MATH240: Discrete Structures

Matthew He

December 17, 2024

Abstract

1 Foundation

1.1 Set Theory

A set is a collection of distinct objects, called its elements or its members.

Set inclusion

Proposition 1. *The sets*

$$A = \{x \in \mathbb{Z} : \textit{there exists } k \in \mathbb{Z}\textit{such that } x = 2k+1\}$$

and

$$B = \{x \in \mathbb{Z} : \textit{there exists } k \in \mathbb{Z}\textit{such that } x = 2l+5\}$$

are equal. (Both are different ways of describing the set of all odd integers.)

Proof. Let $x \in A$. Then there exists $k \in \mathbb{Z}$ such that $x = 2k+1$. Letting $l = k-2$, we found that l is an integer. Furthermore,

$$x = 2k+1 = 2(k-2)+5 = 2l+5.$$

We have found l such that $x = 2l+5$, so $x \in B$. This shows that $A \subseteq B$.

On the other hand, let $x \in B$. Then there exists $l \in \mathbb{Z}$ such that $x = 2l+5$. Now we let $k = l+2$; $k \in \mathbb{Z}$ since $l \in \mathbb{Z}$. We have

$$x = 2l+5 = 2(l+2)+1 = 2k+1.$$

This shows that $x \in A$, so $B \subseteq A$. This combined with the previous paragraph show $A = B$. □

Set operations

- The *union* of two sets A and B, denoted by $A \cup B$, is the set of all elements that are in A or in B.
- The *intersection* of two sets A and B, denoted by $A \cap B$, is the set of all elements that are in both A and B.
- The *difference* of two sets A and B, denoted by $A \backslash B$, is the set of all elements that are in A but not in B.
- The *symmetric difference* of two sets A and B, denoted by $A \triangle B$, is the set of all elements that are in exactly one of A and B.

1.2 Propositional Logic

Conditional and biconditional. The *conditional* logical relation IF q THEN p is denoted by $p \Rightarrow q$. Within the conditional statement, p is called the *antecedent*, which is the assumption. The *consequent* is q, which is the conclusion. The *biconditional* logical relation p IF AND ONLY IF q is denoted by $p \Leftrightarrow q$. It asserts that variables p and q are logically equivalent.

The logical formula of these two conditionals are: $p \Rightarrow q \equiv \neg p \vee q$ and $q \Rightarrow p \equiv \neg q \vee p$.

p	q	$p \Rightarrow q$	$p \Leftrightarrow q$
1	1	1	1
1	0	0	0
0	1	1	0
0	0	1	1

Table 1: Truth table for conditional and biconditional relations

A formula is called a

- *tautology* if it is true for all possible truth values of its variables.
- *contradiction* if it is false for all possible truth values of its variables.
- *contingency* if it is neither a tautology nor a contradiction.
- *satisfiable* if there is an assignment of truth values to its variables that makes it true.
- *falsifiable* if there is an assignment of truth values to its variables that makes it false.

Tutorial

1. $p \Rightarrow p$: tautology, since $\equiv \neg p \vee p = 1$
2. $(p \Rightarrow q) \Rightarrow p$: contingency, true if $p = q = 1$, false if $p = 0$.
(When $p = 0$, $p \Rightarrow q$is always true, and a $ture \Rightarrow false(p)$ is always false.)
3. $(\neg(p \wedge \neg q)) \vee p \equiv 1$: tautology.
4. $(p \Leftrightarrow q) \wedge (p \Leftrightarrow \neg q)$: contradictory.

1.3 Predicate Logic

Definition 1 (Predicate). *A predicate* is a statement containing some number of variable coming from a universe **u**.

Negating quantifiers. By De Morgan's laws, we have

$$\neg(\forall x \in U, P(x)) \equiv \exists x \in U, \neg P(x),$$

$$\neg(\exists x \in U, P(x)) \equiv \forall x \in U, \neg P(x).$$

Let's do some example. We express the statement, "There is a nonzero real numnber such that every real number is not its inverse or is negative." In the universe $\mathbb{R}$, ths formula corresponding to the statement is

$$\exists x : (x \neq 0 \wedge (\forall y : xy \neq 1 \vee y < 0)).$$

Example 1

Prove that a logical formula is satisfiable iff. its negation is falsifiable.

Proof. A logical formula is satisfiable iff there is an assignment of all the variables which makes the formula true. By definition of the negation, this assignment makes the negation of our formula false, which means the negation is falsifiable. Proof of the converse is analogous.

If a formula is true then its negation is false. Let's deriv the negation of the formula above:

$$\begin{aligned}\neg(\exists x :&(x \neq 0 \wedge (\forall y : xy \neq 1 \vee y < 0)))\\ &\equiv \forall x : \neg(x \neq 0 \wedge (\forall y : xy \neq 1 \vee y < 0))\\ &\equiv \forall x : (x = 0 \vee \neg(\forall y : xy \neq 1 \vee y < 0))\\ &\equiv \forall x : (x = 0 \vee \exists y : \neg(xy \neq 1 \vee y < 0))\\ &\equiv \forall x : (x = 0 \vee \exists y : xy = 1 \wedge y \geq 0).\end{aligned}$$

Negating a formula in predicate logic is entirely mechanical. The $\neg$ symbol moves from left to right, flipping the quantifiers and negating the predicates as it goes.

