv.$$

Because C was a simple cycle and a, b appear only on those forced subpaths, this operation produces a simple cycle in G. That resulting cycle contains the edges $xy = e$ and $uv = e'$, as desired.

(F)⇒(D). Ask a toddler on the street.

(D)⇒(C). We prove the contrapositive: ¬(C) ⇒ ¬(D).

Assume ¬(C). Then there exist vertices $x, y \in V(G)$ that do not lie on a common cycle. Assume for contradiction that (D) holds.

Since $\delta(G) \geq 1$, both x and y are incident with at least one edge. Choose an edge $e \in E(G)$ incident with x, and an edge $e' \in E(G)$ incident with y.

If there were a cycle C containing both e and e', then C would contain both endpoints of e, hence it would contain x; similarly it would contain y. Thus x and y would lie on a common cycle, contradicting our choice of x, y.

Therefore no cycle contains both e and e', so (D) fails. This proves ¬(C) ⇒ ¬(D), hence (D) ⇒ (C).

We have shown (A) ⟺ (B) ⟺ (C) and (C)⇒(F)⇒(D)⇒(C), so all listed conditions are equivalent. □

Corollary 21.14. If G is 2-connected, then the graph G' obtained by subdividing any edge of G is also 2-connected.

Proof. Let $e = uv \in E(G)$, and let G' be obtained from G by subdividing e with a new vertex w, creating edges $e_1 = uw$ and $e_2 = wv$.

We use (F) from the previous theorem: a graph is 2-connected if and only if it has minimum degree at least 2 and every pair of edges lies on a common cycle. First note that subdividing an edge does not create any vertex of degree < 2 (the new vertex w has degree 2, and all other vertices keep their degrees), hence

$$\delta(G') \geq 2.$$

It remains to verify condition (F) for G'. Let $g, h \in E(G')$ be arbitrary. We will show that there is a cycle in G' containing both g and h.

Define the following simple operation on cycles: if C is a cycle in G and $e \in E(C)$, let C' denote the cycle in G' obtained from C by subdividing the edge $e = uv$ with the u–v path $u - w - v$ (i.e., replace e by e_1, e_2). If $e \notin E(C)$, we simply view C as a cycle in G'.

Now consider cases according to how $\{g, h\}$ intersects $\{e_1, e_2\}$.

Case 1: $\{g, h\} \cap \{e_1, e_2\} = \emptyset$. Then $g, h \in E(G) \setminus \{e\}$. Since G is 2-connected, condition (F) holds in G, so there exists a cycle C in G containing both g and h. If $e \notin E(C)$, then C is also a cycle in G' containing g and h. If $e \in E(C)$, replace e by the path $u - w - v$ to obtain C', a cycle in G' containing g and h.

Case 2: $|\{g, h\} \cap \{e_1, e_2\}| = 1$. Without loss of generality, $g = e_1$ and $h \neq e_2$. Then $h \in E(G') \setminus \{e_1, e_2\} \subseteq E(G) \setminus \{e\}$. Because G satisfies (F), there exists a cycle C in G containing the two edges e and h. Replacing e by the path $u - w - v$ yields a cycle C' in G' that contains h and also contains both e_1 and e_2, in particular it contains $g = e_1$ and h.

Case 3: $\{g, h\} = \{e_1, e_2\}$. Since G is 2-connected, (F) implies there exists a cycle C in G containing e. Replacing e by the path $u - w - v$ produces a cycle C' in G' containing both e_1 and e_2, hence containing g and h.

In all cases, g and h lie on a common cycle in G', so G' satisfies condition (F). Therefore G' is 2-connected. □

Definition 21.23 (Ear)**.** Let G be a graph. An *ear* in G is a path P with distinct endvertices s, t such that $\deg_G(s) \geq 3$, $\deg_G(t) \geq 3$, and every internal vertex of P has degree 2 in G.

21.8 Whitney's Ear Decomposition

Definition 21.24 (Ear decomposition)**.** An *ear decomposition* of a graph G is a sequence $(P_0, P_1, \ldots, P_k)$ of subgraphs whose edge-sets partition $E(G)$ and such that:

(a) P_0 is a cycle of length at least 3, and

(b) for each $i = 1, \ldots, k$, the graph P_i is a path whose endvertices lie in $V(P_0 \cup \cdots \cup P_{i-1})$ while its internal vertices are not in $V(P_0 \cup \cdots \cup P_{i-1})$.

Equivalently, if we set $G_i := P_0 \cup \cdots \cup P_i$, then G_i is obtained from G_{i-1} by *adding an ear* P_i.

Theorem 21.15 (Whitney's Ear Decomposition Theorem)**.** A graph G is 2-connected if and only if G has an ear decomposition. Moreover, if G is 2-connected, then every cycle C in G of length at least 3 can be chosen as the initial ear P_0 of some ear decomposition of G.

Proof. **(⇐)** Suppose $(P_0, P_1, \ldots, P_k)$ is an ear decomposition of G. We prove the stronger claim that for every $i \in \{0, 1, \ldots, k\}$, the partial union

$$G_i := P_0 \cup P_1 \cup \cdots \cup P_i$$

is 2-connected.

For $i = 0$, $G_0 = P_0$ is a cycle of length at least 3, hence 2-connected.

Assume $i \geq 1$ and that G_{i-1} is 2-connected. By definition, G_i is obtained from G_{i-1} by adding the ear P_i, i.e., a path whose endpoints lie in $V(G_{i-1})$ and whose internal vertices are new.

View the operation "add a path" as follows: first add an edge between the two endpoints of P_i, obtaining a graph H; then subdivide this new edge repeatedly to create the internal vertices of P_i and thus recover G_i from H. Adding an edge to a 2-connected graph preserves 2-connectivity, so H is 2-connected. By Corollary 4.7, subdividing an edge of a 2-connected graph preserves

2-connectivity. Therefore G_i is 2-connected. This completes the induction, and in particular $G_k = G$ is 2-connected.

(⇒) Now assume G is 2-connected, and let C be any cycle in G of length at least 3. We will construct an ear decomposition starting with $P_0 = C$.

Set $G_0 := C$. Inductively, suppose we have constructed a subgraph G_{i-1} of G that is 2-connected and satisfies $E(G_{i-1}) \subseteq E(G)$. If $G_{i-1} = G$, stop; the process will yield an ear decomposition.

Otherwise, there exists an edge $e = uv \in E(G) \setminus E(G_{i-1})$ with at least one endpoint in $V(G_{i-1})$. Choose such an edge with $u \in V(G_{i-1})$. Since u lies in G_{i-1}, there is an edge $e' \in E(G_{i-1})$ incident to u.

Because G is 2-connected, it satisfies condition (F) of Theorem 4.6: any two edges of G lie on a common cycle. Hence there exists a cycle Γ in G containing both e and e'.

Traverse Γ starting at u in the direction that uses the edge e first. Continue along Γ until you encounter, for the first time after leaving u, a vertex $w \in V(G_{i-1})$. Let P_i be the u–w subpath of Γ obtained this way.

By the choice of w as the *first* vertex of Γ (after u) that lies in $V(G_{i-1})$, every internal vertex of P_i lies outside $V(G_{i-1})$. Thus P_i is an ear of G_{i-1}. Define

$$G_i := G_{i-1} \cup P_i.$$

Then G_i strictly increases the edge set (it contains e), so the construction must terminate after finitely many steps, producing $G_k = G$.

Finally, define $P_0 := C$ and let $P_1, \dots, P_k$ be the successive added ears. By construction, the P_i partition $E(G)$ (each step adds edges not previously present), and each P_i is an ear of the previous union. Hence $(P_0, P_1, \dots, P_k)$ is an ear decomposition of G with initial cycle C. □

22 Hamiltonian Cycles

Definition 22.1 (Hamilton cycle / Hamiltonian graph)**.** A *spanning cycle* in a graph G is a cycle C with

$$V(C) = V(G).$$

Such a cycle is called a *Hamilton cycle* (or *H-cycle*). A graph is *Hamiltonian* if it contains a Hamilton cycle.

Remark 22.1. Determining whether a graph is Hamiltonian is computationally intractable in general (the decision problem is NP-complete). So, unlike matchings, we typically do not expect a clean "if and only if" characterization that is also easy to check. Instead, we look for *necessary* and *sufficient* conditions that are strong enough to be useful.

Example 22.1. The Petersen graph is not Hamiltonian

Example 22.2 (Complete bipartite graphs)**.** For the complete bipartite graph $K_{r,s}$,

$$K_{r,s} \text{ is Hamiltonian} \quad \Longleftrightarrow \quad r = s \geq 2.$$

Proof. Any cycle in a bipartite graph alternates between the two partite sets, so a spanning cycle can exist only if it uses the same number of vertices from each side. Hence a Hamilton cycle in $K_{r,s}$ forces $r = s$. Also we must have $r = s \geq 2$ to even have a cycle.

Conversely, if $r = s \geq 2$ with bipartition $L = \{\ell_1, \dots, \ell_r\}$ and $R = \{r_1, \dots, r_r\}$, then

$$\ell_1\, r_1\, \ell_2\, r_2 \cdots \ell_r\, r_r\, \ell_1$$

is a Hamilton cycle, since all $\ell_i r_j$ edges exist in $K_{r,r}$. □

22.1 Necessary conditions

Theorem 22.1 (A necessary connectivity condition for Hamiltonicity)**.** If G has a Hamilton cycle, then for every nonempty set $S \subseteq V(G)$,

$$c(G - S) \;\leq\; |S|,$$

where $c(H)$ denotes the number of connected components of a graph H.

Proof. Let C be a Hamilton cycle of G, and fix a nonempty set $S \subseteq V(G)$. Delete the vertices of S from the cycle C. Since removing vertices from a cycle can only break it, the graph $C - S$ is a disjoint union of paths (possibly trivial), say

$$C - S = P_1 \cup \cdots \cup P_q,$$

where q is the number of path components of $C - S$.

Now walk once around the cyclic order of C. Between two consecutive vertices of S on the cycle (in this cyclic order), there is a (possibly empty) segment of vertices from $V(G) \setminus S$. Each *nonempty* such segment forms exactly one of the paths P_i. Since there are exactly $|S|$ "gaps" between consecutive vertices of S on the cycle, we get

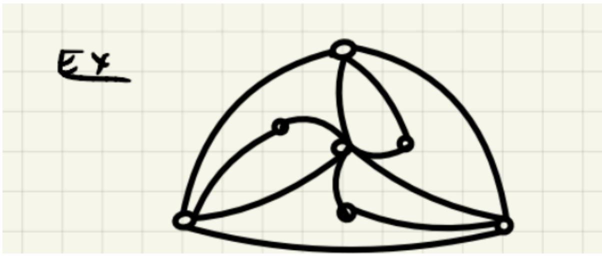
$$q \leq |S|.$$

Finally, note that every component of $G - S$ contains at least one path P_i: indeed, the vertices of $G - S$ are precisely the vertices of $C - S$, and each vertex lies in exactly one P_i. Thus the components of $G - S$ are obtained by possibly *merging* some of the paths P_i together using edges of $G - S$, so the number of components cannot exceed the number of path pieces:

$$c(G - S) \ \leq \ q \ \leq \ |S|.$$

This proves the claim. □

Remark 22.2 (Not sufficient)**.** The condition $c(G - S) \leq |S|$ is a strong obstruction test: a Hamilton cycle cannot "visit" more than $|S|$ different components after removing S, because each component must be entered and exited through vertices of S. However, this condition is *not* sufficient: there exist non-Hamiltonian graphs that still satisfy $c(G - S) \leq |S|$ for every nonempty S (see the example in the figure).



22.2 Ore's Lemma and Dirac's Theorem on Hamiltonian graphs

Lemma 22.2 (Ore's Lemma; Ore (1960))**.** Let x and y be distinct nonadjacent vertices of an n-vertex graph G. If

$$d(x) + d(y) \geq n,$$

then G is Hamiltonian if and only if $G + xy$ is Hamiltonian.

Proof. ($\Rightarrow$) If G is Hamiltonian, then adding the extra edge xy preserves the Hamiltonian cycle, so $G + xy$ is Hamiltonian.

($\Leftarrow$) Assume $G + xy$ has a Hamiltonian cycle, but G does not. In every Hamiltonian cycle of $G + xy$, the edge xy must appear, since otherwise it would also be a Hamiltonian cycle of G. Traverse the Hamiltonian cycle in $G + xy$ from x to y along the x, y-path that lies in G. Index the vertices on this path as

$$x = v_1, v_2, \ldots, v_n = y.$$

Define

$$S = \{ i : v_{i+1} \in N(x) \}, \qquad T = \{ i : v_i \in N(y) \}.$$

Thus $|S| = d(x)$ and $|T| = d(y)$, and both S and T are subsets of $\{1, 2, \ldots, n - 1\}$.

Since

$$|S| + |T| = d(x) + d(y) \geq n,$$

we have

$$|S \cup T| + |S \cap T| = |S| + |T| \geq n.$$

But $S \cup T \subseteq \{1, \ldots, n-1\}$ has size at most $n-1$. Therefore

$$|S \cap T| \geq 1.$$

Choose i in $S \cap T$. Then $v_{i+1} \in N(x)$ and $v_i \in N(y)$, meaning on the x, y-path in G a neighbor of x immediately follows a neighbor of y. Omitting the edge $v_i v_{i+1}$ and using edges xv_{i+1} and $v_i y$ instead produces a spanning cycle entirely in G.

Hence G is Hamiltonian. □

Theorem 22.3 (Dirac). For $n \geq 3$, an n-vertex graph G with $\delta(G) \geq \frac{n}{2}$ is Hamiltonian.

Proof. The requirement $n \geq 3$ is necessary, since K_2 satisfies the degree condition but is not Hamiltonian.

For $n \geq 3$, the complete bipartite graph

$$K_{\lfloor (n-1)/2 \rfloor, \lceil (n+1)/2 \rceil}$$

is not Hamiltonian, yet all of its vertices have degree at least $n/2$. Thus the condition $\delta(G) \geq n/2$ is best possible.

For sufficiency, if $\delta(G) \geq n/2$, then for any $x, y \in V(G)$ we have $d(x) + d(y) \geq n$. Therefore Ore's condition holds, the graph G is Hamiltonian. □

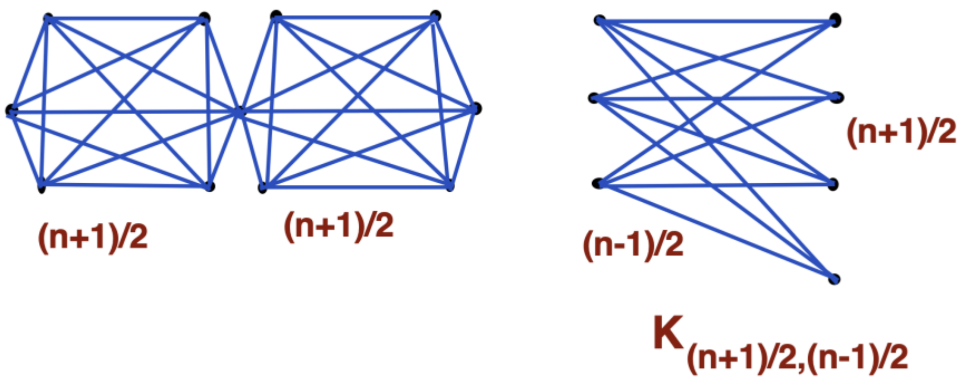


Figure 1: Dirac's bound $\delta(G) \geq n/2$ is sharp; the above graphs do not have Hamilton cycles

22.3 Chvatal's Theorem

Definition 22.2. The Hamiltonian closure of a graph G is the graph $C(G)$ obtained by repeatedly adding an edge uv whenever uv is not already an edge and

$$d(u) + d(v) \geq n,$$

where $n = |V(G)|$.

Lemma 22.4. G has a Hamiltonian cycle if and only if $C(G)$ has a Hamiltonian cycle.

Proposition 22.5 (Chvátal condition)**.** Let $d_n \geq d_{n-1} \geq \cdots \geq d_1$ be the degree sequence of a graph G. Assume that for every $i < \frac{n}{2}$,

$$d_i > i \quad \text{or} \quad d_{n-i} \geq n - i.$$

Then G is Hamiltonian.

Proof. If G is the complete graph K_n, then it is Hamiltonian. Assume $G \neq K_n$. Then some nonedge $uv \notin E(G)$ exists. Choose such a nonedge uv for which $d(u) + d(v)$ is maximum among all nonadjacent pairs. Since G is not complete, we have

$$d(u) + d(v) \leq n - 1. \tag{1}$$

Without loss of generality, let $d(u) \leq d(v)$ and set $i = d(u)$. Thus

$$d(u) = i < \frac{n}{2}.$$

Consider any vertex w not adjacent to u. If $w \notin N(u)$, then by maximality of $d(u) + d(v)$ we must have

$$d(w) \leq d(v).$$

There are $n - 1 - d(u)$ such vertices. In the degree ordering, this means

$$d_{n-i} \leq d(v).$$

But $d(v) \leq n - d(u) = n - i$, so we obtain

$$d_{n-i} \leq n - i.$$

By the hypothesis of the Chvátal condition, since $i < \frac{n}{2}$ and $d_i = i$, we must have

$$d_{n-i} \geq n - i.$$

Combining both inequalities yields

$$d_{n-i} = n - i.$$

Now apply the same argument symmetrically to v. Since $d(u) \leq d(v)$, every nonneighbor of v has degree at most $d(u)$, which implies

$$d_i \leq d(u) < i.$$

This contradicts the assumption that for every $i < \frac{n}{2}$ we have

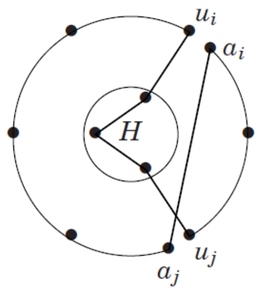
$$d_i > i \quad \text{or} \quad d_{n-i} \geq n - i.$$

Therefore the assumption that G is non-Hamiltonian leads to a contradiction. Hence G is Hamiltonian. □

22.4 Erdős–Chvátal Theorem

Theorem 22.6 (Chvátal–Erdős Theorem)**.** If $\kappa(G) \geq \alpha(G)$, then G has a Hamiltonian cycle (unless G is K_2).

Proof. Let $k = \kappa(G) \geq \alpha(G)$. With $G \neq K_2$, the conditions require $\kappa(G) > 1$, so there is a longest cycle C in G. Since $\delta(G) \geq \kappa(G)$, and since every graph with $\delta(G) \geq 2$ has a cycle of length at least $\delta(G) + 1$, the length of C is at least $k + 1$. Let H be a component of $G - V(C)$. Since $\kappa(G) = k$,at least k vertices of C have neighbors in H.



Let $u_1, \ldots, u_k$ be vertices of C with neighbors in H indexed in order along C. For each i, let a_i be the vertex following u_i along C. If a_i and a_j are adjacent, then we construct a longer cycle by replacing $u_i a_i$ and $u_j a_j$ with $a_i a_j$ and a u_i–u_j path through H (see illustration). Similarly, a_i has no neighbor in H. Hence $\{a_1, \ldots, a_i\}$ plus a vertex of H forms an independent set of size greater than k. This contradiction implies that C is a Hamiltonian cycle. □

Tait, Hamilton cycles, and the Four Color Theorem. In 1880, Peter Guthrie Tait proposed a bold plan to prove the Four Color Conjecture (as it was then called) by translating map-coloring into a statement about cycles. Given a planar map, one can pass to its planar dual graph (we will cover this later). After some modifications, the hard cases can be phrased in terms of *cubic* (3-regular), *bridgeless*, planar graphs. Tait observed that if every such graph had a Hamilton cycle, then the map would be 4-colorable: a Hamilton cycle in a cubic planar graph forces a structure that can be used to produce a 3-edge-coloring, and from that one can derive a 4-coloring of the original map. This became known as *Tait's conjecture*:

> Every 3-regular, 2-edge-connected planar graph is Hamiltonian.

For decades this looked plausibly true and would have implied the Four Color Theorem in a remarkably clean way. But in 1946, W. T. Tutte destroyed the dream by constructing a counterexample: a 3-regular, bridgeless planar graph with *no* Hamilton cycle (the now-famous *Tutte graph*). So Tait's conjecture was false, and the Four Color problem could not be reduced so simply to Hamiltonicity. The Four Color Theorem was finally proved much later (Appel–Haken, 1976) by a very different approach, involving unavoidable sets and computer-assisted checking of reducible configurations.

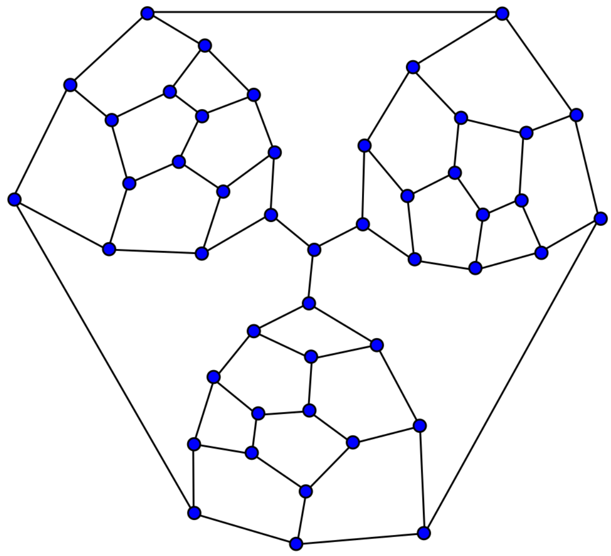


Figure 2: Tutte's counterexample to Tait's conjecture: a 3-regular, bridgeless planar graph with no Hamilton cycle (the "Tutte graph").

Definition 22.3 (Circumference)**.** The *circumference* of a graph G, denoted $c(G)$, is the length (number of vertices) of a longest cycle in G. If G is acyclic, we set $c(G) = 0$.

22.5 Erdős–Gallai Theorem

Theorem 22.7 (Erdős–Gallai (1959))**.** Fix an integer $m \geq 2$. Let G be a graph on n vertices. If

$$e(G) > \frac{m(n-1)}{2},$$

then $c(G) > m$ (equivalently, G contains a cycle of length at least $m + 1$).

Proof. We prove the contrapositive in extremal form: for fixed $m \geq 2$ we show by induction on n that

$$c(G) \leq m \quad \Longrightarrow \quad e(G) \leq \frac{m(n-1)}{2}. \tag{†}$$

Rewriting gives the desired statement.

Reduction to the connected case. If G is disconnected with components $G_1, \ldots, G_r$ ($r \geq 2$), then

$$e(G) = \sum_{i=1}^{r} e(G_i).$$

If (†) failed for G, i.e. $e(G) > \frac{m(n-1)}{2}$, then some component would already violate the corresponding bound: indeed, if for every i we had $e(G_i) \leq \frac{m(|V(G_i)|-1)}{2}$, then summing yields

$$e(G) \leq \frac{m}{2} \sum_{i=1}^{r} (|V(G_i)| - 1) = \frac{m}{2}(n - r) \leq \frac{m(n-2)}{2} < \frac{m(n-1)}{2},$$

a contradiction. Thus it suffices to prove (†) for connected graphs G.

Base case: $n = m + 1$. Assume G is connected, $|V(G)| = n = m + 1$, and $c(G) \le m$. We prove $e(G) \le \frac{m(n-1)}{2} = \frac{(n-1)^2}{2}$.

Suppose for contradiction that $e(G) > \frac{(n-1)^2}{2}$. Then the number of missing edges is

$$\binom{n}{2} - e(G) \;<\; \frac{n(n-1)}{2} - \frac{(n-1)^2}{2} \;=\; \frac{n-1}{2}.$$

If some vertex v had degree $d(v) \le \frac{n}{2} - 1$, then v would have at least

$$(n-1) - \left(\frac{n}{2} - 1\right) = \frac{n}{2}$$

non-neighbors, which would already force at least $\frac{n}{2} \ge \frac{n-1}{2}$ missing edges, a contradiction. Hence $\delta(G) \ge \frac{n}{2}$.

By Dirac's theorem, G is Hamiltonian, so it contains a cycle of length $n = m + 1$, i.e. $c(G) \ge m + 1 > m$, contradicting $c(G) \le m$. Therefore $e(G) \le \frac{(n-1)^2}{2}$, proving (†) in the base case.

Induction step. Assume $n > m + 1$ and that (†) holds for all graphs on fewer than n vertices. Let G be a connected n-vertex graph with $c(G) \le m$. We show $e(G) \le \frac{m(n-1)}{2}$.

Let $P = v_1 v_2 \cdots v_\ell$ be a longest path in G. Among all longest paths, choose P so that the degree of its first vertex v_1 is as large as possible. Set

$$d := d(v_1).$$

If v_1 had a neighbor $x \notin V(P)$, then $x v_1 v_2 \cdots v_\ell$ would be a longer path, contradicting maximality of P. So $N(v_1) \subseteq V(P)$.

"Rotation" set W and bounding degrees inside it: Define

$$W := \{\, v_k : v_1 v_{k+1} \in E(G)\,\}.$$

Since each neighbor of v_1 is some v_{k+1} on P, the map $v_{k+1} \mapsto v_k$ is bijective from $N(v_1)$ to W, so

$$|W| = |N(v_1)| = d.$$

Moreover, for each $v_k \in W$ the path

$$P_k := v_k\, v_{k-1} \cdots v_1\, v_{k+1} \cdots v_\ell$$

is also a longest path (it has the same vertex set and length). By our choice of P (maximizing the degree of the first vertex among longest paths), we must have

$$d(v_k) \;\le\; d(v_1) = d \qquad \text{for all } v_k \in W. \tag{$\star$}$$

Claim: W must be contained among the first m vertices of P. If $v_k \in W$, then $v_1 v_{k+1} \in E(G)$, and the subgraph on

$$v_1, v_2, \ldots, v_{k+1}$$

contains the cycle $v_1 v_2 \cdots v_{k+1} v_1$ of length $k + 1$. Since $c(G) \le m$, we get $k + 1 \le m$, i.e. $k \le m - 1$. Therefore every $v_k \in W$ satisfies $k \le m - 1$.

Let

$$t := \min\{\ell, m\}, \qquad Z := \{v_1, v_2, \ldots, v_t\}.$$

Then $W \subseteq Z$ and $|Z| = t \le m$.

Claim: There are no edges from W to vertices beyond Z. If $\ell < m$, then $Z = V(P)$ and there is nothing to prove. Assume $\ell \ge m$, so $Z = \{v_1, \dots, v_m\}$. We claim that for any $v_k \in W$ and any index $j > m$,

$$v_k v_j \notin E(G). \tag{‡}$$

Indeed, if $v_k v_j \in E(G)$ with $j > m$, then using also $v_1 v_{k+1} \in E(G)$ we obtain the cycle

$$v_1\, v_2\, \cdots\, v_k\, v_j\, v_{j-1}\, \cdots\, v_{k+1}\, v_1,$$

whose length is exactly $j > m$, contradicting $c(G) \le m$. This proves (‡).

Thus *every* edge incident to a vertex of W has its other endpoint in Z.

Delete W and count edges. Let $G_0 := G - W$. Then $|V(G_0)| = n - d$. Let E_W be the set of edges of G with at least one endpoint in W. Then

$$e(G_0) = e(G) - |E_W|. \tag{♣}$$

We now bound $|E_W|$. Because there are no edges from W to $V(G) \setminus Z$ (Step 4), every edge with an endpoint in W lies either inside W or between W and $Z \setminus W$. Write

$$e(W) := e(G[W]), \qquad e[W, Z \setminus W] := \big|\{xy \in E(G) : x \in W,\ y \in Z \setminus W\}\big|.$$

Then

$$|E_W| = e(W) + e[W, Z \setminus W]. \tag{♡}$$

Also,

$$\sum_{w \in W} d(w) = 2e(W) + e[W, Z \setminus W]$$

(each edge inside W is counted twice, and each edge from W to $Z \setminus W$ is counted once). Combining with (♡) gives

$$|E_W| = e(W) + e[W, Z \setminus W] = \frac{1}{2} \sum_{w \in W} d(w) \; + \; \frac{1}{2} e[W, Z \setminus W]. \tag{♠}$$

By ($\star$), each $w \in W$ has $d(w) \le d$, and $|W| = d$, so

$$\sum_{w \in W} d(w) \le d^2.$$

Moreover, each vertex of W can send edges into at most $|Z \setminus W| = t - d$ vertices of $Z \setminus W$, so

$$e[W, Z \setminus W] \le d(t - d).$$

Plugging into (♠) yields

$$|E_W| \le \frac{1}{2} d^2 + \frac{1}{2} d(t - d) = \frac{1}{2} dt \le \frac{1}{2} dm, \tag{◇}$$

since $t \le m$.

From (♣) and (◇), if $e(G) > \frac{m(n-1)}{2}$ then

$$e(G_0) = e(G) - |E_W| > \frac{m(n-1)}{2} - \frac{dm}{2} = \frac{m((n-d)-1)}{2}.$$

By the induction hypothesis applied to the graph G_0 (which has fewer than n vertices), this would force $c(G_0) > m$. But G_0 is a subgraph of G, so $c(G) \ge c(G_0) > m$, contradicting our assumption $c(G) \le m$.

Therefore the assumption $e(G) > \frac{m(n-1)}{2}$ is impossible when $c(G) \le m$. Hence $e(G) \le \frac{m(n-1)}{2}$, completing the induction and proving (†). □

23 Vertex Coloring

23.1 Basics of vertex coloring

Motivation: Vertex-coloring is the "no-conflicts" version of scheduling: vertices are tasks/people/frequencies, edges mean "these two cannot share a label," and colors are the labels. A proper coloring is just an assignment of labels that respects the conflicts. The fewer labels you can get away with, the more structured (or more restrictive) the graph is.

Definition 23.1 (Proper k-coloring). Let $G = (V, E)$ be a graph and let $k \geq 1$. A *(proper) k-coloring* of G is a function

$$f : V(G) \to \{1, 2, \ldots, k\}$$

such that for every edge $xy \in E(G)$,

$$f(x) \neq f(y).$$

Remark 23.1 (Loops and multiple edges). A loop makes proper coloring impossible, so graphs with loops have no proper coloring for any k. Multiple edges do not change anything: if x and y are adjacent once or 10^{100} times, the constraint is still just $f(x) \neq f(y)$. Therefore, when studying vertex-coloring we usually restrict attention to simple graphs.

Definition 23.2 (Color classes). Given a proper k-coloring f, for each color $i \in \{1, \ldots, k\}$ the set

$$f^{-1}(i) = \{v \in V(G) : f(v) = i\}$$

is called the *i-th color class*. Each color class is an independent set.

Proposition 23.1 (Colorings as partitions). A proper k-coloring of G is equivalent to a partition of $V(G)$ into k independent sets (some parts are allowed to be empty).

Remark 23.2 (More colors = more freedom). If G is k-colorable then it is also $(k+1)$-colorable: just reuse the same coloring and ignore the extra color. In particular, any graph on n vertices is n-colorable (color every vertex differently).

Definition 23.3 (Chromatic number). The *chromatic number* of G, denoted $\chi(G)$, is the smallest positive integer k such that G has a proper k-coloring. We say G is *k-colorable* if $\chi(G) \leq k$.

Remark 23.3 (Computational complexity). For each fixed $k \geq 3$, deciding whether a graph is k-colorable is NP-complete. So beyond $k = 2$, we should not expect a clean, fast algorithm that works for all graphs. This is why much of coloring theory focuses on structural sufficient conditions and bounds for $\chi(G)$.

Definition 23.4. We use $\alpha(G)$ for the independence number and $\omega(G)$ for the maximum size of a clique in G, called the clique number.

Proposition 23.2. For every graph G,

$$\chi(G) \ \geq \ \omega(G) \qquad \text{and} \qquad \chi(G) \ \geq \ \frac{|V(G)|}{\alpha(G)}.$$

Proof. **Clique bound.** Let K be a clique in G of maximum size, so $|V(K)| = \omega(G)$. In any proper coloring, adjacent vertices receive different colors, so all vertices of K must receive pairwise distinct colors. Hence at least $\omega(G)$ colors are needed:

$$\chi(G) \geq \omega(G).$$

Independence bound. Let f be a proper $\chi(G)$-coloring of G. Its color classes $f^{-1}(1), \ldots, f^{-1}(\chi(G))$ form a partition of $V(G)$ into independent sets. Each color class is an independent set, so its size is at most $\alpha(G)$:

$$|f^{-1}(i)| \leq \alpha(G) \quad \text{for all } i.$$

Summing over all colors gives

$$|V(G)| \ = \ \sum_{i=1}^{\chi(G)} |f^{-1}(i)| \ \leq \ \sum_{i=1}^{\chi(G)} \alpha(G) \ = \ \chi(G)\,\alpha(G).$$

Rearranging yields

$$\chi(G) \ \geq \ \frac{|V(G)|}{\alpha(G)}.$$

□

Example 23.1. Let $G \sim G(n, \frac{1}{2})$, meaning $V(G) = [n]$ and each edge appears independently with probability $1/2$.

$$\alpha(G) \approx 2\log_2 n, \qquad \omega(G) \approx 2\log_2 n.$$

$$\chi(G) \approx \frac{n}{\alpha(G)} \approx \frac{n}{2\log_2 n}.$$

Recall the *join* of G and H, denoted $G \vee H$, is the graph obtained from the disjoint union $G \cup H$ by adding all edges between $V(G)$ and $V(H)$:

$$V(G \vee H) = V(G) \sqcup V(H), \qquad E(G \vee H) = E(G) \cup E(H) \cup \{xy : \ x \in V(G), \ y \in V(H)\}.$$

Proposition 23.3 (Clique number and chromatic number of a join)**.** For graphs G, H on disjoint vertex sets,

$$\omega(G \vee H) = \omega(G) + \omega(H), \qquad \chi(G \vee H) = \chi(G) + \chi(H).$$

Proof. **Clique number.** Let K be a clique in $G \vee H$. Since all edges between $V(G)$ and $V(H)$ are present, the intersections $K \cap V(G)$ and $K \cap V(H)$ are cliques in G and H, respectively. Hence

$$|K| = |K \cap V(G)| + |K \cap V(H)| \leq \omega(G) + \omega(H),$$

so $\omega(G \vee H) \le \omega(G) + \omega(H)$.

For the reverse inequality, take a maximum clique K_G in G and a maximum clique K_H in H. Then $K_G \cup K_H$ is a clique in $G \vee H$ (all cross-edges are present), and

$$|K_G \cup K_H| = \omega(G) + \omega(H).$$

Thus $\omega(G \vee H) \ge \omega(G) + \omega(H)$, proving equality.

Chromatic number. First we show $\chi(G \vee H) \le \chi(G) + \chi(H)$. Color G properly with colors $\{1, \dots, \chi(G)\}$ and color H properly with fresh colors $\{\chi(G)+1, \dots, \chi(G)+\chi(H)\}$. Because every vertex of G is adjacent to every vertex of H, using disjoint color sets guarantees no conflict across the join edges. Hence this is a proper coloring of $G \vee H$ with $\chi(G) + \chi(H)$ colors.

Now we prove the reverse inequality $\chi(G \vee H) \ge \chi(G) + \chi(H)$. Let f be any proper coloring of $G \vee H$. No color can appear on both sides: if some color c were used on a vertex $x \in V(G)$ and also on a vertex $y \in V(H)$, then xy is an edge of $G \vee H$, contradicting properness. Therefore the set of colors used on $V(G)$ is disjoint from the set of colors used on $V(H)$. In particular, f uses at least $\chi(G)$ colors on $V(G)$ and at least $\chi(H)$ colors on $V(H)$, so in total

$$\chi(G \vee H) \;\ge\; \chi(G) + \chi(H).$$

Combining the two inequalities gives $\chi(G \vee H) = \chi(G) + \chi(H)$. □

Example 23.2 (Joining odd cycles creates a linear gap $\chi - \omega$). Let C_{2t+1} be an odd cycle. Then $\chi(C_{2t+1}) = 3$ and $\omega(C_{2t+1}) = 2$. Let G be the join of k disjoint odd cycles:

$$G = C_{2t_1+1} \;\vee\; C_{2t_2+1} \;\vee\; \cdots \;\vee\; C_{2t_k+1}.$$

Iterating Proposition 23.3 yields

$$\chi(G) = 3k, \qquad \omega(G) = 2k,$$

so $\chi(G) - \omega(G) = k$ grows linearly.

23.2 Greedy coloring

Fix an ordering of the vertices,

$$v_1, v_2, \dots, v_n.$$

Color the vertices one-by-one in this order. When coloring v_i, assign it the smallest positive integer ("the first available color") that is not used by any already-colored neighbor of v_i. This always produces a proper coloring, but the number of colors can depend heavily on the chosen ordering.

Proposition 23.4. For every graph G,

$$\chi(G) \le \Delta(G) + 1,$$

where $\Delta(G)$ is the maximum degree of G. Moreover, for *any* vertex ordering, greedy coloring uses at most $\Delta(G) + 1$ colors.

Proof. Run greedy coloring on an arbitrary ordering $v_1, \dots, v_n$. Fix a step i, and consider the vertex v_i at the moment we are about to color it.

Among the vertices already colored, only the neighbors of v_i impose restrictions: v_i is forbidden from using the colors appearing on those earlier neighbors. But v_i has at most $\deg(v_i) \le \Delta(G)$ neighbors in total, hence at most $\Delta(G)$ earlier neighbors. Therefore at most $\Delta(G)$ colors are forbidden when coloring v_i.

If we allow ourselves the palette of colors $\{1, 2, \dots, \Delta(G) + 1\}$, then by the pigeonhole principle (at most $\Delta(G)$ forbidden colors, but $\Delta(G) + 1$ available colors), there is always at least one color in this palette not used by any earlier neighbor of v_i. Greedy chooses such a color, so the procedure completes using at most $\Delta(G) + 1$ colors.

Hence G is $(\Delta(G) + 1)$-colorable, and therefore $\chi(G) \le \Delta(G) + 1$. □

Example 23.3 (A tree where greedy uses k colors)**.** Fix $k \ge 2$. We construct a tree T_k and a vertex ordering for which the greedy (first-fit) algorithm uses exactly k colors, even though $\chi(T_k) = 2$.

Construction of T_k. Define T_1 to be a single vertex. For $k \ge 2$, assume $T_1, \dots, T_{k-1}$ have been constructed. Create a new vertex r_k (the "root") and for each $i \in \{1, \dots, k-1\}$ connect r_k to the root r_i of T_i by a *subdivision edge*: introduce a new vertex $s_{k,i}$ and add edges

$$r_k s_{k,i}, \qquad s_{k,i} r_i.$$

Equivalently, we attach each earlier tree T_i to r_k by a path of length 2.

The resulting graph is a tree (we attach trees by paths and create no cycles), so $\chi(T_k) = 2$.

The ordering. Order the vertices in the following way. For $i = 1, 2, \dots, k-1$, list all vertices of T_i first (in an order that will be specified inductively), then list the subdivision vertex $s_{k,i}$. Finally, list the new root r_k last.

Claim. In this ordering, greedy uses color i on the root r_i for every $i = 1, \dots, k$.

Proof by induction on k. For $k = 1$ it is trivial. Assume the claim holds for $T_1, \dots, T_{k-1}$.

Consider the greedy coloring on T_k with the ordering described above. By the induction hypothesis, when the algorithm colors the copy of T_i, its root r_i receives color i.

Next, the vertex $s_{k,i}$ appears *after* all of T_i has been colored, and $s_{k,i}$ is adjacent to r_i. So $s_{k,i}$ cannot use color i; in particular, greedy assigns $s_{k,i}$ some color in $\{1, \dots, k-1\} \setminus \{i\}$ (possibly reusing a color from earlier trees).

Finally, the last vertex is r_k. It is adjacent to every $s_{k,i}$ for $i = 1, \dots, k-1$. We claim that for each color $j \in \{1, \dots, k-1\}$, at least one neighbor of r_k has color j. Indeed, take $i = j$. The vertex $s_{k,j}$ is forbidden from using color j (because it is adjacent to r_j), so greedy assigns it the smallest available color *different from* j. Over the collection $\{s_{k,1}, \dots, s_{k,k-1}\}$, every color $1, \dots, k-1$ appears at least once: if a color j were missing entirely from the neighbors of r_k, then when coloring $s_{k,j}$ the color j would have been available and (being the smallest not forbidden by its neighbors) greedy would have used it, contradiction.

Thus, when we color r_k, all colors $1, 2, \dots, k-1$ are present among its already-colored neighbors. Therefore greedy cannot use any of these colors on r_k, and it is forced to introduce a new color:

$$f(r_k) = k.$$

This completes the induction.

Hence greedy uses exactly k colors on the tree T_k even though $\chi(T_k) = 2$.

23.3 Brooks Theorem

Theorem 23.5 (Brooks' Theorem)**.** Let G be a connected graph with maximum degree $\Delta(G)$. If G is neither a complete graph nor an odd cycle, then

$$\chi(G) \le \Delta(G).$$

Equivalently, every connected graph satisfies

$$\chi(G) = \Delta(G) + 1 \quad \text{if and only if} \quad G \text{ is complete or an odd cycle.}$$

We will prove this later.

Proposition 23.6 (Welsh–Powell)**.** If G has degree sequence $d_1, d_2, \ldots, d_n$, then

$$\chi(G) \le 1 + \max_i \ \min\{d_i,\ i-1\}.$$

Proof. Consider the greedy coloring of G in the order $v_1, v_2, \ldots, v_n$.

Fix an index i and look at the moment we color v_i. A color is forbidden for v_i only if it appears on an *earlier neighbor* of v_i. Thus the number of forbidden colors is at most the number of earlier neighbors of v_i.

But v_i has at most d_i neighbors in total, so it has at most d_i earlier neighbors. Also there are only $i-1$ earlier vertices altogether, so it has at most $i-1$ earlier neighbors. Therefore the number of earlier neighbors of v_i is at most

$$\min\{d_i,\ i-1\}.$$

Hence at most $\min\{d_i, i-1\}$ colors are forbidden when coloring v_i.

Greedy always chooses the smallest available color, so it never needs more than

$$1 + \min\{d_i,\ i-1\}$$

colors to color v_i (one more than the number of forbidden colors). Since this holds for every i, the total number of colors used by the greedy algorithm is at most

$$\max_{1\le i\le n} (1 + \min\{d_i,\ i-1\}) = 1 + \max_{1\le i\le n} \min\{d_i,\ i-1\}.$$

Because $\chi(G)$ is the minimum possible number of colors, it is at most the number produced by greedy. Thus the stated bound holds. □

23.4 Degeneracy and Szekeres-Wilf Theorem

Definition 23.5. G is k-degenerate if every subgraph of G has a vertex of degree at most k .

Lemma 23.7 (Degeneracy $\iff$ existence of a k-ordering)**.** Let G be a graph and let $k \geq 0$. The following are equivalent:

1. G is k-degenerate, i.e. every (nonempty) subgraph of G has a vertex of degree at most k.
2. There exists an ordering $v_1, \ldots, v_n$ of $V(G)$ such that for every $2 \leq i \leq n$,

$$\left|N(v_i) \cap \{v_1, \ldots, v_{i-1}\}\right| \leq k,$$

equivalently, v_i has at most k edges to earlier vertices.

Proof. **(i)⇒(ii).** Assume G is k-degenerate. We construct the ordering backwards. Let $G_n := G$. Since G_n is a subgraph of G, it has a vertex of degree at most k; choose one and call it v_n. Delete v_n to obtain $G_{n-1} := G_n - v_n$. Again G_{n-1} is a subgraph of G, so it has a vertex of degree at most k; choose one and call it v_{n-1}. Continue until all vertices are chosen, giving an ordering $v_1, \ldots, v_n$.

Fix $i \geq 2$. When v_i was chosen, it belonged to the current graph G_i, whose vertex set is exactly $\{v_1, \ldots, v_i\}$. By construction, $\deg_{G_i}(v_i) \leq k$. But $\deg_{G_i}(v_i)$ counts precisely the neighbors of v_i among $\{v_1, \ldots, v_{i-1}\}$, hence

$$|N(v_i) \cap \{v_1, \ldots, v_{i-1}\}| = \deg_{G_i}(v_i) \leq k,$$

proving (ii).

(ii)⇒(i). Assume there is an ordering $v_1, \ldots, v_n$ such that each v_i has at most k neighbors among earlier vertices.

Let H be any nonempty subgraph of G (you may take H induced if you like; the argument still works). Choose v_j to be the vertex of H with *largest* index in the ordering among vertices of H. Then every neighbor of v_j inside H must appear earlier in the ordering (since no vertex of H has index larger than j). Therefore

$$\deg_H(v_j) \;\leq\; |N_G(v_j) \cap \{v_1, \ldots, v_{j-1}\}| \;\leq\; k.$$

So H contains a vertex of degree at most k. Since H was arbitrary, every subgraph of G has minimum degree at most k, i.e. G is k-degenerate.

Thus (i) and (ii) are equivalent. □

Theorem 23.8 (Szekeres–Wilf)**.** If G is d-degenerate, then G is $(d+1)$-colorable.

Proof. Let

$$d := \max_{H \subseteq G} \delta(H).$$

We claim that G is d-degenerate, i.e. every subgraph of G contains a vertex of degree at most d.

Indeed, let H be any nonempty subgraph of G. By definition of minimum degree, H has a vertex of degree exactly $\delta(H)$, hence certainly a vertex of degree at most $\delta(H)$. But by the choice of d we have $\delta(H) \leq d$, so H contains a vertex of degree at most d. Since H was arbitrary, G is d-degenerate.

By the degeneracy-ordering lemma, there exists an ordering of the vertices

$$v_1, v_2, \ldots, v_n$$

such that for each $i \geq 2$,

$$\left|N(v_i) \cap \{v_1, \ldots, v_{i-1}\}\right| \leq d.$$

Now run the greedy (first-fit) coloring in this order. When coloring v_i, only earlier neighbors forbid colors, so at most d colors are forbidden. Therefore among the palette $\{1, 2, \ldots, d+1\}$ there is always at least one available color, and the greedy algorithm produces a proper $(d+1)$-coloring of G.

Hence $\chi(G) \leq d + 1 = 1 + \max_{H \subseteq G} \delta(H)$, as claimed. □

23.5 Gallai-Roy Theorem

Theorem 23.9 (Gallai–Roy)**.** For every orientation D of G, if $\ell(D)$ denotes the length of the longest path in D, then

$$\chi(G) \leq 1 + \ell(D).$$

23.6 Mycielski's Construction

The point of Mycielski's construction is to take a triangle-free graph with large chromatic number and increase the chromatic number up by 1 while staying triangle-free. This is how you manufacture graphs with χ as large as you want but still no triangles, showing that χ can be arbitrarily large even if ω is not.

Example 23.4 (Mycielski graph)**.** Let G be a graph with vertex set

$$V(G) = \{v_1, \ldots, v_n\}.$$

Define a disjoint copy of the vertex set

$$V'(G) = \{u_1, \ldots, u_n\},$$

and add one new vertex w. The *Mycielski graph* of G, denoted $\mu(G)$, is the graph with vertex set

$$V(\mu(G)) = V(G) \cup V'(G) \cup \{w\},$$

and edges determined as follows:

1. (**Keep the old graph.**) The induced subgraph on $V(G)$ is G:

$$\mu(G)[V(G)] = G.$$

2. (**Connect w to all copies.**) The neighborhood of w is exactly the copy set:

$$N_{\mu(G)}(w) = V'(G).$$

3. (**Copy each old neighborhood.**) For each $j \in \{1, \ldots, n\}$, connect u_j to exactly the copies of the neighbors of v_j:

$$N_{\mu(G)}(u_j) = \{\, u_i : v_i \in N_G(v_j) \,\}.$$

   Equivalently, in edge language:

$$u_j u_i \in E(\mu(G)) \iff v_j v_i \in E(G).$$

There are *no* edges between $V(G)$ and $V'(G)$, and no edges from w to $V(G)$. Define $M_3 := C_5$, and then recursively

$$M_{k+1} := \mu(M_k).$$

Theorem 23.10 (Mycielski)**.** Given a graph G, let $G' = \mu(G)$ be the graph obtained from G by Mycielski's construction. If $\chi(G) = k$ and G is C_3-free, then

$$\chi(G') = k + 1 \qquad \text{and} \qquad G' \text{ is } C_3\text{-free.}$$

Proof. Write $V(G) = \{v_1, \dots, v_n\}$ and let $U = \{u_1, \dots, u_n\}$ be a disjoint copy. The vertex set of G' is $V(G') = V(G) \cup U \cup \{w\}$, and the edges are:

- $G'[V(G)] = G$;
- w is adjacent to every vertex in U and to no vertex in $V(G)$;
- for all i, j, we have $u_i v_j \in E(G')$ iff $v_i v_j \in E(G)$.

In particular, U is independent in G'.

(1) Triangle-freeness. Assume G is C_3-free. Suppose for contradiction that G' contains a triangle T.

If T contains w, then the other two vertices must lie in U (since $N_{G'}(w) = U$). But U is independent, so this is impossible.

Hence T does not contain w. If T lies entirely in $V(G)$, then it is a triangle in G, contradiction. Therefore T contains at least one vertex of U; say $u_i \in V(T)$. Since U is independent and $w \notin V(T)$, the other two vertices of T must lie in $V(G)$, say v_j and v_ℓ. Then $u_i v_j, u_i v_\ell \in E(G')$, which by construction implies $v_i v_j, v_i v_\ell \in E(G)$. Also $v_j v_\ell \in E(G')$ forces $v_j v_\ell \in E(G)$ because $G'[V(G)] = G$. Thus $v_i v_j v_\ell$ is a triangle in G, contradiction. Hence G' is C_3-free.

(2) Upper bound $\chi(G') \le k + 1$. Let φ be a proper k-coloring of G with color set $\{1, \dots, k\}$. Extend φ to G' by

$$\varphi(u_i) := \varphi(v_i) \quad (1 \le i \le n), \qquad \varphi(w) := k + 1.$$

This is proper: w sees only vertices in U, all colored in $\{1, \dots, k\}$; there are no edges inside U; and for an edge $u_i v_j$ we have $v_i v_j \in E(G)$, so $\varphi(u_i) = \varphi(v_i) \ne \varphi(v_j)$. Therefore $\chi(G') \le k + 1$.

(3) Lower bound $\chi(G') \ge k + 1$. Assume for contradiction that G' has a proper k-coloring ψ with color set $\{1, \dots, k\}$. By permuting color names if needed, we may assume

$$\psi(w) = k.$$

Then every neighbor of w must avoid color k, so

$$\psi(u_i) \in \{1, \dots, k - 1\} \quad \text{for all } i.$$

Let

$$A := \{\, v_i \in V(G) : \psi(v_i) = k \,\}.$$

Note A is independent in G because it is a single color class of ψ and G is an induced subgraph of G' on $V(G)$.

Define a new coloring θ of G with colors in $\{1, \dots, k - 1\}$ by

$$\theta(v_i) := \begin{cases} \psi(u_i), & v_i \in A, \\ \psi(v_i), & v_i \notin A. \end{cases}$$

This is well-defined and uses only $\{1, \dots, k - 1\}$ since $\psi(u_i) \in \{1, \dots, k - 1\}$ and $\psi(v_i) \ne k$ when $v_i \notin A$.

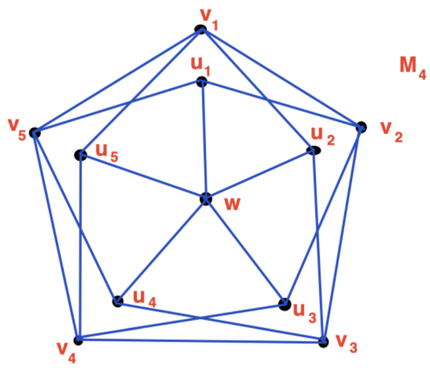


Figure 3: Mycielski graph M_4

We claim θ is proper on G. Consider any edge $v_iv_j \in E(G)$.

If neither endpoint lies in A, then $\theta(v_i) = \psi(v_i)$ and $\theta(v_j) = \psi(v_j)$, and these are different because ψ is proper on $G'[V(G)] = G$.

If both endpoints lie in A, this cannot happen since A is independent.

So suppose $v_i \in A$ and $v_j \notin A$. Then

$$\theta(v_i) = \psi(u_i), \qquad \theta(v_j) = \psi(v_j).$$

Because $v_iv_j \in E(G)$, the construction of G' gives $u_iv_j \in E(G')$. Since ψ is proper on G', we have $\psi(u_i) \neq \psi(v_j)$, hence $\theta(v_i) \neq \theta(v_j)$. Thus θ is a proper $(k-1)$-coloring of G.

This contradicts $\chi(G) = k$. Therefore G' is not k-colorable, so $\chi(G') \geq k+1$.

Combining (2) and (3) yields $\chi(G') = k+1$. □

Theorem 23.11. If G is triangle-free on n vertices, then

$$\chi(G) \leq 2\lceil\sqrt{n}\rceil.$$

(In particular, if n is a perfect square then $\chi(G) \leq 2\sqrt{n}$.)

Proof. Set $t := \lceil\sqrt{n}\rceil$. We color G in two phases.

Phase I (peel off large neighborhoods). Initialize $H := G$. While H has a vertex v with $d_H(v) \geq t$, do:

- introduce a new color (a brand new color never used before);
- color every vertex in the open neighborhood $N_H(v)$ with this new color;
- delete the vertices of $N_H(v)$ from H (leave v uncolored for now).

This is a proper coloring step because G is triangle-free: if two vertices $x, y \in N_H(v)$ were adjacent, then vxy would form a triangle. Hence $N_H(v)$ is an independent set, so it can safely receive one color.

Each iteration colors (and deletes) at least t vertices, so Phase I uses at most

$$\frac{n}{t} \leq \frac{n}{\sqrt{n}} = \sqrt{n} \leq t$$

colors.

Phase II (finish the low-degree remainder). When Phase I stops, the remaining uncolored graph H satisfies $\Delta(H) \leq t - 1$ (since there is no vertex of degree $\geq t$). By the greedy bound $\chi(H) \leq \Delta(H) + 1$, we can color H using at most

$$\Delta(H) + 1 \leq (t - 1) + 1 = t$$

additional colors.

Total. Phase I uses $\leq t$ colors and Phase II uses $\leq t$ colors, so $\chi(G) \leq 2t = 2\lceil\sqrt{n}\rceil$. □

24 Color-critical graphs

Definition 24.1 (Color-critical and k-critical). A graph G is *color-critical* if every proper subgraph $H \subsetneq G$ satisfies
$$\chi(H) < \chi(G).$$
If $\chi(G) = k$, we also say that G is *k-critical*.

Remark 24.1 (Immediate consequences). If G is color-critical, then:
$$\forall v \in V(G):\ \chi(G - v) < \chi(G), \qquad \forall e \in E(G):\ \chi(G - e) < \chi(G).$$
In particular, G has no isolated vertex (deleting an isolated vertex does not change χ).

Example 24.1. 1. The unique 1-critical graph is K_1.

2. The unique 2-critical graph is K_2.

3. The 3-critical graphs are exactly the odd cycles C_{2t+1}.

Lemma 24.1 (Minimum degree bound). If G is k-critical, then
$$\delta(G) \geq k - 1.$$

Proof. Let $v \in V(G)$. Since G is k-critical, we have $\chi(G - v) \leq k - 1$. Fix a proper $(k-1)$-coloring c of $G - v$.

If some color in $\{1, \ldots, k-1\}$ does *not* appear on $N_G(v)$, then we could assign v that missing color and obtain a proper $(k-1)$-coloring of G, contradicting $\chi(G) = k$. Hence every one of the $k - 1$ colors appears on $N_G(v)$, so $|N_G(v)| \geq k - 1$, i.e. $\deg(v) \geq k - 1$.

Since v was arbitrary, $\delta(G) \geq k - 1$. □

Proposition 24.2. Let $\chi(G) = k$.

1. If $\chi(G - v) < \chi(G)$ for some $v \in V(G)$, then there exists a proper k-coloring f of G such that:
$$f(v) = k \text{ is used only on } v, \qquad \text{and} \qquad \{1, 2, \ldots, k-1\} \subseteq f\big(N_G(v)\big).$$

2. If $\chi(G - e) < \chi(G)$ for some edge $e = uv \in E(G)$, then in *every* proper $(k-1)$-coloring g of $G - e$ we have
$$g(u) = g(v).$$

Proof. 1. Since $\chi(G - v) \leq k - 1$, pick a proper $(k-1)$-coloring of $G - v$ and call it f. Extend it to G by assigning $f(v) := k$. Now v is the unique vertex of color k. If some color $c \in \{1, \ldots, k-1\}$ did *not* appear on $N_G(v)$, then we could recolor v with c and obtain a proper $(k-1)$-coloring of G, contradicting $\chi(G) = k$. Hence every color $1, \ldots, k-1$ appears on $N_G(v)$.

2. Let $e = uv$. Take any proper $(k-1)$-coloring g of $G - e$. If $g(u) \neq g(v)$, then g is also a proper $(k-1)$-coloring of G (because adding the edge uv would still connect different colors), contradicting $\chi(G) = k$. Therefore $g(u) = g(v)$ in every $(k-1)$-coloring of $G - e$. □

24.1 Connectivity properties of color-critical graphs

Proposition 24.3. If G is k-critical with $k \geq 2$, then G is 2-connected. Equivalently, G has no cut-vertex.

Proof. First, G is connected: otherwise $\chi(G) = \max_i \chi(G_i)$ over components, so deleting a vertex from a component not attaining the maximum would not decrease χ, contradicting criticality.

Now suppose for contradiction that G has a cut-vertex v. Then $G - v$ is disconnected; write its components as $H_1, \dots, H_t$ with $t \geq 2$, and let $G_i := G[V(H_i) \cup \{v\}]$.

For each i, G_i is a proper subgraph of G, so by k-criticality

$$\chi(G_i) \leq k - 1.$$

Fix proper $(k-1)$-colorings c_i of each G_i. By permuting the color names inside each c_i (allowed since colors are just labels), we may assume that all c_i assign the *same* color to v.

Now define a coloring c of G by setting $c|_{V(G_i)} := c_i$ for each i. This is well-defined because the only overlap between the vertex sets of the G_i is the single vertex v, and we forced agreement there. No edge joins H_i to H_j for $i \neq j$, so c is a proper $(k-1)$-coloring of all of G. This contradicts $\chi(G) = k$.

Therefore G has no cut-vertex. Since G is connected, it is 2-connected. □

Theorem 24.4 (Dirac)**.** Every k-critical graph G is $(k-1)$-edge-connected

Proof. Let G be k-critical. Let F be a minimum edge-cut, so $G - F$ is disconnected. Choose one component of $G - F$ with vertex set X, and put $Y := V(G) \setminus X$. Then $F = E_G(X, Y)$.

Since X and Y are proper nonempty vertex subsets, the induced subgraphs $G[X]$ and $G[Y]$ are proper subgraphs of G, hence

$$\chi(G[X]) \leq k - 1 \qquad \text{and} \qquad \chi(G[Y]) \leq k - 1$$

by k-criticality. Fix proper $(k-1)$-colorings of $G[X]$ and $G[Y]$. Let

$$X = X_1 \cup \cdots \cup X_{k-1}, \qquad Y = Y_1 \cup \cdots \cup Y_{k-1}$$

be the corresponding color classes (so each X_i and Y_j is independent).

Now build a bipartite graph B with left part $\{1, \dots, k-1\}$ (the colors on X) and right part $\{1, \dots, k-1\}$ (the colors on Y), where we join i to j in B iff there is *no* edge of G between X_i and Y_j. (Think: pairing color i on X with color j on Y would be "safe across the cut".)

Claim: B has no perfect matching. Indeed, if M were a perfect matching, then each j on the right is matched to a unique i on the left. Recolor every vertex of Y_j with color i. This is just a permutation of colors inside Y, so it remains a proper $(k-1)$-coloring of $G[Y]$. And because $(i, j) \in M$ implies there are no edges between X_i and Y_j, there are no monochromatic edges across the cut. Thus we obtain a proper $(k-1)$-coloring of all of G, contradicting $\chi(G) = k$. So B has no perfect matching.

By Hall's theorem, there exists a nonempty set $S \subseteq \{1, \dots, k-1\}$ such that

$$|N_B(S)| < |S|.$$

Let $T := \{1, \dots, k-1\} \setminus N_B(S)$. Then

$$|T| = (k-1) - |N_B(S)| \geq (k-1) - (|S|-1) = k - |S|.$$

Moreover, if $i \in S$ and $j \in T$, then $j \notin N_B(S)$, so (i, j) is *not* an edge of B. By definition of B, this means there *is* at least one edge of G between X_i and Y_j.

Each cut-edge in $E_G(X, Y)$ lies between a unique pair (X_i, Y_j), so the previous paragraph implies

$$|E_G(X, Y)| \;\geq\; |S| \cdot |T| \;\geq\; |S|\,(k - |S|).$$

For integers $1 \leq |S| \leq k-1$, the minimum of $|S|(k - |S|)$ occurs at the endpoints $|S| = 1$ or $|S| = k-1$, and equals $k-1$. Hence

$$|F| = |E_G(X, Y)| \geq k - 1.$$

Since F was a minimum edge-cut, every edge-cut has size at least $k-1$, so G is $(k-1)$-edge-connected. □

24.2 Hajós construction (building k-critical graphs of connectivity 2)

The Hajós construction is a standard way to *manufacture* new k-critical graphs from old ones, while keeping the vertex-connectivity as small as possible (namely 2).

Example 24.2 (Hajós)**.** Let G_1 and G_2 be vertex-disjoint graphs, and fix edges $x_1y_1 \in E(G_1)$ and $x_2y_2 \in E(G_2)$. Form a new graph G^* by:

1. deleting the edges x_1y_1 and x_2y_2;
2. identifying (gluing) the vertices x_1 and x_2 into a single new vertex x^*;
3. adding the edge y_1y_2.

Remark 24.2 (Why $\kappa(G^*) = 2$)**.** By construction, $G^* - \{x^*, y_1\}$ is disconnected: removing x^* separates the two "halves", and removing y_1 in addition kills the only remaining link y_1y_2 from the G_1-side to the G_2-side. Thus $\{x^*, y_1\}$ is a separating set, so $\kappa(G^*) \leq 2$ (and in the usual applications one checks G^* is 2-connected, hence $\kappa(G^*) = 2$).

Theorem 24.5 (Hajós preserves k-criticality)**.** If G_1 and G_2 are k-critical, then the Hajós graph G^* is also k-critical.

Proof. Write x^* for the identified vertex, and note that

$$V(G^*) = \big(V(G_1) \cup V(G_2)\big) \setminus \{x_1, x_2\} \;\cup\; \{x^*\}.$$

Step 1: $\chi(G^*) \geq k$. Suppose for contradiction that G^* has a proper $(k-1)$-coloring f.

Restrict f to $V(G_1) \setminus \{x_1\}$ and give x_1 the color $f(x^*)$. Since the only edge of G_1 that we deleted was x_1y_1, this assignment fails to be a proper $(k-1)$-coloring of G_1 *only if* x_1 and y_1 receive the same color. Hence

$$f(x^*) = f(y_1).$$

Applying the same argument on the G_2 side yields

$$f(x^*) = f(y_2).$$

Therefore $f(y_1) = f(y_2)$, contradicting that y_1y_2 is an edge of G^* and f is proper. Thus G^* is not $(k-1)$-colorable, so $\chi(G^*) \geq k$.

Step 2: every edge deletion lowers the chromatic number. Let $e \in E(G^*)$. We show $\chi(G^* - e) \leq k - 1$.

Case 2a: $e = y_1y_2$. Since G_i is k-critical, the graph $G_i - x_iy_i$ is $(k-1)$-colorable. Let f_i be a proper $(k-1)$-coloring of $G_i - x_iy_i$. Because x_i and y_i are *nonadjacent* in $G_i - x_iy_i$, we may (if necessary) permute the colors within f_i so that

$$f_i(x_i) = f_i(y_i).$$

Now permute the colors of f_2 so that $f_2(x_2) = f_1(x_1)$. After identifying x_1 and x_2 into x^*, define a coloring f of $G^* - y_1y_2$ by

$$f(v) = \begin{cases} f_1(v) & v \in V(G_1) \setminus \{x_1\}, \\ f_2(v) & v \in V(G_2) \setminus \{x_2\}, \\ f_1(x_1) = f_2(x_2) & v = x^*. \end{cases}$$

This is well-defined and proper: inside each side it agrees with a proper coloring; across the cut there is *no* edge except y_1y_2, which we removed. Hence $\chi(G^* - y_1y_2) \leq k - 1$.

Case 2b: e is any other edge. By symmetry, assume $e \in E(G_1)$ after the construction (this includes edges incident to x^* that came from G_1).

Again, k-criticality of G_1 gives a proper $(k-1)$-coloring f_1 of $G_1 - e$. Also k-criticality of G_2 gives a proper $(k-1)$-coloring f_2 of $G_2 - x_2y_2$, and as above we may choose f_2 so that $f_2(x_2) = f_2(y_2)$. Permute colors in f_2 so that $f_2(x_2) = f_1(x_1)$, and then glue the colorings into a coloring of $G^* - e$ exactly as in Case 2a. The only potential conflict across the two sides is the edge y_1y_2, but y_1 and y_2 may have different colors (and if $e \neq y_1y_2$ we did not delete that edge), so the resulting coloring is still proper.

Thus in all cases $\chi(G^* - e) \leq k - 1$.

Conclusion. We have shown $\chi(G^*) \geq k$ and $\chi(G^* - e) \leq k - 1$ for every $e \in E(G^*)$, which is exactly that G^* is k-critical. □

24.3 Proof of Brooks Theorem

We now return to the proof of Brooks Theorem.

> **Theorem 24.6** (Brooks' Theorem)**.** Let G be a graph with maximum degree Δ. If $\Delta \geq 3$ and G contains no clique $K_{\Delta+1}$, then G is Δ-colorable (i.e. $\chi(G) \leq \Delta$).

Proof. We may assume G is connected, since different components can be colored independently using the same palette of Δ colors.

Assume for contradiction that G is a counterexample with the fewest vertices. Thus G is not complete and not an odd cycle, $\Delta(G) = \Delta$, and $\chi(G) > \Delta$. Since $\chi(G) \leq \Delta + 1$ always, we have

$$\chi(G) = \Delta + 1.$$

By minimality, for every vertex u the graph $G - u$ is Δ-colorable, so $\chi(G - u) \leq \Delta < \Delta + 1 = \chi(G)$. Hence G is $(\Delta + 1)$-*critical*.

By previously proved results about k-critical graphs (applied with $k = \Delta + 1$), we have:

$$\delta(G) \geq (\Delta + 1) - 1 = \Delta \quad \text{and} \quad G \text{ is 2-connected.}$$

Step 2: Find a, b at distance 2 with $G - a - b$ connected.

Since also $\Delta(G) = \Delta$, the minimum-degree bound forces G to be Δ-*regular*. Moreover, 2-connectedness implies that for every vertex u, the graph $G - u$ is connected.

If $\Delta = |V(G)| - 1$, then every vertex has degree $|V(G)| - 1$, so G is complete, contrary to hypothesis. Hence

$$\Delta \leq |V(G)| - 2.$$

Fix any vertex $v \in V(G)$. Since v is not adjacent to every other vertex, pick a vertex $t \notin N(v) \cup \{v\}$, and let

$$v = p_0, p_1, p_2, \ldots, p_\ell = t$$

be a shortest v–t path. Then $\ell \geq 2$, so $p_2 \neq v$ exists and $p_2 \notin N(v)$ (otherwise vp_2 would shorten the path). Set

$$a := v, \qquad b := p_2.$$

Then $\text{dist}(a, b) = 2$, and a and b have a common neighbor p_1.

It remains to ensure $G - a - b$ is connected. Consider $G - a = G - v$, which is connected because G is 2-connected. If $(G - v) - b$ is connected, we are done. Otherwise, b is a cut-vertex of $G - v$. Let B be an endblock of $G - v$ not containing p_1 (equivalently, not containing the neighbor of v on the chosen shortest path), and let z be the unique cut-vertex of $G - v$ in B.

Because G is 2-connected, the vertex v must have a neighbor in $B - \{z\}$; otherwise z would separate $B - \{z\}$ from the rest of G. Choose such a neighbor $a' \in B - \{z\}$ of v, and set $a := a'$ while keeping $b := p_2$ as above. Then a and b are nonadjacent (they lie in distinct blocks of $G - v$) and still have distance 2 through v. Moreover, removing a deletes a vertex from an endblock B but leaves the block attached through z, so $(G - v) - a$ remains connected; hence $G - a - b$ is connected as well.

Thus we have vertices a, b with $\text{dist}(a, b) = 2$ such that $H := G - a - b$ is connected. Let v be a common neighbor of a and b (so $a - v - b$ is a path).

Step 3: Greedy coloring in reverse order. Let $H := G - a - b$, which is connected. Choose an ordering

$$x_1 = v,\ x_2,\ \ldots,\ x_m$$

of $V(H)$ such that each x_i ($i \geq 2$) has a neighbor among $\{x_1, \ldots, x_{i-1}\}$ (e.g. a rooted spanning tree order).

We now Δ-color G.

(i) Precolor a and b. Assign both a and b color 1. This is legal because a and b are nonadjacent.

(ii) Color $x_m, x_{m-1}, \ldots, x_2$ greedily. When coloring x_i ($i \geq 2$), it has a neighbor among $\{x_1, \ldots, x_{i-1}\}$ that is still uncolored (since we color in reverse). Hence at most $\deg_G(x_i) - 1 \leq \Delta - 1$ of its neighbors are already colored, so some color in $\{1, \ldots, \Delta\}$ is available.

(iii) Color $x_1 = v$ last. All neighbors of v are now colored, including a and b, and both a, b have color 1. Since v has degree Δ, the colors appearing on $N(v)$ use at most $\Delta - 1$ distinct colors, so some color in $\{1, \ldots, \Delta\}$ is missing from $N(v)$. Color v with that missing color.

This yields a proper Δ-coloring of G, contradicting $\chi(G) = \Delta + 1$. Therefore G is Δ-colorable.

□

24.4 List coloring

Definition 24.2. For a graph G, a list assignment L assigns to each vertex $v \in V(G)$ a set $L(v)$ of colors allowed at v. An L-coloring of G is a propositioner coloring f such that $f(v) \in L(v)$ for every vertex v.

Definition 24.3. A graph G is k-choosable (or list k-colorable) if for every list assignment L satisfying $|L(v)| \geq k$ for all $v \in V(G)$, the graph G has an L-coloring. The list chromatic number $\chi_\ell(G)$ is the minimum k such that G is k-choosable.

Theorem 24.7. Let $m = \binom{2k-1}{k}$. Then $K_{m,m}$ is not k-choosable. Hence

$$\chi_\ell(K_{m,m}) > k, \qquad \chi(K_{m,m}) = 2.$$

Proof. Let the color set be $\{1, 2, \ldots, 2k-1\}$, and let L assign to each vertex every k-subset of this set. Since there are $\binom{2k-1}{k} = m$ such subsets, this defines distinct lists for the m vertices in each part of $K_{m,m}$.

Suppose toward a contradiction that $K_{m,m}$ has an L-coloring using the colors $\{1, \ldots, 2k-1\}$. Let the two partite sets be X and Y. Since each list has size k, every vertex must receive one of the k colors in its list. In particular, each part must use at least k colors; otherwise some vertex in that part would have no available color from its list.

Without loss of generality, assume the left part X uses at most $k-1$ distinct colors. Let B be the set of colors used on X; then $|B| \leq k-1$. Therefore the remaining colors

$$\{1, 2, \ldots, 2k-1\} \setminus B$$

form a set of size at least k.

Hence Y must use at least one color from this remaining set. But any such color appears in at least one list on X as well, because every k-subset occurs as a list on both sides. Since $K_{m,m}$ is complete bipartite, this forces some edge between X and Y to have both endpoints receiving the same color, contradicting propositioner coloring. Thus no L-coloring exists. Since L assigns lists of size k, the graph $K_{m,m}$ is not k-choosable, and therefore $\chi_\ell(K_{m,m}) > k$. □

Theorem 24.8 (List-coloring version of Brooks' theorem)**.** If G is connected, not complete, and not an odd cycle, then

$$\chi_\ell(G) \leq \Delta(G).$$

25 Edge Coloring

25.1 Basics of edge-coloring

Definition 25.1 (k-edge-coloring)**.** A *(proper) k-edge-coloring* of a graph G is a function

$$f : E(G) \longrightarrow \{1, 2, \ldots, k\}$$

such that for every color $i \in \{1, \ldots, k\}$, the set

$$f^{-1}(i) := \{ e \in E(G) : f(e) = i \}$$

is a *matching* (that is, no two edges in $f^{-1}(i)$ share an endpoint). The sets $f^{-1}(i)$ are called the *color classes* of f.

Remark 25.1 (Immediate observations)**.** 1. If G has a loop, then G has no proper k-edge-coloring for any k. Indeed, a loop is incident to its own endpoint twice, so it cannot share a color class with anything, and in particular it violates the matching condition.

2. Multiple edges *do* affect edge-coloring: if two parallel edges join the same pair of vertices, they are incident at both ends and hence must receive different colors.

3. Equivalently, f is a proper edge-coloring iff for every vertex $v \in V(G)$, the edges incident to v all receive distinct colors.

Remark 25.2 (Partition viewpoint)**.** Giving a k-edge-coloring is the same thing as partitioning the edge set into k matchings:

$$E(G) = M_1 \sqcup M_2 \sqcup \cdots \sqcup M_k,$$

where $M_i = f^{-1}(i)$ is the ith color class.

Definition 25.2 (Edge chromatic number)**.** The *edge chromatic number* (also called the *chromatic index*) of G is

$$\chi'(G) := \min\{ k \in \mathbb{Z}_{>0} : G \text{ has a proper } k\text{-edge-coloring} \}.$$

We say G is *k-edge-colorable* if $\chi'(G) \le k$.

Lemma 25.1 (Trivial lower bound)**.** For every graph G,

$$\chi'(G) \ge \Delta(G).$$

Proof. Let v be a vertex of maximum degree $\Delta(G)$. In any proper edge-coloring, all $\Delta(G)$ edges incident to v must receive pairwise distinct colors, so at least $\Delta(G)$ colors are needed. □

Example 25.1 (Cycles). For the cycle C_n,

$$\chi'(C_n) = \begin{cases} 2, & n \text{ even}, \\ 3, & n \text{ odd}. \end{cases}$$

Example 25.2 (Complete graphs). For the complete graph K_n,

$$\chi'(K_n) = \begin{cases} n-1, & n \text{ even}, \\ n, & n \text{ odd}. \end{cases}$$

Example 25.3 (Bipartite graphs (Kőnig's line coloring theorem)). If G is bipartite, then

$$\chi'(G) = \Delta(G).$$

Example 25.4 (Petersen graph). The Petersen graph is 3-regular, but it is *not* 3-edge-colorable:

$$\chi'(\text{Petersen}) = 4.$$

(Equivalently: Petersen has no decomposition of its 15 edges into 3 perfect matchings.)

Example 25.5 (3-regular graphs with a bridge). If G is 3-regular and has a cut-edge (bridge), then G is not 3-edge-colorable, so

$$\chi'(G) = 4.$$

Proposition 25.2. $\chi'(G) \leq 2\Delta(G) - 1$

Proof. An edge-coloring of G is exactly a vertex-coloring of its line graph $L(G)$: each edge of G becomes a vertex of $L(G)$, and two vertices of $L(G)$ are adjacent iff the corresponding edges in G share an endpoint. Therefore

$$\chi'(G) = \chi(L(G)).$$

Now apply greedy coloring to $L(G)$. When a vertex is colored, the only colors forbidden are those already used on its previously colored neighbors. Hence if every vertex has at most D previously colored neighbors, then $D + 1$ colors always suffice.

So it remains to bound the degree in $L(G)$. Let $e = uv$ be an edge of G. In $L(G)$, the vertex e is adjacent to all edges incident to u except e itself (there are $\deg(u) - 1$ of them) and all edges incident to v except e (itself) (there are $\deg(v) - 1$ of them). Thus

$$\deg_{L(G)}(e) = (\deg(u) - 1) + (\deg(v) - 1) \leq 2\Delta(G) - 2.$$

Therefore, in a greedy coloring of $L(G)$, when we color e there are at most $2\Delta(G) - 2$ forbidden colors, so one more color always exists. Hence

$$\chi(L(G)) \leq (2\Delta(G) - 2) + 1 = 2\Delta(G) - 1,$$

and using $\chi'(G) = \chi(L(G))$ we conclude

$$\chi'(G) \leq 2\Delta(G) - 1.$$

□

25.2 Shannon's Theorem

Theorem 25.3 (Shannon, 1949). Let G be a loopless multigraph with maximum degree $\Delta := \Delta(G)$. Then
$$\chi'(G) \le \left\lfloor \frac{3\Delta}{2} \right\rfloor.$$

Proof. If $\Delta \le 1$ the statement is immediate. Assume $\Delta \ge 2$ and set
$$s := \left\lfloor \frac{3\Delta}{2} \right\rfloor, \qquad M := \{1, 2, \dots, s\}.$$

We prove the theorem by induction on $|E(G)|$ over all loopless multigraphs with maximum degree at most Δ.

Base case. If $|E(G)| \le s$, color every edge with a distinct color from M.

Induction step. Let G have m edges, and assume every loopless multigraph with maximum degree $\le \Delta$ and at most $m-1$ edges is s-edge-colorable. Pick an edge $e = uv$ and let $G_1 := G - e$. By induction, G_1 has a proper edge-coloring
$$f : E(G_1) \to M.$$

For a vertex $x \in V(G)$ define the set of *free colors at x* by
$$C(x) := \{\, c \in M : \text{no edge incident to } x \text{ has color } c \text{ under } f \,\}.$$

Because incident edges receive distinct colors,
$$|C(x)| = s - d_{G_1}(x) \ \ge\ s - \Delta\ \ge \left\lfloor \frac{\Delta}{2} \right\rfloor \qquad \text{for all } x,$$

and moreover
$$|C(u)|, |C(v)|\ \ge\ s - (\Delta - 1) = \left\lfloor \frac{\Delta}{2} \right\rfloor + 1, \tag{9}$$

since $d_{G_1}(u), d_{G_1}(v) \le \Delta - 1$.

If $C(u) \cap C(v) \neq \emptyset$, we are done. Indeed, choose $c \in C(u) \cap C(v)$ and color e with c.

So assume from now on that
$$C(u) \cap C(v) = \emptyset. \tag{10}$$

For colors $a, b \in M$, let $H_{a,b}$ be the subgraph of G_1 consisting of all edges colored a or b. Every component of $H_{a,b}$ is a path or an even cycle with colors alternating along it.

Claim 1: For any $a \in C(u)$ and $b \in C(v)$, the vertices u and v lie in the same component of $H_{a,b}$; equivalently, there is an a–b alternating u–v path.

Proof. Suppose not. Then u lies in some component K of $H_{a,b}$ that does not contain v. Swap the colors a and b on every edge of K (this preserves a proper edge-coloring). Since $a \in C(u)$, the vertex u had no incident a-edge in K, so after swapping, u has no incident b-edge either; i.e. $b \in C(u)$ in the new coloring. But $b \in C(v)$ always (we did not touch the component of v), so now $b \in C(u) \cap C(v)$, contradicting (12). □

Pick any colors $a \in C(u)$ and $b \in C(v)$. By Claim 1, there is an alternating a–b path from u to v. In particular, since b is free at v, the last edge on this path entering v must have color a. Let that edge be vw, so

$$f(vw) = a. \tag{11}$$

Claim 2: $C(v) \cap C(w) = \emptyset$.

Proof. If some color t were free at both v and w, we could recolor the edge vw with t, making a free at v (because (11) was the only a-edge incident to v along that a–b chain). Since $a \in C(u)$, this would create $a \in C(u) \cap C(v)$, contradicting (12). □

Claim 3: $C(u) \cap C(w) \neq \emptyset$.

Proof. Using (9) and $|C(w)| \geq \lfloor \Delta/2 \rfloor$,

$$|C(u)| + |C(v)| + |C(w)| \;\geq\; 3\left\lfloor \frac{\Delta}{2} \right\rfloor + 2 \;>\; \left\lfloor \frac{3\Delta}{2} \right\rfloor \;=\; |M|.$$

But by (12) and Claim 2, the set $C(v)$ is disjoint from $C(u) \cup C(w)$, so if $C(u)$ and $C(w)$ were also disjoint then

$$|C(u)| + |C(v)| + |C(w)| = |C(v)| + |C(u) \cup C(w)| \leq |C(v)| + |M \setminus C(v)| = |M|,$$

contradiction. Hence $C(u) \cap C(w) \neq \emptyset$. □

Let $c \in C(u) \cap C(w)$. Apply Claim 1 again, now to the pair (c, b): there is a c–b alternating u–v path P. Since c is free at w, the path P cannot pass through w (because an internal vertex on a c–b alternating path must be incident to both colors, and w is incident to no c-edge). Therefore the edge of P incident to v is *not* vw.

Perform a Kempe swap on P (swap colors c and b along P). After this swap:

- c becomes free at v (because b was free at v and P ends at v), and
- c remains free at w (since P avoids w).

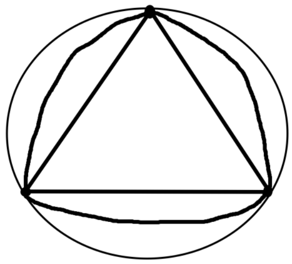
So now $c \in C(v) \cap C(w)$, contradicting Claim 2.

This contradiction shows our earlier assumption (12) was impossible. Hence $C(u) \cap C(v) \neq \emptyset$, and we can color the missing edge $e = uv$ with a common free color. Thus G is s-edge-colorable.

By induction, every loopless multigraph G satisfies $\chi'(G) \leq s = \lfloor 3\Delta(G)/2 \rfloor$. □

Example 25.6 (Shannon's Triangle). In the multigraph obtained by replacing each edge of K_3 with k parallel edges, every edge is adjacent to all others. Therefore all edges must receive distinct colors. Each vertex has degree $2k$, and

$$\chi'(G) = 3k = \frac{3}{2}\Delta(G).$$



Graph S_6

$\Delta(S_k) = k$ and $\chi'(S_k) = \left\lfloor \frac{3k}{2} \right\rfloor$.

Figure 4: Shannon's Triangle demonstrates the bound is sharp

25.3 Vizing's Theorem

Theorem 25.4 (Vizing (multigraph form)). Let G be a *loopless* multigraph. Write

$$\Delta := \Delta(G) \qquad \text{and} \qquad \mu := \mu(G) := \max_{uv} m_G(uv)$$

(where $m_G(uv)$ is the number of parallel edges between u and v). Then

$$\chi'(G) \le \Delta + \mu.$$

(If G has a loop, then $\chi'(G)$ is undefined/infinite since a loop is adjacent to itself.)

Proof. Set $k := \Delta + \mu$ and prove by induction on $|E(G)|$.

Induction setup. The case $|E(G)| = 0$ is trivial. Assume $|E(G)| \ge 1$ and pick an edge $e = xy$. By induction, $G - e$ has a proper k-edge-coloring

$$c : E(G - e) \longrightarrow [k] := \{1, 2, \dots, k\}.$$

For a vertex v, define the set of *missing colors* at v (w.r.t. c) by

$$C(v) := \{\, t \in [k] : \text{no edge incident to } v \text{ has color } t \,\}.$$

Since at most $d_{G-e}(v) \le \Delta$ colors appear at v and $k \ge \Delta + 1$, we always have $C(v) \ne \emptyset$. In particular,

$$|C(x)| \ge k - d_{G-e}(x) \ge k - (\Delta - 1) = \mu + 1.$$

If $C(x) \cap C(y) \ne \emptyset$, choose $t \in C(x) \cap C(y)$ and color e with t. So assume for contradiction that

$$C(x) \cap C(y) = \emptyset. \tag{12}$$

Alternating paths and the "if it doesn't hit x, we win" lemma. Fix colors $a, b \in [k]$. Because the coloring is proper, from any vertex there is *at most one* incident edge of color a and at most one of color b. Hence there is a unique maximal path starting at a vertex v that begins with an a-colored edge and then alternates $a, b, a, b, \dots$; call it the *a/b-path from v*.

Lemma 25.5. Let uv be an uncolored edge and let c be a proper k-edge-coloring of $G - uv$. If $a \in C(u)$ and $b \in C(v)$ and the a/b-path from v does *not* end at u, then one can modify c (by swapping a and b along that path) so that a becomes missing at *both* u and v, and then color uv with a.

So, because we are assuming G is *not* k-edge-colorable, we may use the contrapositive of Lemma 25.5 in the following form:

() Whenever we have a proper k-edge-coloring of $G - uv$ and choose $a \in C(u)$ and $b \in C(v)$, the a/b-path from v *must* end at u.

Build a Vizing fan at x. Choose any color $a \in C(x)$. Let $y_0 := y$. Build a maximal sequence of *distinct* neighbors $y_0, y_1, \dots, y_s$ of x such that for each $i \geq 1$,

$$c(xy_i) \in C(y_{i-1}).$$

(Maximal means: you cannot extend the sequence by adding a new neighbor y_{s+1} with $c(xy_{s+1}) \in C(y_s)$.)

Now define colorings $c_0, c_1, \dots, c_s$ as follows: $c_0 := c$, and for $1 \leq i \leq s$ define c_i by *rotating* colors along the fan:

$$c_i(xy_j) := c_0(xy_{j+1}) \quad (0 \leq j < i), \qquad c_i(xy_i) \text{ is undefined (i.e. } xy_i \text{ is uncolored),}$$

and $c_i(e') := c_0(e')$ for every other edge e'. By construction, c_i is a proper k-edge-coloring of $G - xy_i$: when we recolor xy_j with $c_0(xy_{j+1})$, that color was missing at y_j. Also, the set $C(x)$ is the same for all c_i (we only permute colors on edges incident to x).

Pick a missing color at the last fan vertex. Choose any color $b \in C_{c_0}(y_s)$ (missing at y_s under c_0). Then b is still missing at y_s under every c_i (we never recolor edges incident to y_s except possibly xy_s, and deleting an edge cannot *remove* a missing color).

If $b \in C(x)$, then in c_s the edge xy_s is uncolored and both endpoints miss b, so we color xy_s with b and obtain a k-edge-coloring of G, contradiction. Hence

$$b \notin C(x). \tag{13}$$

So some edge incident to x has color b under c_0; write that edge as xy_i for some $1 \leq i \leq s$. Moreover, by maximality of the fan: since b is missing at y_s, the neighbor at the other end of the b-colored edge from x must already be in the fan, so such an i exists with $1 \leq i \leq s$.

Under the rotation to c_s, the color $b = c_0(xy_i)$ is shifted one step left, so

$$c_s(xy_{i-1}) = b.$$

The alternating-path contradiction. Consider the a/b-path P from y_s in the coloring c_s. Here $a \in C(x)$ (still) and $b \in C(y_s)$, so property () applied to the uncolored edge xy_s (in c_s) implies: P must end at x.

Because a is missing at x, the final edge of P entering x must have color b. But the only b-colored edge incident to x in c_s is xy_{i-1}, so P ends with the edge $y_{i-1}x$.

Now look at the coloring c_{i-1}. In c_{i-1} the edge xy_{i-1} is uncolored. Also, by the fan property, the color $c_0(xy_i) = b$ was chosen so that it is missing at y_{i-1} under c_0, and we did not recolor any edge incident to y_{i-1} except possibly xy_{i-1}. Therefore b is missing at y_{i-1} under c_{i-1}.

Let P' be the a/b-path from y_{i-1} under c_{i-1}. The internal edges of P avoid x (since x is the endpoint), and we only changed colors on edges incident to x when going from c_{i-1} to c_s. Hence the a/b-alternating walk from y_{i-1} in c_{i-1} follows exactly the edges of P in reverse order until it reaches y_s.

But b is missing at y_s, so P' stops at y_s and does *not* reach x. This contradicts property () applied to the uncolored edge xy_{i-1} in c_{i-1} (with $a \in C(x)$ and $b \in C(y_{i-1})$).

This contradiction shows that (12) is impossible. Hence $C(x) \cap C(y) \neq \emptyset$, and we can color $e = xy$ with a common missing color. Therefore G has a proper k-edge-coloring, i.e. $\chi'(G) \leq k = \Delta + \mu$. □

Definition 25.3 (Vizing class). A simple graph G is

$$\textit{Class 1} \iff \chi'(G) = \Delta(G), \qquad \textit{Class 2} \iff \chi'(G) = \Delta(G) + 1.$$

Example 25.7 (Standard examples).
- Bipartite graphs are Class 1 (Kőnig's line coloring theorem below).
- Odd cycles C_{2t+1} are Class 2: $\Delta = 2$ but $\chi' = 3$.
- Complete graphs satisfy

$$\chi'(K_n) = \begin{cases} n-1, & n \text{ even (Class 1)}, \\ n, & n \text{ odd (Class 2)}. \end{cases}$$

25.4 Konig's Line Coloring Theorem

Theorem 25.6 (Kőnig (line coloring)). If G is bipartite, then

$$\chi'(G) = \Delta(G).$$

Proof. Let $\Delta = \Delta(G)$. The lower bound $\chi'(G) \geq \Delta$ is immediate.

For the upper bound, it suffices to show that every bipartite graph with maximum degree Δ has a proper Δ-edge-coloring.

We may assume G is Δ-regular: if not, add dummy vertices and dummy edges (still bipartite) to obtain a Δ-regular bipartite supergraph $\widehat{G}$. Any Δ-edge-coloring of $\widehat{G}$ restricts to one of G by deleting dummy edges.

Now let $\widehat{G}$ be Δ-regular with bipartition (X, Y). By Hall's Marriage Theorem, $\widehat{G}$ has a perfect matching: indeed for any $S \subseteq X$,

$$\Delta|S| = e(S, N(S)) \leq \Delta|N(S)| \quad \Longrightarrow \quad |N(S)| \geq |S|.$$

So there exists a matching M_1 saturating X, hence perfect.

Remove M_1. The remaining graph is $(\Delta - 1)$-regular and bipartite, so by the same argument it has a perfect matching M_2. Iterating, we obtain pairwise edge-disjoint perfect matchings

$$E(\widehat{G}) = M_1 \dot\cup M_2 \dot\cup \cdots \dot\cup M_\Delta.$$

Color edges in M_i with color i. Each M_i is a matching, so this is a proper Δ-edge-coloring of $\widehat{G}$, and hence of G.

Therefore $\chi'(G) \leq \Delta$, and combined with $\chi'(G) \geq \Delta$ we get $\chi'(G) = \Delta$. □

26 Planar graphs

26.1 Basics of planar graphs

Definition 26.1 (Polygonal curve). A *polygonal curve* in $\mathbb{R}^2$ is a curve obtained by concatenating finitely many straight line segments.

Definition 26.2 (Drawing of a (multi)graph). Let G be a (multi)graph. A *drawing* of G is a function

$$\varphi : V(G) \cup E(G) \longrightarrow \mathbb{R}^2$$

such that

1. for each vertex $v \in V(G)$, $\varphi(v) \in \mathbb{R}^2$ is a point;
2. if $v \neq v'$, then $\varphi(v) \neq \varphi(v')$ (distinct vertices map to distinct points);
3. for each edge $e = xy \in E(G)$, $\varphi(e)$ is a polygonal curve whose endpoints are $\varphi(x)$ and $\varphi(y)$.

(For a loop xx, the curve $\varphi(xx)$ starts and ends at $\varphi(x)$.)

Definition 26.3 (Crossing). A *crossing* in a drawing φ is a point $p \in \varphi(e) \cap \varphi(f)$ for two *distinct* edges $e \neq f$ such that p is not the image of a common endpoint of e and f. Equivalently, p is a common *internal* point of the two edge-curves.

Definition 26.4 (Planar graph and plane graph). A (multi)graph G is *planar* if it has a drawing with no crossings. A *plane graph* is a pair (G, φ) where φ is a crossing-free drawing of G (i.e. a specific planar embedding has been chosen).

Example 26.1. K_4 is planar (it has a crossing-free drawing).

K_5 and $K_{3,3}$ are not planar: no matter how you draw them in the plane, some pair of edges must cross.

Definition 26.5 (Faces). Let (G, φ) be a plane graph. A *face* of (G, φ) is a connected component of

$$\mathbb{R}^2 \setminus \varphi\big(V(G) \cup E(G)\big).$$

There is always one unbounded face, called the *outer face*.

Definition 26.6 (Length of a face). Let F be a face of a plane graph (G, φ). The *length* $\ell(F)$ is the total length of the closed walk(s) in G that trace the boundary of F. Equivalently, $\ell(F)$ is the number of edge-sides incident with F, counting multiplicity (a bridge contributes twice, once for each side).

Definition 26.7 (Dual graph). Let (G, φ) be a plane graph with face set $\mathcal{F}$. The *dual* G^* is the graph defined by

$$V(G^*) = \mathcal{F}, \qquad E(G^*) \leftrightarrow E(G),$$

where each edge $e \in E(G)$ corresponds to an edge $e^* \in E(G^*)$ joining the two faces on the two sides of e (if e borders the same face on both sides, then e^* is a loop).

Remark 26.1. The dual depends on the chosen embedding: different plane drawings of the same planar graph can yield non-isomorphic dual graphs.

Proposition 26.1 (Handshake for faces). Let (G, φ) be a plane multigraph with edge set $E(G)$ and faces $\mathcal{F} = \{F_1, \ldots, F_f\}$. Then

$$\sum_{i=1}^{f} \ell(F_i) = 2|E(G)|.$$

Proof. Traverse the boundary of each face as a closed walk in G. Each step of such a walk uses an *edge-side* (an incidence of an edge with a face). By definition, $\ell(F_i)$ counts the number of edge-sides on the boundary of F_i, with multiplicity.

Every edge e in a plane drawing has exactly two sides. If e is not a bridge, it is incident with two (not necessarily distinct) faces, contributing 1 to each. If e is a bridge, then both sides of e are incident with the same face, so e contributes 2 to that one face. In all cases, each edge contributes exactly 2 to the total sum of face-lengths. Hence $\sum_{F \in \mathcal{F}} \ell(F) = 2|E(G)|$. □

26.2 Euler's Formula

Theorem 26.2 (Euler's Formula). If (G, φ) is a connected plane multigraph with $n = |V(G)|$ vertices, $m = |E(G)|$ edges, and f faces, then

$$n - m + f = 2.$$

Proof. We induct on m (the number of edges). The claim is immediate for $m = 0$: connectedness forces $n = 1$ and $f = 1$, so $n - m + f = 1 - 0 + 1 = 2$.

Assume $m \geq 1$ and that the statement holds for all connected plane multigraphs with fewer than m edges.

Case 1: G has a non-loop edge that is not a bridge. Pick such an edge e. Deleting e keeps the graph connected (since e is not a bridge), and in a plane embedding the deletion of e merges exactly two faces into one. Thus, for $G' := G - e$ we have

$$n' = n, \qquad m' = m - 1, \qquad f' = f - 1.$$

By the induction hypothesis, $n' - m' + f' = 2$, and substituting gives

$$n - m + f = n' - (m' + 1) + (f' + 1) = n' - m' + f' = 2.$$

Case 2: every non-loop edge of G is a bridge. Then the underlying simple graph is a tree (with possibly some loops attached). Contract (or delete) a bridge edge e connecting two distinct vertices. Contracting a bridge preserves the number of faces (bridges do not lie on a cycle, so they do not separate two distinct faces), and it reduces both n and m by 1. For $G' := G/e$,

$$n' = n - 1, \qquad m' = m - 1, \qquad f' = f.$$

Again, by induction $n' - m' + f' = 2$, hence

$$n - m + f = (n' + 1) - (m' + 1) + f' = n' - m' + f' = 2.$$

In either case, Euler's formula holds for G, completing the induction. □

Corollary 26.3 (Euler for k components)**.** If a plane multigraph G has k connected components and n, m, f denote its numbers of vertices, edges, and faces (in a fixed plane drawing), then

$$n - m + f = k + 1.$$

Proof. Let the components be $G_1, \ldots, G_k$, with parameters (n_i, m_i, f_i). Applying Theorem 26.2 to each component,

$$n_i - m_i + f_i = 2 \qquad (1 \le i \le k).$$

Summing gives $\sum_i n_i - \sum_i m_i + \sum_i f_i = 2k$.

Now $\sum_i n_i = n$ and $\sum_i m_i = m$. For faces, each component has its own outer face, but in the union of all components these k outer faces merge into a single global outer face. Hence

$$\sum_{i=1}^{k} f_i = f + (k - 1).$$

Substituting into the summed Euler equalities yields

$$n - m + \big(f + (k - 1)\big) = 2k,$$

so $n - m + f = k + 1$. □

Theorem 26.4 (Edge bounds for planar graphs)**.** Let G be a simple planar graph with $n \ge 3$ vertices and m edges.

(i) $m \le 3n - 6$.

(ii) If G is triangle-free, then $m \le 2n - 4$.

Proof. Fix a plane embedding of G with f faces. Since G is simple, the boundary walk of any face has length at least 3 (no loops, no parallel edges), so

$$\ell(F) \ge 3 \quad \text{for all faces } F.$$

Summing over faces and using Proposition 26.1 gives

$$3f \le \sum_F \ell(F) = 2m.$$

Euler's formula (Theorem 26.2) gives $f = 2 - n + m$, hence

$$3(2 - n + m) \le 2m \implies 6 - 3n + 3m \le 2m \implies m \le 3n - 6,$$

proving (i).

For (ii), if G is triangle-free then every face has length at least 4, so $4f \leq \sum_F \ell(F) = 2m$. Again substitute $f = 2 - n + m$:

$$4(2 - n + m) \leq 2m \implies 8 - 4n + 4m \leq 2m \implies m \leq 2n - 4.$$

□

Corollary 26.5. The graphs K_5 and $K_{3,3}$ are nonplanar.

Proof. We use the planar edge bounds from Theorem 26.4.

1) K_5 is nonplanar. The complete graph K_5 is simple with

$$n = 5, \qquad m = \binom{5}{2} = 10.$$

If K_5 were planar, then by Theorem 26.4(i) we would have

$$m \leq 3n - 6 = 3 \cdot 5 - 6 = 9,$$

but $m = 10 > 9$, a contradiction. Hence K_5 is not planar.

2) $K_{3,3}$ is nonplanar. The complete bipartite graph $K_{3,3}$ is simple with

$$n = 6, \qquad m = 3 \cdot 3 = 9.$$

Moreover $K_{3,3}$ is bipartite, so it contains no odd cycle, in particular no triangle. Thus $K_{3,3}$ is triangle-free.

If $K_{3,3}$ were planar, then by Theorem 26.4(ii) we would have

$$m \leq 2n - 4 = 2 \cdot 6 - 4 = 8,$$

but $m = 9 > 8$, a contradiction. Hence $K_{3,3}$ is not planar. □

Theorem 26.6 (Cycles ↔ bonds in the dual)**.** Let G be a connected *plane* multigraph (i.e. a planar multigraph with a fixed crossing-free drawing), and let G^* be its dual. For an edge set $X \subseteq E(G)$ write $X^* := \{e^* : e \in X\} \subseteq E(G^*)$.

Then an edge set $C \subseteq E(G)$ is the edge set of a cycle in G (equivalently: $G[C]$ is connected and 2-regular) if and only if C^* is a *bond* in G^* (a minimal nonempty edge cut).

Theorem 26.7. Let G be a *connected plane multigraph* (i.e. G is embedded in the plane with no crossings), and let G^* be its planar dual. The following are equivalent:

(A) G is bipartite.

(B) Every face of G has even length.

(C) G^* is Eulerian (equivalently, every vertex of G^* has even degree).

Proof. **(A)⇒(B).** If G is bipartite, then every closed walk in G has even length (colors must alternate). The boundary of each face is a closed walk in G (possibly repeating vertices/edges in a multigraph), hence its length is even.

(B)⇔(C). Vertices of G^* correspond to faces of G. Fix a face F of G and let v_F be the corresponding vertex of G^*. Each time an edge e of G appears on the boundary walk of F, the dual edge e^* is incident with v_F. (If e is a bridge of G, then e^* is a loop in G^* and contributes 2 to $\deg_{G^*}(v_F)$, exactly matching the fact that e appears twice on the boundary walk of the unique incident face.) Therefore

$$\deg_{G^*}(v_F) = \ell(F) \qquad \text{for every face } F \text{ of } G.$$

Hence all face-lengths are even iff all degrees in G^* are even. Since G is connected, G^* is connected, so "all degrees even" is exactly the Eulerian condition.

(B)⇒(A). Assume every face has even length. If G were not bipartite, it would contain an odd cycle C. As a simple closed curve in the plane, C has an interior region. Let $F_1, \ldots, F_t$ be the faces of G lying strictly inside C, and let E_{int} be the set of edges of G that lie strictly inside C (i.e. not on C).

Every interior edge is incident with two interior faces, hence is counted twice in the sum $\sum_{i=1}^{t} \ell(F_i)$, while each edge of C is incident with exactly one interior face, hence is counted once. Thus

$$\sum_{i=1}^{t} \ell(F_i) = 2|E_{\text{int}}| + |C|.$$

The left-hand side is even (sum of even numbers), and $2|E_{\text{int}}|$ is even, so $|C|$ must be even. This contradicts that C is an odd cycle. Therefore G has no odd cycle, hence G is bipartite.

We have shown (A)⇒(B)⇔(C) and (B)⇒(A), so all three conditions are equivalent. □

26.3 Outerplanar graphs

Definition 26.8 (Outerplanar graphs). A (multi)graph G is *outerplanar* if it has a planar drawing in which *every vertex lies on the boundary of the outer face*. A specific such drawing is called an *outerplane* embedding of G.

Proposition 26.8 (2-connected outerplanar ⇒ Hamiltonian cycle). If G is a 2-connected outerplanar graph, then G has a Hamilton cycle.

Proof. Fix an outerplane embedding of G, so all vertices lie on the boundary of the outer face. Let W be the closed walk obtained by traversing the boundary of the outer face.

We claim that W is in fact a *simple* cycle (no vertex repeats). Indeed, if some vertex v appeared at least twice on this boundary walk, then the edges of G incident with v would split locally into at least two separate "intervals" of W, and the portion of G drawn between two consecutive appearances of v would be separated from the rest by v. That makes v a cut-vertex, contradicting that G is 2-connected.

Hence the boundary of the outer face is a cycle containing every vertex of G, i.e. a Hamilton cycle. □

Example 26.2. K_4 and $K_{2,3}$ are planar but not outerplanar. (Equivalently: they are exactly the two forbidden minors for outerplanarity; we will prove this characterization later.)

Proposition 26.9 (Outerplanar graphs have a low-degree vertex)**.** Every simple outerplanar graph G has a vertex of degree at most 2. In fact, if $|V(G)| \geq 4$ then G has two *nonadjacent* vertices x, y with $d_G(x) \leq 2$ and $d_G(y) \leq 2$ (so in particular $xy \notin E(G)$).

Proof. Take an outerplane embedding of G and add edges one-by-one (without creating crossings) until no further edge can be added while keeping all vertices on the outer face. This produces a *maximal outerplanar* supergraph H on the same vertex set.

Facts about maximal outerplanar graphs:

- the boundary of the outer face is a Hamilton cycle C containing all vertices (by the previous proposition applied to each 2-connected block; maximal outerplanar graphs are 2-connected when $|V| \geq 3$);
- every bounded face of H is a triangle (otherwise we could add a chord inside a face and contradict maximality).

Now consider the *weak dual* of H: its vertices are the bounded faces of H, with two faces adjacent if they share an edge. Because all bounded faces are triangles and the embedding is outerplane, this weak dual is a tree. Therefore it has a leaf face F.

A leaf face F shares exactly one edge with another bounded face, so in H the triangle F has exactly one *internal* edge and its other two edges lie on the outer Hamilton cycle C. Let v be the vertex of F opposite the internal edge. Then v is incident only with those two outer-face edges, so $d_H(v) = 2$.

Doing the same with a different leaf face gives another vertex u with $d_H(u) = 2$. These two degree-2 vertices can be chosen nonadjacent (in a triangulated polygon, distinct ears are never adjacent unless the graph is very small; for $|V| \geq 4$ we can pick two leaf faces whose ear vertices are not consecutive on C).

Finally, since G is a subgraph of H on the same vertices,

$$d_G(v) \leq d_H(v) = 2 \qquad \text{and} \qquad d_G(u) \leq d_H(u) = 2,$$

and if u, v are nonadjacent in H then certainly $uv \notin E(G)$. □

26.4 Maximal planar graphs

Definition 26.9 (Maximal planar)**.** A simple planar graph G is *maximal planar* if G is planar and for every nonedge $uv \notin E(G)$, the graph $G + uv$ is nonplanar. Equivalently: G is planar and you cannot add any new edge without destroying planarity.

Definition 26.10 (Triangulation)**.** A *triangulation* (or *maximal plane graph*) is a plane embedding of a simple graph in which every face (including the outer face) has boundary a triangle.

Remark 26.2 (Plane vs. sphere)**.** A plane embedding is equivalent to an embedding on the sphere S^2: adding a single "point at infinity" to $\mathbb{R}^2$ turns the outer face into an ordinary face. This is why triangulations are usually defined as "every face is a triangle," outer face included.

Theorem 26.10. Let G be a connected simple plane graph on $n \geq 3$ vertices and m edges. The following are equivalent:

(A) $m = 3n - 6$.

(B) G is a triangulation (every face is a triangle).

(C) G is maximal planar.

Proof. We use two standard facts for connected plane graphs:

$$\sum_F \ell(F) = 2m \qquad \text{and} \qquad n - m + f = 2,$$

where $\ell(F)$ is the length of face F and f is the number of faces.

(B)⇒(A). If every face is a triangle then $\ell(F) = 3$ for all faces, so

$$2m = \sum_F \ell(F) = 3f \quad \Rightarrow \quad f = \frac{2m}{3}.$$

Plug into Euler:

$$n - m + \frac{2m}{3} = 2 \quad \Rightarrow \quad m = 3n - 6.$$

(A)⇒(B). In any simple plane graph, every face has length at least 3, so

$$2m = \sum_F \ell(F) \ \geq \ 3f.$$

Using Euler, $f = 2 - n + m$, hence

$$2m \geq 3(2 - n + m) = 6 - 3n + 3m \quad \Rightarrow \quad m \leq 3n - 6.$$

If $m = 3n - 6$ holds, then all inequalities above must be equalities. In particular $\sum_F \ell(F) = 3f$, which forces $\ell(F) = 3$ for every face F. So G is a triangulation.

(B)⇒(C). Assume G is a triangulation and suppose we try to add a new edge uv. In any plane embedding, a new edge can be drawn without crossings only if u and v lie on the boundary of a common face. But every face boundary is a triangle, so any two vertices on that boundary are already adjacent. Hence no new edge can be added while preserving planarity, i.e. G is maximal planar.

(C)⇒(B). Assume G is maximal planar. If some face F has boundary length $\ell(F) \geq 4$, then the boundary walk of F contains two nonconsecutive vertices on the face boundary; adding the diagonal between them inside F creates no crossings, producing a larger planar graph. This contradicts maximality. Therefore every face has length 3, so G is a triangulation.

Combining the implications yields (A)⇔(B)⇔(C). □

26.5 Kuratowski and Wagner's Theorems

Definition 26.11 (Subdivision of an edge). Let G be a graph and let $e = uv \in E(G)$. A *subdivision* of e is obtained by deleting e and replacing it by a path

$$u - x_1 - x_2 - \cdots - x_t - v$$

where the new vertices $x_1, \dots, x_t$ are distinct and have degree 2 in the resulting graph. More generally, a *subdivision of* G is any graph obtained from G by subdividing some (possibly zero) edges.

Remark 26.3. Subdividing edges preserves planarity: if G has a plane drawing, then we can place the new degree-2 vertices along the drawn curve for each subdivided edge, so no crossings are created.

Theorem 26.11 (Kuratowski, 1930). A graph G is planar if and only if G contains no subgraph that is a subdivision of K_5 or $K_{3,3}$.

Definition 26.12 (Edge contraction). Let G be a simple graph and let $e = xy \in E(G)$. The *contraction* of e is the graph G/xy obtained by deleting x and y, adding a new vertex $x * y$, and joining $x * y$ to every vertex in

$$N_G(x) \cup N_G(y) \setminus \{x, y\}.$$

(If this produces parallel edges, we keep only one so that G/xy remains simple.)

Definition 26.13 (Minor). A graph H is a *minor* of a graph G if H can be obtained from G by a sequence of the following operations:

1. deleting a vertex,
2. deleting an edge,
3. contracting an edge.

We write $H \preceq G$ to indicate that H is a minor of G.

Proposition 26.12. If G is planar, then contracting any edge of G yields a planar graph.

Proof. Take a plane drawing of G. Contracting an edge xy can be realized by "shrinking" the drawn curve for xy to a point, merging x and y into a single vertex. This operation does not create crossings, so the resulting graph is still planar. □

Theorem 26.13 (Wagner). A graph G is planar if and only if G has no minor isomorphic to K_5 or $K_{3,3}$.

Proof. We prove both directions by contrapositive.

($\Rightarrow$ contrapositive.) Assume that G contains a minor H with $H \cong K_5$ or $H \cong K_{3,3}$. Since K_5 and $K_{3,3}$ are nonplanar, H is nonplanar. By Observation 26.12, planarity is preserved under edge contractions, and it is clearly preserved under deletions. Therefore, if G were planar then every minor of G would be planar. Since H is a minor of G and H is not planar, it follows that G is not planar.

($\Leftarrow$ contrapositive.) Assume that G is not planar. By Kuratowski's Theorem, G contains a subgraph G' that is a subdivision of some $H \in \{K_5, K_{3,3}\}$. But if G' is a subdivision of H, then contracting each subdivided path back to a single edge produces H. Hence H is a minor of G'. Finally, every subgraph of G is a minor of G (just delete the other vertices and edges), so H is a minor of G. Thus G contains K_5 or $K_{3,3}$ as a minor. □

Theorem 26.14 (Outerplanar characterization)**.** A graph G is outerplanar if and only if G contains no subdivision of K_4 and no subdivision of $K_{2,3}$.

Proof. Let $G^\star$ be the graph obtained from G by adding one new vertex y adjacent to every vertex of G (so y is a universal vertex for G).

Lemma 26.15. $G^\star$ is planar if and only if G is outerplanar.

Proof of Lemma. ($\Rightarrow$) Suppose $G^\star$ is planar. Take a planar embedding of $G^\star$ and choose the outer face so that it contains y on its boundary. Delete the vertex y and its incident edges. In the remaining drawing of G, every vertex of G lies on the boundary of the outer face (because each was adjacent to y), hence G is outerplanar.

($\Leftarrow$) Suppose G is outerplanar. Then G has an embedding in which all vertices lie on the boundary of the outer face. Place a new vertex y in that outer face and draw edges from y to every vertex of G inside the outer face without crossings. This yields a planar embedding of $G^\star$. □

By the claim, G is outerplanar $\iff$ $G^\star$ is planar. By Kuratowski's Theorem, $G^\star$ is planar if and only if it contains no subdivision of K_5 and no subdivision of $K_{3,3}$.

($\Rightarrow$) Assume G is outerplanar. Then $G^\star$ is planar, so $G^\star$ has no subdivision of K_5 or $K_{3,3}$. If G contained a subdivision of K_4, then (since y is adjacent to every vertex of G) adding y would turn it into a subdivision of K_5 inside $G^\star$, contradiction. Similarly, if G contained a subdivision of $K_{2,3}$, then adding y (adjacent to all branch vertices) yields a subdivision of $K_{3,3}$ in $G^\star$, contradiction. Hence G contains no subdivision of K_4 or $K_{2,3}$.

($\Leftarrow$) Assume G is not outerplanar. Then by the lemma $G^\star$ is not planar, so by Kuratowski $G^\star$ contains a subdivision of K_5 or of $K_{3,3}$.

If $G^\star$ contains a subdivision of K_5, then deleting the universal vertex y from that subdivision leaves a subdivision of K_4 contained in $G = G^\star - y$.

If $G^\star$ contains a subdivision of $K_{3,3}$, then y cannot lie in one of the two bipartition classes (since y is adjacent to *every* vertex of G), so deleting y leaves a subdivision of $K_{2,3}$ inside G.

Thus G contains a subdivision of K_4 or $K_{2,3}$, completing the contrapositive.

Therefore G is outerplanar if and only if it contains no subdivision of K_4 and no subdivision of $K_{2,3}$. □

Theorem 26.16 (Fáry)**.** Every planar graph has a plane drawing in which every edge is drawn as a straight line segment (i.e., a straight-line embedding).

Definition 26.14 (Convex embedding)**.** A plane drawing of a planar graph is a *convex embedding* if every face (including the outer face) is bounded by a convex polygon.

Theorem 26.17 (Tutte, Convex Embedding Theorem)**.** Let G be a 3-connected planar graph. Then G admits a convex straight-line embedding in the plane. Equivalently, G has a plane straight-line drawing in which every face boundary is a convex polygon.

Example 26.3. Petersen graph is not planar

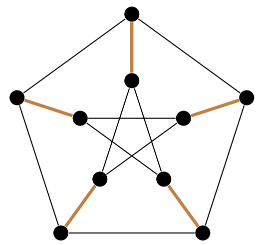


Figure 5: The Petersen graph. Brown spokes are the edges to be contracted in Proof (2).

We present three independent proofs that the Petersen graph is not planar.

1. Subdivision of $K_{3,3}$. The Petersen graph contains a subdivision of $K_{3,3}$. Indeed, removing every other vertex on the outer 5-cycle and following the spokes into the inner 5-cycle produces three vertices on one side and three on the other, with internally disjoint paths between them. Since $K_{3,3}$ is nonplanar and subdivisions preserve nonplanarity, the Petersen graph is nonplanar.
2. After contracting the entire inner 5-cycle, the resulting graph becomes a copy of K_5. Since edge contraction preserves nonplanarity, the original Petersen graph is also nonplanar.
3. Edge-counting argument. The Petersen graph has

$$n = 10, \qquad m = 15.$$

If a simple planar graph has $n \geq 3$, then $m \leq 3n - 6$. For $n = 10$, this gives

$$m \leq 3(10) - 6 = 24.$$

This inequality is satisfied, so we refine the bound: the Petersen graph is cubic and has girth 5. A planar graph of girth g satisfies

$$m \leq \frac{g}{g-2}(n-2).$$

For $g = 5$, this becomes

$$m \leq \frac{5}{3}(10-2) = \frac{5}{3} \cdot 8 = \frac{40}{3} < 14.$$

But the Petersen graph has $m = 15$, violating the planar girth bound. Thus it cannot be planar.

26.6 Four Color Theorem

The *Four Color Problem* (posed in 1852) asked whether every planar map can be colored with at most four colors so that adjacent regions receive different colors. Over the next century it helped *create* modern graph theory: repeated "proofs" (notably Kempe's 1879 argument) were later found to have subtle gaps, and the search for a correct proof forced people to invent sharper structural ideas about planar graphs.

In the mid-1970s, Kenneth Appel and Wolfgang Haken at UIUC developed a strategy that reduces the theorem to checking a *finite* list of unavoidable configurations in any minimal counterexample, and then verifying (by computer) that each such configuration is *reducible* (cannot occur in a minimal counterexample). The computer part was not a cute afterthought: the argument required checking a large family of cases that was beyond reasonable hand-verification.

When Appel and Haken announced success in 1976, it became the first mainstream, high-profile computer-assisted proof, and it triggered a real (and frankly healthy) fight about verification, standards of certainty, and how the community should audit proofs that are longer than a human can reliably re-check end-to-end. A famous insult-compliment attributed to His Majesty, Paul Erdős, who spent time in Urbana-Champaign: after seeing the computer-heavy proof, he supposedly said,

> "I'm not sure God has this proof in *The Book*. Maybe He has it in *The File*."

(Translation for those unfortunate *epsilons* who do not understand fluent Erdős: "It's correct and historic, but it is not what I would call *beautiful*")

After the announcement, skeptics didn't just attack the math; they attacked the **physics** of the computer. Critics argued: "How do you know a cosmic ray didn't hit a vacuum tube in the IBM 360 and flip a bit from `FALSE` to `TRUE`?". This sounds like a joke, but it was a serious philosophical debate. Appel and Haken had to argue that the probability of a cosmic ray flipping a specific bit in a way that preserved the program's syntax *and* produced a valid-looking Reducible" result was statistically lower than the probability of a human making a typo in a 500-page handwritten proof. They essentially introduced the concept of Probabilistic Proof" to mainstream topology. The math world hated it.

Theorem 26.18 (Four Color Theorem (Appel–Haken, 1976))**.** Every planar graph is 4-colorable. Equivalently, for every planar graph G,

$$\chi(G) \leq 4.$$

Remark 26.4 (Map-coloring form)**.** In any plane embedding, the faces can be colored with at most 4 colors so that any two faces sharing an edge receive different colors. Equivalently, the dual graph G^* satisfies $\chi(G^*) \leq 4$.

A bit of UIUC four-color lore

There is a story that floats around about the *very* end of the Appel–Haken computer check.

Disclaimer for the overly responsible reader: I have only ever heard this as oral lore; it may be apocryphal or embellished. Please treat this as confidential department heritage folklore: the kind of thing you hear in a hallway. Please do *not* cite this in a journal article.

The legend goes like this:

June 21, 1976: the last night of the computer run. They were basically done, but not *all the way* done. Appel and Haken were in the Digital Computer Laboratory (DCL) watching an IBM 360 grind through the final batch of configurations. For each configuration, the program prints `TRUE` (good, proof lives) or `FALSE` (bad, proof collapses).

By this point, the machine had been chewing through case checks for something like **1,200 hours** of CPU time. The administration, allegedly, had begun to ask questions of the form "how much longer?" and "what is this *for*?" and "why does DCL have its own weather system?" It was not subtle that the proof had acquired a second adversary: not a planar counterexample, but the **electric bill**. In 1976, computer time was billed internally at hundreds of dollars per hour.

At around this point (so the story says), Haken's daughter, Dorothea Blostein (who wasn't just a visitor, but a paid research assistant and UIUC undergrad), visited the lab over at DCL and showed up with champagne.

The vibe was tense. Logic dictates you wait for the output. But according to the story, they looked at the machine, looked at the months of successful runs, and decided: "Screw it. The algorithm hasn't failed yet." They popped the cork *while* the mainframe was still processing, toasting their victory before the final verification was actually complete.

If the computer had spit out `FALSE` ten minutes later, . . .

Luckily, it didn't.

26.7 Five Color Theorem

In 1879 Alfred Kempe published what was widely accepted as a proof of the Four Color Theorem, and for about a decade the problem was treated as essentially closed. Then, in 1890, Percy Heawood found a genuine structural flaw in Kempe's argument, resurrecting map coloring as a live problem.

Here is the part that makes the Five Color Theorem historically satisfying: Heawood did not merely break Kempe's proof and walk away. Kempe's ideas were powerful but not powerful enough to reach four. He extracted from it what *was* salvageable, repaired the method, and proved a clean, unconditional replacement: five colors alwa

Theorem 26.19 (Heawood's Five Color Theorem)**.** Every planar graph is properly vertex-colorable with at most 5 colors. Equivalently, if G is planar then $\chi(G) \le 5$.

Proof. We may assume G is connected (color components independently with the same palette). We prove the statement by induction on $n := |V(G)|$.

Lemma 1 (Low-degree vertex). Every simple planar graph has a vertex of degree at most 5.

Proof of Lemma 1. If $n \le 2$ the claim is obvious. For $n \ge 3$, Euler's formula gives $n - m + f = 2$. In a simple planar graph every face has length at least 3, so $\sum_F \ell(F) \ge 3f$. Also $\sum_F \ell(F) = 2m$ (each edge borders two faces), hence $2m \ge 3f$ and therefore $f \le \frac{2m}{3}$. Substitute into Euler:

$$2 = n - m + f \le n - m + \frac{2m}{3} = n - \frac{m}{3},$$

so $m \le 3n - 6$. Thus the average degree is

$$\frac{1}{n} \sum_{v \in V(G)} d(v) = \frac{2m}{n} \le \frac{2(3n-6)}{n} = 6 - \frac{12}{n} < 6.$$

Hence some vertex has degree ≤ 5. □

Induction step. Let $v \in V(G)$ be a vertex with $d(v) \leq 5$ (exists by Lemma 1), and set $H := G - v$. Then H is planar with $n-1$ vertices, so by induction H has a proper 5-coloring.

We try to extend this coloring to v.

Case 1: $d(v) \leq 4$. At most 4 colors appear on $N(v)$, so one of the 5 colors is missing there. Color v with a missing color and we are done.

Case 2: $d(v) = 5$. If fewer than 5 distinct colors appear on $N(v)$, we again color v with a missing color. So assume the five neighbors of v use *all five* colors.

Fix a planar embedding and list the neighbors in cyclic order around v:

$$v_1, v_2, v_3, v_4, v_5.$$

Relabel the colors so that

$$c(v_i) = i \quad \text{for } i = 1, 2, 3, 4, 5.$$

For colors $i \neq j$, let H_{ij} be the subgraph of H induced by vertices colored i or j. Each component of H_{ij} is bipartite and therefore has exactly two possible i/j-colorings (swap i and j within that component). Such components are called *Kempe components*, and paths inside them are *i–j Kempe chains*.

Claim 2. If v_1 and v_3 lie in different components of H_{13}, then we can recolor H (preserving propriety) so that color 1 is missing on $N(v)$.

Proof. Swap colors 1 and 3 on the Kempe component of H_{13} that contains v_1. This keeps the coloring proper (we only permute colors inside a 1/3-induced component), and it changes $c(v_1)$ from 1 to 3 while $c(v_3)$ stays 3 (since v_3 is in a different component). Hence no neighbor of v has color 1 anymore, so we can color v with 1.

So assume *the opposite*: v_1 and v_3 are in the *same* component of H_{13}. Then there exists a 1–3 Kempe chain P_{13} in H joining v_1 to v_3.

Now consider the closed curve in the embedding formed by the edges vv_1, vv_3, and the chain P_{13}. By the Jordan curve theorem, this closed curve separates the plane into an "inside" and an "outside" region. Because the neighbors occur around v in the order v_1, v_2, v_3, v_4, v_5, the vertices v_2 and v_4 lie on *different* sides of that curve. In particular, any path in the drawing from v_2 to v_4 that avoids v must cross the curve.

Claim 3. v_2 and v_4 lie in different components of H_{24}.

Proof. If they were in the same component of H_{24}, there would be a 2–4 Kempe chain P_{24} in H joining v_2 to v_4. This chain uses only vertices of colors 2 and 4, so it is disjoint from the 1–3 chain P_{13}. Also it avoids v because $v \notin H$. But then P_{24} would give a curve in the embedding connecting v_2 to v_4 without crossing the separating curve built from $vv_1 \cup P_{13} \cup vv_3$, contradicting planarity. Hence no such P_{24} exists, i.e. v_2 and v_4 are in different components of H_{24}.

By Claim 3, we may swap colors 2 and 4 on the H_{24}-component containing v_2. This keeps the coloring proper and changes $c(v_2)$ from 2 to 4 while leaving $c(v_4) = 4$ unchanged. Therefore color 2 is now missing on $N(v)$, so we color v with 2.

In all cases the 5-coloring of H extends to a 5-coloring of G. This completes the induction.

26.8 Discharging method

The following proposition illustrates the discharging method, which was used in the proof of Four Color Theorem

Proposition 26.20. Every planar graph G contains either

1. a vertex of degree at most 4, or
2. a vertex v of degree 5 that has at least two neighbors of degree at most 6.

Proof. Assume for contradiction that G is a planar graph with neither (1) nor (2). In particular,

$$\delta(G) \geq 5, \tag{14}$$

and every 5-vertex has *at most one* neighbor of degree ≤ 6.

Fix a plane embedding of G and add edges inside faces until every face is a triangle; let T be the resulting plane triangulation on the *same* vertex set. (Adding edges preserves planarity and does not decrease any vertex degree.)

We claim that T also has neither (1) nor (2). Indeed, if T had a vertex of degree ≤ 4, then the same vertex in G would also have degree ≤ 4 (degrees only increase), contradicting (14). If T had a vertex v of degree 5, then v also has degree 5 in G (since degrees only increase), so no added edge is incident with v; hence $N_T(v) = N_G(v)$. Moreover, any neighbor with degree ≤ 6 in T also has degree ≤ 6 in G. Thus (2) in T would imply (2) in G, contradiction. Therefore it suffices to derive a contradiction for the triangulation T.

So from now on assume G itself is a plane triangulation with neither (1) nor (2).

Step 1: Initial charge and total charge. Define the initial charge

$$\mu(x) := d(x) - 6 \qquad (x \in V(G)).$$

Since G is a plane triangulation with $n := |V(G)| \geq 3$, we have $e(G) = 3n - 6$ (standard: $2e = \sum_F \ell(F) = 3f$ and $n - e + f = 2$). Hence

$$\sum_{x \in V(G)} \mu(x) = \sum_{x \in V(G)} (d(x) - 6) = 2e(G) - 6n = (6n - 12) - 6n = -12.$$

So the total initial charge is *negative*.

Step 2: Discharging rule (redistribution). Redistribute charge along edges as follows:

> Every vertex x with $d(x) \geq 7$ sends $\frac{1}{4}$ to each neighbor y with $d(y) = 5$. Vertices of degree ≤ 6 send nothing.

Let $\mu'(x)$ be the final charge of x after this redistribution. Because we only move charge between vertices, the total charge is unchanged:

$$\sum_{x \in V(G)} \mu'(x) = \sum_{x \in V(G)} \mu(x) = -12.$$

We will show that, under our counterexample assumptions, *every* vertex has $\mu'(x) \geq 0$, which contradicts the total being -12.

Step 3: Verify $\mu'(x) \geq 0$ for every vertex. We consider cases by $d(x)$.

Case 1: $d(x) \le 4$. This cannot occur by assumption (otherwise we already have outcome (1)).

Case 2: $d(x) = 5$. By assumption, x has *at most one* neighbor of degree ≤ 6. Since $d(x) = 5$, it follows that x has *at least four* neighbors of degree ≥ 7, and each such neighbor sends $\frac{1}{4}$ to x. Thus

$$\mu'(x) \ \ge \ (5-6) + 4 \cdot \frac{1}{4} \ = \ 0.$$

Case 3: $d(x) = 6$. Vertex x neither sends nor receives charge (it is not a 5-vertex and does not send). Hence

$$\mu'(x) = \mu(x) = 6 - 6 = 0.$$

Case 4: $d(x) \ge 8$. In the worst case, x sends $\frac{1}{4}$ to *every* neighbor (i.e. all neighbors have degree 5), so x sends at most $\frac{1}{4}d(x)$ in total. Therefore

$$\mu'(x) \ \ge \ (d(x) - 6) - \frac{1}{4}d(x) \ = \ \frac{3}{4}d(x) - 6 \ \ge \ \frac{3}{4} \cdot 8 - 6 \ = \ 0.$$

Case 5: $d(x) = 7$. Write the neighbors of x in their cyclic order around x in the embedding:

$$N(x) = \{u_1, u_2, \ldots, u_7\},$$

where $u_i u_{i+1} \in E(G)$ (indices mod 7), since G is a triangulation. Let t be the number of 5-neighbors among the u_i. We claim $t \le 4$.

If $t \ge 5$, then on the 7-cycle $u_1 u_2 \cdots u_7 u_1$ there must exist three consecutive vertices of degree 5 (because with only two non-5 vertices, you cannot separate five 5-vertices from forming a length-3 consecutive block). So for some i, the vertices u_{i-1}, u_i, u_{i+1} all have degree 5. Then u_i is a 5-vertex with two neighbors u_{i-1} and u_{i+1} of degree $5 \le 6$, which is exactly outcome (2), contradicting our assumption. Hence indeed $t \le 4$.

Therefore x sends at most $t \cdot \frac{1}{4} \le 4 \cdot \frac{1}{4} = 1$ total charge, and so

$$\mu'(x) \ \ge \ (7-6) - 1 \ = \ 0.$$

We have shown $\mu'(x) \ge 0$ for every vertex x. Summing gives $\sum_x \mu'(x) \ge 0$, contradicting $\sum_x \mu'(x) = -12$. This contradiction proves that our initial assumption was false, so G must contain either a vertex of degree at most 4 or a 5-vertex with at least two neighbors of degree at most 6. □

27 Ramsey Theory

Ramsey Theory constitutes a profound segment of combinatorial mathematics, showing that complete disorder is an impossibility within sufficiently large systems. It investigates the conditions under which a structure, no matter how chaotic or random its organization, must inevitably contain a substructure exhibiting a high degree of regularity.

The fundamental question Ramsey Theory poses is: "How many elements must a system possess to guarantee the existence of a specific monochromatic property?" This generalizes the Pigeonhole Principle to structural collisions within partitions of graphs and sets.

27.1 Graph Ramsey Theory

Definition 27.1 (2-edge-coloring and monochromatic copy). A *2-edge-coloring* of a graph G is a map $c : E(G) \to \{\text{red}, \text{blue}\}$. A subgraph $H \subseteq G$ is *monochromatic* if all its edges have the same color.

Remark 27.1. The coloring is not necessarily a proper edge coloring

Definition 27.2 (Ramsey number $R(s,t)$). For integers $s, t \geq 2$, the *Ramsey number* $R(s,t)$ is the smallest n such that every red/blue coloring of the edges of K_n contains either a red K_s or a blue K_t.

Theorem 27.1 (Ramsey). Every red/blue coloring of the edges of K_6 contains a monochromatic triangle. Equivalently, $R(3,3) = 6$.

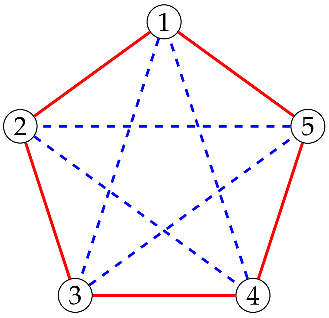
Proof. Let the vertices of K_6 be $\{v, 1, 2, 3, 4, 5\}$ and consider the 5 edges incident to v. By the pigeonhole principle, at least 3 of these edges have the same color.

WLOG, assume v is joined by *red* edges to three vertices, say a, b, c. Now look at the triangle on $\{a, b, c\}$:

- If any of ab, bc, ca is red, say ab is red, then vab is a red triangle.
- If none of ab, bc, ca is red, then all three are blue, so abc is a blue triangle.

In either case there is a monochromatic triangle. Hence every 2-edge-coloring of K_6 forces a monochromatic K_3. □

Proposition 27.2. There exists a red/blue coloring of $E(K_5)$ with *no* monochromatic triangle. Hence $R(3,3) > 5$.

Proof. 

□

Theorem 27.3.

$$R(p, q) \leq R(p-1, q) + R(p, q-1).$$

Proof. Let

$$N = R(p-1, q), \qquad M = R(p, q-1),$$

and consider any red/blue coloring of the edges of the complete graph

$$K_{N+M}.$$

Choose a vertex x. Among the $N + M - 1$ edges incident to x, the Pigeonhole Principle implies that x has at least N red edges or at least M blue edges.

Case 1: x has at least N red incident edges. Let S be the set of neighbors of x joined to x by red edges; then $|S| \geq N = R(p-1, q)$. By the definition of $R(p-1, q)$, the induced subgraph on S contains either

1. a red $(p-1)$-clique, or
2. a blue q-clique.

If it contains a blue q-clique, we are done. If it contains a red $(p-1)$-clique with vertices $\{v_1, \ldots, v_{p-1}\}$, then adding x yields a red p-clique, since all edges xv_i are red.

Thus in this case there is a monochromatic K_p (red) or K_q (blue).

Case 2: x has at least M blue incident edges. Let T be the set of neighbors of x joined to x by blue edges; then $|T| \geq M = R(p, q-1)$. By the definition of $R(p, q-1)$, the induced subgraph on T contains either

1. a red p-clique, or
2. a blue $(q-1)$-clique.

If it contains a red p-clique, we are done. If it contains a blue $(q-1)$-clique with vertices $\{u_1, \ldots, u_{q-1}\}$, then adding x produces a blue q-clique, since all edges xu_i are blue.

Thus in this case as well there is a monochromatic K_p (red) or K_q (blue).

Since every red/blue coloring of K_{N+M} produces a red K_p or a blue K_q, we conclude that

$$R(p, q) \leq N + M = R(p-1, q) + R(p, q-1).$$

□

Corollary 27.4.

$$R(p,q) \le \binom{p+q-2}{p-1}.$$

Proof. A 2-edge-coloring of K_p either has an edge of the second color or does not, so

$$R(p,2) = p.$$

Hence the upper bound holds with equality when q (or p) is 2. This provides the basis for induction on $p+q$.

By Theorem 27.1 and the induction hypothesis,

$$R(p,q) \le R(p-1,q) + R(p,q-1) \le \binom{p+q-3}{p-2} + \binom{p+q-3}{p-1}.$$

By Pascal's Identity,

$$\binom{p+q-3}{p-2} + \binom{p+q-3}{p-1} = \binom{p+q-2}{p-1}.$$

Therefore,

$$R(p,q) \le \binom{p+q-2}{p-1}.$$

□

27.2 Erdős lower bound for diagonal Ramsey numbers

To prove a *lower bound* on $R(p,p)$, we want to exhibit *one* coloring of K_n with *no* monochromatic K_p. Constructing such a coloring explicitly is hard, so we do the classic Erdős probabilistic method proof: pick a *random* coloring and prove it has the desired property with positive probability. If something happens with positive probability, then it happens for at least one outcome, so an appropriate coloring *exists*.

Theorem 27.5 (Erdős, 1947)**.**

$$R(p,p) \;\ge\; \frac{1}{e\sqrt{2}}\, p\, 2^{p/2}(1-o(1)).$$

Proof. Let the edges of K_n be colored independently at random, where each edge is assigned the color red or blue with probability $\frac{1}{2}$. The expected number of red p-cliques is

$$\mathbb{E}\big[\#\{\text{red } p\text{-cliques}\}\big] \;=\; \binom{n}{p}\left(\frac{1}{2}\right)^{\binom{p}{2}}.$$

Similarly, the expected number of blue p-cliques is

$$\mathbb{E}\big[\#\{\text{blue } p\text{-cliques}\}\big] \;=\; \binom{n}{p}\left(\frac{1}{2}\right)^{\binom{p}{2}}.$$

Hence the expected number of monochromatic p-cliques is

$$\mathbb{E}\big[\#\{\text{monochromatic } p\text{-cliques}\}\big] \;=\; \binom{n}{p}\, 2\left(\frac{1}{2}\right)^{\binom{p}{2}}.$$

If

$$\binom{n}{p} 2 \left(\frac{1}{2}\right)^{\binom{p}{2}} < 1,$$

then there exists a coloring with *no* monochromatic p-clique. Hence such an n is a lower bound on $R(p,p)$.

Using the estimate

$$\binom{n}{p} < \left(\frac{ne}{p}\right)^p,$$

it suffices to have

$$\left(\frac{ne}{p}\right)^p < 2^{\binom{p}{2}-1}.$$

Now we solve this inequality for n. Taking natural logs,

$$p\left(\ln n + 1 - \ln p\right) < \left(\binom{p}{2} - 1\right)\ln 2 = \left(\frac{p(p-1)}{2} - 1\right)\ln 2.$$

Divide by p:

$$\ln n < \frac{p-1}{2}\ln 2 - \frac{\ln 2}{p} - 1 + \ln p.$$

Exponentiating gives

$$n < \exp\left(\ln p - 1\right)\,\exp\left(\frac{p-1}{2}\ln 2\right)\,\exp\left(-\frac{\ln 2}{p}\right) = \frac{p}{e}\,2^{(p-1)/2}\,2^{-1/p}.$$

Since $2^{(p-1)/2} = 2^{p/2}/\sqrt{2}$ and $2^{-1/p} = 1 - o(1)$ as $p \to \infty$, we get

$$n < (1 - o(1))\,\frac{p}{e\sqrt{2}}\,2^{p/2}.$$

Therefore, for such n there exists a 2-coloring of $E(K_n)$ with no monochromatic K_p, so

$$R(p,p) \geq (1 - o(1))\,\frac{p}{e\sqrt{2}}\,2^{p/2}.$$

□

27.3 General Ramsey's Theorem

Definition 27.3 (r-uniform k-coloring)**.** Let S be a set and $r \in \mathbb{N}$. Write

$$\binom{S}{r} := \{\, A \subseteq S : |A| = r \,\}$$

for the family of r-subsets of S. A *k-coloring of* $\binom{S}{r}$ is a map

$$f : \binom{S}{r} \longrightarrow [k] := \{1, 2, \ldots, k\}.$$

Definition 27.4 (Homogeneous set). Given a k-coloring $f : \binom{S}{r} \to [k]$, a subset $T \subseteq S$ is *i-homogeneous* (or *monochromatic of color i*) if

$$f(A) = i \qquad \text{for every } A \in \binom{T}{r}.$$

Equivalently, all r-subsets of T receive the same color i.

Definition 27.5 (Ramsey number). Fix $k, r \in \mathbb{N}$ and target sizes $p_1, \dots, p_k \in \mathbb{N}$. We write

$$R(p_1, \dots, p_k; r)$$

for the minimum N such that for every k-coloring $f : \binom{[N]}{r} \to [k]$ there exists some $i \in [k]$ and an i-homogeneous set $T \subseteq [N]$ with $|T| = p_i$.

Remark 27.2. When $r = 2$, a k-coloring of $\binom{[N]}{2}$ is just an edge-coloring of the complete graph K_N with k colors, and a homogeneous set is a monochromatic clique. For example,

$$R(3, 3; 2) = 6.$$

Also note that adding a parameter 2 does nothing in the graph case:

$$R(3, 3, 2; 2) = R(3, 3; 2),$$

since any two vertices already form a monochromatic K_2 in whatever color the edge has.

Theorem 27.6 (Ramsey, 1930). For all $k, r \in \mathbb{N}$ and all $p_1, \dots, p_k \in \mathbb{N}$,

$$R(p_1, \dots, p_k; r) < \infty.$$

Proof. We prove finiteness by induction on the uniformity r; for fixed r we use a secondary induction on $\sum_i p_i$ (equivalently, on $\sum_i (p_i - 1)$).

Base case $r = 1$. A k-coloring of $\binom{[N]}{1}$ is just a coloring of the points $1, 2, \dots, N$. By the pigeonhole principle, if

$$N = (p_1 - 1) + \cdots + (p_k - 1) + 1 = p_1 + \cdots + p_k - k + 1,$$

then some color class has size at least p_i, giving an i-homogeneous set of size p_i. Hence $R(p_1, \dots, p_k; 1) \le N$.

Vacuous case. If some $p_i < r$, then any set T of size p_i is automatically i-homogeneous, because $\binom{T}{r} = \varnothing$. Thus $R(p_1, \dots, p_k; r) \le p_i < \infty$. So we may assume from now on that $r \ge 2$ and $p_i \ge r$ for all i.

Inductive step. Fix $r \ge 2$ and assume all Ramsey numbers for uniformity $r - 1$ are finite (induction on r), and also that all numbers of the form $R(p_1, \dots, p_i - 1, \dots, p_k; r)$ are finite (secondary induction, since the sum of the parameters decreases).

For each $i \in [k]$, define

$$q_i := R(p_1, \dots, p_{i-1},\ p_i - 1,\ p_{i+1}, \dots, p_k;\ r),$$

and set

$$M := R(q_1, \dots, q_k;\ r-1), \qquad N := M+1.$$

We claim $R(p_1, \dots, p_k; r) \le N$.

Let $f : \binom{[N]}{r} \to [k]$ be an arbitrary k-coloring. Write $v := N$. Define a new coloring of $(r-1)$-subsets of $[N-1]$ by

$$f' : \binom{[N-1]}{r-1} \to [k], \qquad f'(A) := f(A \cup \{v\}).$$

By the definition of $M = R(q_1, \dots, q_k; r-1)$, there exists a color $i \in [k]$ and a set $T \subseteq [N-1]$ with $|T| = q_i$ such that T is i-homogeneous for f', i.e.

$$\forall A \in \binom{T}{r-1} \ : \ f(A \cup \{v\}) = f'(A) = i.$$

Now restrict the original coloring f to $\binom{T}{r}$. Because $|T| = q_i$ and $q_i = R(p_1, \dots, p_i - 1, \dots, p_k; r)$, by the definition of q_i one of the following occurs:

1. there is a color $j \ne i$ and a j-homogeneous set $U \subseteq T$ with $|U| = p_j$ (done); or
2. there is an i-homogeneous set $S \subseteq T$ with $|S| = p_i - 1$.

In case (2), we claim $S \cup \{v\}$ is i-homogeneous for f and has size p_i. Indeed, take any r-subset $B \in \binom{S\cup\{v\}}{r}$. If $v \notin B$, then $B \in \binom{S}{r}$, so $f(B) = i$ because S is i-homogeneous in f. If $v \in B$, write $B = A \cup \{v\}$ where $A \in \binom{S}{r-1} \subseteq \binom{T}{r-1}$; then $f(B) = f(A \cup \{v\}) = i$ by the i-homogeneity of T for f'. Thus all r-subsets of $S \cup \{v\}$ have color i, as required.

Therefore every k-coloring of $\binom{[N]}{r}$ produces an i-homogeneous set of size p_i for some i, so $R(p_1, \dots, p_k; r) \le N < \infty$. □

Remark 27.3. The proof yields the bound

$$R(p_1, \dots, p_k; r) \ \le \ 1 + R\big(R(p_1 - \delta_{1i}, \dots, p_k - \delta_{ki}; r) \text{ for } i \in [k];\ r-1\big),$$

i.e. if $q_i = R(p_1, \dots, p_i - 1, \dots, p_k; r)$ then

$$R(p_1, \dots, p_k; r) \le 1 + R(q_1, \dots, q_k; r-1).$$

27.4 Erdős–Szekeres on points in convex position

Theorem 27.7 (Erdős–Szekeres, 1935). For each integer $m \ge 3$, there exists an integer $N(m)$ such that every set of $N(m)$ points in the plane in general position contains m points in convex position.

Proof. A set of points is in general position if no three lie on a line. The theorem states that if the total number of points is sufficiently large, then one can always find m of them forming the vertex set of a convex polygon.

The classical bounds satisfy

$$2^{m-2} \le N(m) \le \binom{2m-4}{m-2} \le 2^{2m},$$

and it remains an open problem whether $N(m) = 2^{m+o(m)}$.

The proof is based on the following two lemmas.

Lemma 27.8. Every set of five points in general position contains four points in convex position.

Proof of lemma 1. Given any five points in general position, either one of them lies inside the convex hull of the other four, or none of them does. In the first case, the remaining four already form a convex quadrilateral. In the second case, all five points lie on the boundary of their convex hull, and any four consecutive vertices again give four points in convex position. In either case four convex-position points exist. □

Lemma 27.9. If every four-element subset of a point set P of size m is in convex position, then the entire set P is in convex position.

Proof of lemma 2. Assume for contradiction that P is not in convex position. Then some point of P lies strictly inside the convex hull of the remaining points. Let v be such a point. Consider any three points on the hull that form a triangle containing v. Together with v they form four points that fail to be in convex position, contradicting the assumption that every four-subset is convex. Therefore all points must lie on the boundary of the convex hull, and P is in convex position. □

To prove the theorem, consider $N = R(m, 5)$, the Ramsey number for red-blue colorings of m- and 5-subsets. Given any N points in general position, color each 4-element subset red if it is in convex position and blue otherwise. By the choice of N there is either a red m-set or a blue 5-set.

If there is a blue 5-set, Claim 1 shows this is impossible, since every five points contain four convex-position points. Thus no blue 5-set exists. Therefore the Ramsey argument produces a red m-set, meaning every 4-subset of these m points is in convex position. By Claim 2 the entire set of m points is in convex position. This establishes the existence of $N(m)$. □

27.5 Schur's Theorem

Theorem 27.10 (Schur, 1916)**.** Given $k > 0$, there exists an integer s_k such that every k-coloring of $\{1, 2, \ldots, S_k\}$ has a monochromatic solution (x, y, z) to

$$x + y = z.$$

Proof. Let $r_k = R_k(3)$ be the Ramsey number such that every k-edge-coloring of K_{r_k} contains a monochromatic triangle. We show that $s_k \le r_k$.

Let

$$f : \{1, 2, \ldots, r_k - 1\} \longrightarrow [k]$$

be a k-coloring of the integers $\{1, 2, \ldots, r_k - 1\}$. We use f to define a k-edge-coloring f' of K_{r_k} as follows. Let $V(K_{r_k}) = \{1, 2, \ldots, r_k\}$, and for each edge ij define

$$f'(ij) = f(|i - j|).$$

By the definition of r_k, the coloring f' contains a monochromatic triangle with vertices a, b, c. Without loss of generality, assume

$$a < b < c.$$

Define

$$x = b - a, \qquad y = c - b, \qquad z = c - a.$$

Then x, y, z are positive integers, and

$$f(x) = f(|b - a|) = f'(ab), \qquad f(y) = f(|c - b|) = f'(bc), \qquad f(z) = f(|c - a|) = f'(ac).$$

Since the edges ab, bc, and ac form a monochromatic triangle under f', we have

$$f(x) = f(y) = f(z).$$

Finally, by construction,

$$x + y = (b - a) + (c - b) = c - a = z.$$

Thus (x, y, z) is a monochromatic solution to the equation $x + y = z$.

□

28 Probabilistic Method

The probabilistic method is an indirect strategy: to prove that a combinatorial object *exists*, we define a random process that produces such objects and show that the probability of success is *positive*. If Pr(success) > 0, then at least one successful outcome must exist.

28.1 Two basic tools: the first moment and alteration

Lemma 28.1 (First moment method)**.** Let X be a nonnegative integer-valued random variable. If $\mathbb{E}[X] < 1$, then $\Pr(X = 0) > 0$. In particular, there exists an outcome with $X = 0$.

Proof. If $X \geq 1$ then $X \geq 1 \cdot \mathbf{1}_{\{X \geq 1\}}$, hence $\mathbb{E}[X] \geq \Pr(X \geq 1)$. If $\mathbb{E}[X] < 1$, then $\Pr(X \geq 1) < 1$, so $\Pr(X = 0) > 0$. □

Lemma 28.2 (Alteration principle)**.** Suppose a random structure contains "bad" substructures, and let X be the number of bad substructures. If we can destroy *every* bad substructure by deleting at most one vertex per bad substructure, then there exists an outcome in which, after deleting at most X vertices, no bad substructure remains. In particular, there exists an outcome with at least $n - X$ vertices and no bad substructure.

Proof. Given an outcome, delete one vertex from each bad substructure (choosing arbitrarily). Then every bad substructure is destroyed, and we deleted at most X vertices. □

28.2 Hypergraphs and Property B (2-colorability)

Definition 28.1 (Hypergraph, k-uniform)**.** A *hypergraph* is a pair $H = (V, E)$ where V is a vertex set and $E \subseteq 2^V$ is a family of subsets called *hyperedges*. It is *k-uniform* if every edge has size k.

Definition 28.2 (Proper 2-coloring / Property B)**.** A *proper 2-coloring* of a hypergraph $H = (V, E)$ is a map $\varphi : V \to \{\text{red}, \text{blue}\}$ such that no edge is monochromatic (i.e. every $e \in E$ contains at least one red and at least one blue vertex). A hypergraph is *2-colorable* if it admits such a coloring. This property is also called *Property B*.

Definition 28.3. Let $f(k)$ be the minimum number of edges in a *non*-2-colorable k-uniform hypergraph.

Theorem 28.3 (Erdős)**.** For $k \geq 2$,

$$2^{k-1} \leq f(k) \leq \binom{2k-1}{k}.$$

Proof. **Upper bound.** Let V be a set of size $2k-1$, and let E be the set of all k-subsets of V. Then $|E| = \binom{2k-1}{k}$. In any red/blue coloring of V, one color class has size at least k, so it contains a monochromatic k-subset, which is an edge of H. Hence this hypergraph is not 2-colorable, so $f(k) \le \binom{2k-1}{k}$.

Lower bound. Let $H = (V, E)$ be k-uniform with $|E| = m$. Color each vertex independently red/blue, each with probability 1/2. For a fixed edge e (with $|e| = k$),

$$\Pr(e \text{ is monochromatic}) = \Pr(e \text{ all red}) + \Pr(e \text{ all blue}) = 2 \cdot 2^{-k} = 2^{1-k}.$$

Let X be the number of monochromatic edges. By linearity of expectation,

$$\mathbb{E}[X] = m \cdot 2^{1-k}.$$

If $m < 2^{k-1}$ then $\mathbb{E}[X] < 1$, so by Lemma 28.1 there exists a coloring with $X = 0$, i.e. with no monochromatic edge. Thus any non-2-colorable k-uniform hypergraph must satisfy $m \ge 2^{k-1}$. □

Remark 28.1 (Sharper, more recent bounds). The true order of magnitude of $f(k)$ is still subtle. There are significant improvements known on both sides, for example

$$f(k) \;\le\; (1 + o(1)) \cdot C\, k^2 2^k \qquad \text{and} \qquad f(k) \;\ge\; c\, 2^k \cdot \frac{\sqrt{k}}{\log k}$$

for absolute constants $C, c > 0$. (These require deeper ideas beyond the basic first-moment argument.)

Theorem 28.4 (Alteration-method bound for $R(k, k)$). For every n and k,

$$R(k, k) \;>\; n - \binom{n}{k} 2^{1-\binom{k}{2}}.$$

Consequently, choosing $n = (1 - o(1))\frac{k}{e\sqrt{2}} 2^{k/2}$ yields

$$R(k, k) \;>\; (1 - o(1))\frac{k}{e\sqrt{2}} 2^{k/2}.$$

Proof. Randomly 2-color the edges of K_n (each edge independently red/blue with probability 1/2). For each k-set $S \in \binom{[n]}{k}$, let A_S be the event that $K_n[S]$ is monochromatic. A fixed S spans $\binom{k}{2}$ edges, so

$$\Pr(A_S) = 2 \cdot \left(\tfrac{1}{2}\right)^{\binom{k}{2}} = 2^{1-\binom{k}{2}}.$$

Let $X = \sum_S \mathbf{1}_{A_S}$ be the number of monochromatic K_k's. By linearity of expectation,

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

Now *alter* the graph by deleting one vertex from each monochromatic K_k. After deleting at most X vertices, the remaining induced subgraph has no monochromatic K_k. Therefore there exists a coloring with at least $n - X$ vertices and no monochromatic K_k, so $R(k, k) > n - X$. Taking expectations gives

$$R(k, k) \;>\; n - \mathbb{E}[X] \;=\; n - \binom{n}{k} 2^{1-\binom{k}{2}}.$$

The stated asymptotic choice of n comes from making the expected number of monochromatic K_k's smaller than n (and then optimizing using $\binom{n}{k} \approx (ne/k)^k$). □

28.3 Lovász Local Lemma (LLL) when union bound is too weak

Definition 28.4 (Dependency graph viewpoint). Let $A_1, \dots, A_N$ be events. A graph D on $[N]$ is a *dependency graph* if each A_i is independent of the collection $\{A_j : j \notin N_D[i]\}$ (i.e. independent of all events outside its closed neighborhood).

Theorem 28.5 (Symmetric Lovász Local Lemma). Suppose $\Pr(A_i) \le p$ for all i, and each A_i depends on at most d others (i.e. has degree $\le d$ in some dependency graph). If

$$e\,p\,(d+1) \le 1,$$

then

$$\Pr\left(\bigcap_{i=1}^{N} \overline{A_i}\right) > 0.$$

Remark 28.2. Compare with the union bound: $\Pr(\bigcup A_i) \le \sum \Pr(A_i)$. LLL is what you use when the A_i are not independent but are "locally dependent": each bad event only interacts with a bounded neighborhood of other bad events.

28.4 Spencer's LLL proof idea for $R(k,k)$

Theorem 28.6 (Spencer, via LLL). For

$$n = (1+o(1))\frac{\sqrt{2}}{e}\,k\,2^{k/2},$$

there exists a red/blue edge-coloring of K_n with no monochromatic K_k. Equivalently, $R(k,k) > n$.

Proof sketch with the key bookkeeping. Color edges of K_n independently red/blue with probability $1/2$. For each k-set S, let A_S be the event that S spans a monochromatic K_k. As before,

$$p := \Pr(A_S) = 2^{1-\binom{k}{2}}.$$

Two events A_S and A_T are independent whenever K_S and K_T share no edges, which happens iff $|S \cap T| \le 1$. Thus A_S depends only on sets T with $|S \cap T| \ge 2$. A crude bound on the number of such T is obtained by choosing the intersection size:

$$d \;\le\; \sum_{i=2}^{k} \binom{k}{i}\binom{n-k}{k-i} \;\le\; \binom{k}{2}\binom{n}{k-2} \;=\; O\!\left(k^2\,\frac{n^{k-2}}{(k-2)!}\right).$$

Now apply Theorem 28.5. The condition $e\,p\,(d+1) \le 1$ becomes, up to lower-order factors, an inequality of the form

$$e \cdot 2^{1-\binom{k}{2}} \cdot \left(k^2\,\frac{n^{k-2}}{(k-2)!}\right) \;\lesssim\; 1,$$

and plugging $n = (1 + o(1))\frac{\sqrt{2}}{e}k2^{k/2}$ makes this true. Hence with positive probability none of the A_S occur, i.e. no monochromatic K_k exists in the coloring, so $R(k,k) > n$. □

28.5 Erdős: large girth and large chromatic number

Definition 28.5. The *girth* $g(G)$ of a graph G is the length of its shortest cycle (take $g(G) = \infty$ if G is acyclic). The *independence number* $\alpha(G)$ is the maximum size of an independent set.

Theorem 28.7 (Erdős)**.** For all integers $g \geq 3$ and $k \geq 2$, there exists a graph G with

$$g(G) \geq g \qquad \text{and} \qquad \chi(G) \geq k.$$

Proof. We build G from a random graph and then *alter* it.

Step 1: choose a random graph. Let $G \sim G(n,p)$, meaning each edge appears independently with probability p. We choose

$$p = n^{t-1} \qquad \text{where } 0 < t < \frac{1}{g}.$$

(So p is small: sparse enough to keep short cycles rare, but not so small that huge independent sets become likely.)

Step 2: short cycles are rare. Fix $j \in \{3, 4, \ldots, g\}$. The number of (labeled) j-cycles is at most n^j, and each specific j-cycle appears with probability p^j. Thus the expected number X_j of j-cycles satisfies

$$\mathbb{E}[X_j] \leq n^j p^j = n^{jt}.$$

Let $X = \sum_{j=3}^{g} X_j$ be the number of cycles of length $\leq g$. Then

$$\mathbb{E}[X] \leq \sum_{j=3}^{g} n^{jt} = o(n),$$

since $jt < 1$ for every $j \leq g$. By Markov,

$$\Pr(X \geq n/2) \leq \frac{2\mathbb{E}[X]}{n} \to 0,$$

so for large n there exists a realization with $X < n/2$.

Step 3: large independent sets are unlikely. Fix an integer r. The probability that a fixed r-set is independent is $(1-p)^{\binom{r}{2}} \leq e^{-p\binom{r}{2}}$. By the union bound,

$$\Pr(\alpha(G) \geq r) \leq \binom{n}{r}(1-p)^{\binom{r}{2}} \leq \left(\frac{en}{r}\right)^r \exp\left(-p\binom{r}{2}\right).$$

Choose

$$r = \left\lceil \frac{4 \ln n}{p} \right\rceil.$$

Then $p\binom{r}{2} \asymp (\ln n)^2/p$ dominates the $r\ln(en/r)$ term, and the RHS tends to 0 as $n \to \infty$. Hence for large n there exists a realization with $\alpha(G) < r$.

Step 4: alter to kill short cycles, and count colors. Pick a realization of G for which simultaneously $X < n/2$ and $\alpha(G) < r$. Delete one vertex from each cycle of length $\le g$. This removes all cycles of length $\le g$, so the resulting graph G' satisfies $g(G') \ge g$. We deleted at most $X < n/2$ vertices, hence $|V(G')| \ge n/2$. Also, deleting vertices cannot increase α, so $\alpha(G') \le \alpha(G) < r$. Finally,

$$\chi(G') \ \ge \ \frac{|V(G')|}{\alpha(G')} \ \ge \ \frac{n/2}{r}.$$

Since $r \sim 4(\ln n)/p = 4(\ln n)\, n^{1-t}$, the ratio $(n/2)/r$ tends to ∞ with n. For n large enough we have $\chi(G') \ge k$. Thus G' has girth at least g and chromatic number at least k. □

28.6 Markov, Chebyshev, and the second moment method

Theorem 28.8 (Markov's Inequality)**.** Let $X \ge 0$ be a random variable and let $a > 0$. Then

$$\Pr(X \ge a) \ \le \ \frac{\mathbb{E}[X]}{a}.$$

Proof. Since $X \ge a\mathbf{1}_{\{X \ge a\}}$, taking expectations gives $\mathbb{E}[X] \ge a \ \Pr(X \ge a)$. □

Theorem 28.9 (Chebyshev's Inequality)**.** Let X be a random variable with finite variance. Then for every $t > 0$,

$$\Pr\left(|X - \mathbb{E}[X]| \ge t\right) \ \le \ \frac{\mathrm{Var}(X)}{t^2}.$$

Proof. Apply Markov to the nonnegative random variable $(X - \mathbb{E}[X])^2$:

$$\Pr\left((X - \mathbb{E}[X])^2 \ge t^2\right) \le \frac{\mathbb{E}\left[(X - \mathbb{E}[X])^2\right]}{t^2} = \frac{\mathrm{Var}(X)}{t^2}.$$

□

Lemma 28.10 (Second moment method / Paley–Zygmund (useful form))**.** If $X \ge 0$ and $\mathbb{E}[X^2] < \infty$, then

$$\Pr(X > 0) \ \ge \ \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]}.$$

Proof. By Cauchy–Schwarz,

$$\mathbb{E}[X] = \mathbb{E}\left[X\mathbf{1}_{\{X>0\}}\right] \le \sqrt{\mathbb{E}[X^2]\ \Pr(X > 0)}.$$

Rearrange. □

28.7 Caro–Wei proof of Turán

Theorem 28.11 (Caro–Wei bound)**.** For every graph G,

$$\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{d(v)+1}.$$

Proof. Take a uniformly random ordering (permutation) π of $V(G)$. Let S be the set of vertices that appear before all their neighbors in π. Then S is independent (two adjacent vertices cannot both be first among the two). For a fixed vertex v, among the $d(v)+1$ vertices in $\{v\} \cup N(v)$, each is equally likely to be the earliest in π, so

$$\Pr(v \in S) = \frac{1}{d(v)+1}.$$

Thus by linearity,

$$\mathbb{E}[|S|] = \sum_v \Pr(v \in S) = \sum_v \frac{1}{d(v)+1}.$$

Since $\alpha(G) \geq |S|$ always, we get $\alpha(G) \geq \mathbb{E}[|S|]$. □

Corollary 28.12 (Jensen/Cauchy–Schwarz form)**.** If G has n vertices and m edges, then

$$\alpha(G) \geq \frac{n^2}{2m+n}.$$

Proof. By Cauchy–Schwarz applied to $a_v = d(v)+1 > 0$,

$$\sum_v \frac{1}{a_v} \geq \frac{n^2}{\sum_v a_v}.$$

Now $\sum_v (d(v)+1) = 2m+n$ and Caro–Wei gives the result. □

Theorem 28.13 (Turán)**.** If G is K_{r+1}-free on n vertices, then

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

Proof. Let $\overline{G}$ be the complement. "G is K_{r+1}-free" means $\overline{G}$ has no independent set of size $r+1$, i.e. $\alpha(\overline{G}) \leq r$.

Apply the corollary to $\overline{G}$. Since

$$e(\overline{G}) = \binom{n}{2} - e(G),$$

we get

$$\alpha(\overline{G}) \geq \frac{n^2}{2e(\overline{G})+n} = \frac{n^2}{(n^2-n)-2e(G)+n} = \frac{n^2}{n^2-2e(G)}.$$

Since $\alpha(\overline{G}) \leq r$,

$$\frac{n^2}{n^2-2e(G)} \leq r \quad \Longrightarrow \quad n^2 - 2e(G) \geq \frac{n^2}{r} \quad \Longrightarrow \quad e(G) \leq \left(1-\frac{1}{r}\right)\frac{n^2}{2}.$$

(This matches the Turán density; the exact best bound $t_r(n)$ differs only by $O(n)$.) □

28.8 Random graphs $G(n, p)$: thresholds for isolated vertices and connectivity

Definition 28.6 ($G(n, p)$). $G(n, p)$ is the random graph on vertex set $[n] = \{1, \dots, n\}$ in which each of the $\binom{n}{2}$ edges is present independently with probability $p = p(n)$.

Theorem 28.14 (Isolated vertices threshold). Let X be the number of isolated vertices in $G \sim G(n, p)$ and set

$$p = \frac{\log n + c}{n} \qquad (c \in \mathbb{R} \text{ fixed}).$$

Then $X \Rightarrow \text{Poisson}(e^{-c})$, and in particular

$$\Pr(\text{no isolated vertices}) = \Pr(X = 0) \longrightarrow e^{-e^{-c}}.$$

Proof sketch via factorial moments. For $t \geq 1$, let $(X)_t = X(X-1)\cdots(X-t+1)$. Choose t vertices; they are all isolated iff *every* edge incident to them is absent, including edges among the t themselves. The number of forbidden edges is

$$t(n-t) + \binom{t}{2}.$$

Hence

$$\mathbb{E}[(X)_t] = (n)_t \, (1-p)^{t(n-t)+\binom{t}{2}}.$$

With $p = (\log n + c)/n$ and fixed t, one checks

$$(n)_t \sim n^t, \qquad (1-p)^{t(n-t)} \sim e^{-ptn} \sim e^{-t(\log n + c)} = n^{-t} e^{-ct},$$

and $(1-p)^{\binom{t}{2}} \to 1$. Thus $\mathbb{E}[(X)_t] \to (e^{-c})^t$, which are exactly the factorial moments of $\text{Poisson}(e^{-c})$. Therefore X converges in distribution to that Poisson law, giving $\Pr(X = 0) \to e^{-e^{-c}}$. □

Theorem 28.15 (Connectivity threshold). Let $G \sim G(n, p)$ with $p = (\log n + c)/n$ and fixed $c \in \mathbb{R}$. Then

$$\Pr(G \text{ is connected}) \longrightarrow e^{-e^{-c}}.$$

In particular, the threshold for connectivity is $p \sim (\log n)/n$.

Proof sketch: "only obstruction is isolated vertices". Clearly,

$$\Pr(G \text{ connected}) \leq \Pr(\text{no isolated vertices}) \to e^{-e^{-c}}.$$

It remains to show that

$$\Pr(\text{no isolated vertices but disconnected}) \to 0.$$

If G is disconnected and has no isolated vertices, then it has a component S with size $2 \leq s := |S| \leq n/2$. For a fixed set S of size s, the event "S is a component" implies: (i) there are no edges from S to $V \setminus S$ and (ii) $G[S]$ is connected. Thus by union bound,

$$\Pr(\exists \text{ component of size } s) \leq \binom{n}{s} \Pr(G[S] \text{ connected}) \, (1-p)^{s(n-s)}.$$

Bound $\Pr(G[S]\text{ connected})$ by spanning trees: if $G[S]$ is connected it contains some spanning tree; there are s^{s-2} trees (Cayley) and each appears with probability p^{s-1}, so

$$\Pr(G[S]\text{ connected}) \;\le\; s^{s-2}\,p^{s-1}.$$

Hence

$$\Pr(\exists\text{ component of size } s) \;\le\; \binom{n}{s} s^{s-2}p^{s-1}(1-p)^{s(n-s)}.$$

Now plug $p = (\log n + c)/n$. The factor $(1-p)^{s(n-s)} \le e^{-ps(n-s)} \le e^{-s(\log n+c)/2} = n^{-s/2}e^{-cs/2}$ (for $s \le n/2$), which kills the $\binom{n}{s} \le (en/s)^s$ term strongly, and the remaining $s^{s-2}p^{s-1}$ is at most polynomial in $\log n$ times $n^{-(s-1)}$. Summing over $s = 2, \ldots, \lfloor n/2 \rfloor$ gives a total $o(1)$. Therefore the probability of being disconnected without isolated vertices vanishes, and

$$\Pr(G\text{ connected}) \sim \Pr(\text{no isolated vertices}) \to e^{-e^{-c}}.$$

□

29 Partially Ordered Sets

29.1 Structure of Posets

Definition 29.1 (Partially ordered set (poset)). A *partially ordered set* (or *poset*) is a pair $(P, \leq)$ where P is a set and $\leq$ is a binary relation on P such that for all $x, y, z \in P$:

1. (*Reflexive*) $x \leq x$.
2. (*Antisymmetric*) If $x \leq y$ and $y \leq x$, then $x = y$.
3. (*Transitive*) If $x \leq y$ and $y \leq z$, then $x \leq z$.

We say x and y are *comparable* if $x \leq y$ or $y \leq x$. Otherwise, they are *incomparable*.

Example 29.1 (The divisibility poset). Let $P = \mathbb{Z}_{\geq 1}$. Define a relation $\preceq$ on P by

$$a \preceq b \iff a \mid b.$$

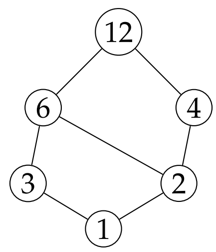
Then $(\mathbb{Z}_{\geq 1}, \preceq)$ is a poset, called the *divisibility poset*. More generally, for a fixed $n \in \mathbb{Z}_{\geq 1}$, the set of positive divisors of n,

$$D(n) = \{d \in \mathbb{Z}_{\geq 1} : d \mid n\},$$

ordered by divisibility is a finite poset.

Definition 29.2 (Hasse diagram). Let $(P, \leq)$ be a finite poset. The *Hasse diagram* of P is the directed graph obtained from the relation $\leq$ by drawing an edge $x \longrightarrow y$ exactly when $x < y$ and y *covers* x (i.e. $x < y$ and there is no $z \in P$ with $x < z < y$), and omitting all edges implied by transitivity. By convention, the diagram is usually drawn with edges pointing upward, so the arrows are often suppressed. When viewed as a directed graph, transitivity means H has no directed cycles.

Example 29.2. The following diagram is a Hasse diagram for the poset $([12], |)$ restricted to $\{2, 3, 4, 6, 12\}$.



We have $6 \preceq 12$ since $6 \mid 12$.

We do not draw an edge between 3 and 12 since 6 lies between them in the order.

29.2 Dilworth's Theorem

Definition 29.3 (Chain and antichain). Let $(P, \leq)$ be a poset. A subset $C \subseteq P$ is a *chain* if for all $x, y \in C$, we have $x \leq y$ or $y \leq x$ (i.e. every pair is comparable).

A subset $A \subseteq P$ is an *antichain* if for all distinct $x, y \in A$, neither $x \leq y$ nor $y \leq x$ (i.e. every distinct pair is incomparable).

Theorem 29.1 (Dilworth)**.** Let $(P, \leq)$ be a finite poset. Let

$$w = \max\{\, |A| : A \subseteq P \text{ is an antichain} \,\}$$

be the *width* of P. Then the minimum number of chains whose disjoint union is P (i.e. a partition of P into chains) is exactly w.

Proof. Let $n = |P|$.

(1) Easy inequality. If P is partitioned into k chains, then any antichain meets each chain in at most one element, so $|A| \leq k$ for every antichain A. Hence $w \leq k$, and therefore

$$w \leq k_{\min},$$

where $k_{\min}$ denotes the minimum number of chains in a chain-partition of P.

(2) Build a bipartite graph. Form a bipartite graph $G = (L \cup R, E)$ where $L = \{x_L : x \in P\}$ and $R = \{x_R : x \in P\}$, and put an edge

$$x_L y_R \in E \iff x < y \text{ in } P.$$

Let $\nu(G)$ be the size of a maximum matching, and $\tau(G)$ the size of a minimum vertex cover. By Kőnig's theorem for bipartite graphs,

$$\nu(G) = \tau(G).$$

Chain partitions $\iff$ matchings. We claim

$$k_{\min} = n - \nu(G).$$

Indeed, if P is partitioned into k chains, write each chain as

$$c_1 < c_2 < \cdots < c_t.$$

Add the edges $c_{1,L}c_{2,R}, c_{2,L}c_{3,R}, \ldots, c_{t-1,L}c_{t,R}$ to a set M. Across all chains this produces a matching (each element appears in at most one chosen left endpoint and at most one chosen right endpoint), and $|M| = \sum(t-1) = n - k$. Thus $\nu(G) \geq n - k$, so $k \geq n - \nu(G)$, hence $k_{\min} \geq n - \nu(G)$.

Conversely, given a matching M of size m, interpret each matched edge $x_L y_R$ as a directed link $x \to y$. Because M is a matching, every vertex has outdegree ≤ 1 and indegree ≤ 1 under these links, so the links decompose P into vertex-disjoint directed paths; each such path is a chain in P (since every link respects $<$). Adding isolated vertices as length-1 paths, we get a chain-partition with exactly $n - m$ chains. Taking $m = \nu(G)$ gives $k_{\min} \leq n - \nu(G)$. So $k_{\min} = n - \nu(G)$, as claimed.

(3) Extract a large antichain from a minimum vertex cover. Let C be a minimum vertex cover in G, so $|C| = \tau(G)$. Define

$$A = \{\, x \in P : x_L \notin C \text{ and } x_R \notin C \,\}.$$

Then A is an antichain: if $x < y$ with $x, y \in A$, the edge $x_L y_R$ exists but neither endpoint lies in C, contradicting that C covers all edges.

Let $T = \{x \in P : x_L \in C \text{ or } x_R \in C\}$. Then $|T| \le |C| = \tau(G)$ (since each vertex of C contributes to at most one element of P, but some elements might contribute *two* vertices). Hence

$$|A| \;=\; n - |T| \;\ge\; n - \tau(G).$$

Therefore $w \ge |A| \ge n - \tau(G)$.

Using $\nu(G) = \tau(G)$ and the formula for $k_{\min}$,

$$k_{\min} \;=\; n - \nu(G) \;=\; n - \tau(G) \;\le\; w.$$

Combined with (1), $w \le k_{\min}$, we conclude $k_{\min} = w$. □

29.3 LYM Inequality and Sperner's Theorem

Definition 29.4 (Boolean lattice and Sperner family)**.** Let $[n] = \{1, 2, \ldots, n\}$ and let $2^{[n]}$ be the family of all subsets of $[n]$, ordered by inclusion $\subseteq$. A family $\mathcal{F} \subseteq 2^{[n]}$ is called a *Sperner family* (or an *antichain*) if it contains no two distinct sets with one contained in the other:

$$\forall A, B \in \mathcal{F},\; A \ne B \implies \neg(A \subseteq B) \text{ and } \neg(B \subseteq A).$$

Equivalently, $\mathcal{F}$ is an antichain in the poset $(2^{[n]}, \subseteq)$.

Theorem 29.2 (LYM inequality (Lubell–Yamamoto–Meshalkin))**.** If $\mathcal{F} \subseteq 2^{[n]}$ is a Sperner family, then

$$\sum_{A \in \mathcal{F}} \frac{1}{\binom{n}{|A|}} \;\le\; 1.$$

Lubell's proof. A *maximal chain* in $(2^{[n]}, \subseteq)$ is a chain of length $n+1$ of the form

$$\varnothing = C_0 \subset C_1 \subset \cdots \subset C_n = [n], \qquad |C_k| = k.$$

Every permutation $\pi = (\pi_1, \ldots, \pi_n)$ of $[n]$ determines a maximal chain by $C_k = \{\pi_1, \ldots, \pi_k\}$, and every maximal chain arises from exactly 1 permutation in this way (up to the obvious correspondence), so choosing a uniformly random permutation is equivalent to choosing a uniformly random maximal chain.

Fix $A \subseteq [n]$ with $|A| = k$. We compute the probability that a random maximal chain contains A. The chain contains A exactly when the first k elements of the random permutation are precisely the elements of A (in some order). The number of permutations with this property is $k!(n-k)!$, hence

$$\Pr(A \text{ lies on the random chain}) \;=\; \frac{k!(n-k)!}{n!} \;=\; \frac{1}{\binom{n}{k}}.$$

Let X be the random variable counting how many sets of $\mathcal{F}$ lie on the random maximal chain. Since $\mathcal{F}$ is Sperner, a chain can meet $\mathcal{F}$ in at most one set, so $X \le 1$ always. By linearity of expectation,

$$\mathbb{E}[X] = \sum_{A \in \mathcal{F}} \Pr(A \text{ lies on the random chain}) = \sum_{A \in \mathcal{F}} \frac{1}{\binom{n}{|A|}}.$$

But also $\mathbb{E}[X] \le 1$ because $X \le 1$ surely. Therefore

$$\sum_{A \in \mathcal{F}} \frac{1}{\binom{n}{|A|}} \le 1,$$

which is the LYM inequality. □

Theorem 29.3 (Sperner). If $\mathcal{F} \subseteq 2^{[n]}$ is a Sperner family, then

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

Moreover, equality is achieved by taking $\mathcal{F}$ to be the entire middle level $\binom{[n]}{\lfloor n/2 \rfloor}$ (and also $\binom{[n]}{\lceil n/2 \rceil}$ when n is odd).

Proof. Let $M = \max_{0 \le k \le n} \binom{n}{k} = \binom{n}{\lfloor n/2 \rfloor}$. For every $A \in \mathcal{F}$ we have $\binom{n}{|A|} \le M$, hence

$$\frac{1}{\binom{n}{|A|}} \ge \frac{1}{M}.$$

Summing over $A \in \mathcal{F}$ gives

$$\sum_{A \in \mathcal{F}} \frac{1}{\binom{n}{|A|}} \ge \sum_{A \in \mathcal{F}} \frac{1}{M} = \frac{|\mathcal{F}|}{M}.$$

By the LYM inequality, the left-hand side is ≤ 1, so

$$\frac{|\mathcal{F}|}{M} \le 1 \implies |\mathcal{F}| \le M = \binom{n}{\lfloor n/2 \rfloor}.$$

Taking $\mathcal{F} = \binom{[n]}{\lfloor n/2 \rfloor}$ is clearly an antichain and has size M, so the bound is tight. □

29.4 Erdős–Ko–Rado Theorem and Katona circle method

Definition 29.5 (Intersecting family). A family $\mathcal{F} \subseteq \binom{[n]}{k}$ is *intersecting* if

$$\forall A, B \in \mathcal{F}, \qquad A \cap B \neq \emptyset.$$

A *star* is an intersecting family of the form

$$\mathcal{S}_i = \{A \in \binom{[n]}{k} : i \in A\} \quad (i \in [n]).$$

Theorem 29.4 (Erdős–Ko–Rado). Assume $n \ge 2k$ and let $\mathcal{F} \subseteq \binom{[n]}{k}$ be intersecting. Then

$$|\mathcal{F}| \le \binom{n-1}{k-1}.$$

Moreover, equality is attained by a star $\mathcal{S}_i$.

Katona's circle method. A *cyclic order* on $[n]$ means a permutation written around a circle, where rotations are identified. The number of cyclic orders is $(n-1)!$.

Fix a cyclic order σ. After labeling the positions around the circle as $1, 2, \ldots, n$ in clockwise order, define the *k-intervals* (also called *cyclic consecutive k-sets*) by

$$I_\sigma(t) \;=\; \{t, t+1, \ldots, t+k-1\} \pmod{n}, \qquad t = 1, 2, \ldots, n.$$

Let $\mathcal{I}(\sigma) = \{I_\sigma(t) : 1 \le t \le n\}$ be the set of all k-intervals in σ.

Lemma 29.5. If $n \ge 2k$ and $\mathcal{G} \subseteq \mathcal{I}(\sigma)$ is intersecting, then $|\mathcal{G}| \le k$.

Proof of Lemma. Rotate the labels so that some member of $\mathcal{G}$ is $I_\sigma(1) = \{1, 2, \ldots, k\}$. Let $I_\sigma(t) \in \mathcal{G}$ be any other interval. If $t \in \{k+1, k+2, \ldots, n-k+1\}$, then

$$I_\sigma(t) = \{t, t+1, \ldots, t+k-1\} \subseteq \{k+1, k+2, \ldots, n\}$$

and hence $I_\sigma(t) \cap I_\sigma(1) = \varnothing$, contradicting that $\mathcal{G}$ is intersecting. Therefore every t with $I_\sigma(t) \in \mathcal{G}$ must lie in

$$\{1, 2, \ldots, k\} \;\cup\; \{n-k+2, \ldots, n\}.$$

But since we already forced $1 \in \{t : I_\sigma(t) \in \mathcal{G}\}$ by rotation, we may keep the labels fixed so that 1 is the smallest start index among those in $\mathcal{G}$, which rules out $\{n-k+2, \ldots, n\}$. Hence all start indices in $\mathcal{G}$ lie in $\{1, 2, \ldots, k\}$, so $|\mathcal{G}| \le k$. □

Now do a double count. Let

$$X \;=\; \{(A, \sigma) : \sigma \text{ a cyclic order on } [n],\ A \in \mathcal{F},\ A \in \mathcal{I}(\sigma)\}.$$

Upper bound on $|X|$. For each fixed σ, the subfamily $\mathcal{F} \cap \mathcal{I}(\sigma)$ is intersecting, so by the Lemma it has size at most k. Since there are $(n-1)!$ cyclic orders,

$$|X| \;=\; \sum_\sigma |\mathcal{F} \cap \mathcal{I}(\sigma)| \;\le\; k\,(n-1)!.$$

Exact count of $|X|$ *by fixing* $A \in \mathcal{F}$. Fix a particular k-set $A \subseteq [n]$. Count cyclic orders σ for which A is consecutive, i.e. $A \in \mathcal{I}(\sigma)$. Treat A as a single block plus the $n-k$ remaining elements as singletons. Then we have $n-k+1$ objects arranged cyclically, giving $(n-k)!$ cyclic orders on the objects. Inside the block, the k elements of A can appear in any of $k!$ linear orders around the circle. Thus the number of cyclic orders with A consecutive is

$$k!\,(n-k)!.$$

Therefore

$$|X| \;=\; \sum_{A \in \mathcal{F}} k!\,(n-k)! \;=\; |\mathcal{F}|\,k!\,(n-k)!.$$

Combine the two counts.

$$|\mathcal{F}|\,k!\,(n-k)! \;=\; |X| \;\le\; k\,(n-1)!,$$

so

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

Finally, the star $\mathcal{S}_i = \{A \in \binom{[n]}{k} : i \in A\}$ is intersecting and has size $|\mathcal{S}_i| = \binom{n-1}{k-1}$, so the bound is tight. □