1.4 *Proof Techniques*

Proofing statement of the form $p \Rightarrow q$

We assume p is true, prove q, which is showing the case that p is true be q is false can't happen.

Proposition 2. *If n is an odd integer, then n^2 is an odd integer.*

Write the proposition in predicate formula:

$$u = \mathbb{Z}, \ \ \forall n : ((\exists k : n = 2k + 1) \Rightarrow (\exists n^2 = 2l + 1))$$

proof. Let n be ana integer. Assume that n is odd, that is, there exist k such that $n = 2k + 1$.
Then, $n^2 = (2k + 1)^2 = 2(2k^2 + 2k) + 1$. Let $l = 2k^2 + 2k$, then $n^2 = 2l + 1$, thus, it's odd.

To disprove a statement: prove its negation is true

Proposition 3. *Disprove the statement: $\exists x \forall y : x + y \neq 0$.*

Proof. To disprove, we shall prove its negation: $\forall x \exists y : x + y = 0 \equiv \neg(\exists x \forall y : x + y \neq 0)$ Let $x \in \mathbb{R}$ be given. Pick $y = -x$, then $x + y = 0$. Q.E.D

Since $p \wedge \neg p \equiv 0$

Converse and Contrapositive

Definition 2.

1. *the **converse** of $p \Rightarrow q$ is $q \Rightarrow p$*
   NB. $p \Rightarrow q \not\equiv q \Rightarrow p$

2. *the **contrapositive** of $p \Rightarrow q$ is$\neg q \Rightarrow \neg p$*
   $p \Rightarrow q \equiv \neg p \vee q \equiv \neg q \Rightarrow \neg p$

Proofs by contradiction

You assume something is true, and get something nonsense.

$$\neg p \Rightarrow 0 \equiv \neg(\neg p) \vee 0 \equiv p$$

Proposition 4. *There is no least positive rational number.*

$$u = \mathbb{Q} : \neg(\exists x : x > 0 \wedge (\forall y : y > 0 \Rightarrow x \leq y))$$

Proof. Suppose, for a contradiction that the proposition is false, that is, there exist $x \in \mathbb{Q}$ such that $x > 0$ and for all $y \in \mathbb{Q}$ with $y > 0, x \leq y$.
Let $y = \frac{x}{2}$, we have $\frac{x}{2} > 0$ since $x > 0$. Then $x \leq y$, so $x \leq \frac{x}{2}$.
Divide through by x (because x > 0) to get $1 \leq \frac{1}{2}$.
the contradiction completes the proof. Q.E.D

Tutorial: Prove or disprove:

1. $\forall n \in \mathbb{N}, \exists m \in \mathbb{N}, n + m = 0$
   False. Let's proof its negation: $\exists n \in \mathbb{N}, \forall m \in \mathbb{N}, n + m \neq 0$, it's true, e.g., we can let $n = 2$.
2. $\forall n \in \mathbb{N}, \exists m \in \mathbb{Z}, n + m = 0$
   Ture, Let $n \in \mathbb{N}$, then choose $m = -n \in \mathbb{Z}$, and we get $m + n = 0$.
3. $\forall n \in \mathbb{N}, \exists k \in \mathbb{N}, (k \geq m) \Rightarrow (k \geq 5n)$:
   The statement is equivalent to $(k < m) \vee (k \geq 5n)$.
   For any $n \in \mathbb{N}$, choose $m = 5n$, then for all $k \in \mathbb{N}$, we have that statement is true.

Case Analysis

Proposition 5. *prop. There exists irrational numbers a,b such that a^b is rational.*

1.5 Functions

Definition 3 (Surjective and injective)**.** *A function is surjective if*

$$\forall b \in B, \exists a \in A, f(a) = b$$

A function is injective if

$$\forall a_1, a_2 \in A, a_1 \neq a_2 \to f(a_1) \neq f(a_2)$$

or $f(a_1) = f(a_2) \to a_1 = a_2$.

Theorem 1. *Let $a_1, a_2, \ldots, a_n$ be a finite sequence (repeats allowed) of real numbers. Let*

$$a = \frac{1}{n} \sum_{i=1}^{n} a_i.$$

be the average value of the sequence and let m be the maximum value of the sequence attains. Then $m \geq a$.

Theorem 2. *Let A and B be finite sets with $|A| = m$ and $|B| = n$. For every function $f : A \to B$, then there is some $b \in B$ such that there are at least $\lceil m/n \rceil$ elements in A that get mapped to b.*

Corollary 1 (The pigeonhole principle)**.** *Let $n \geq 2$. If n pigeons nest in $n - 1$ holes, there is at least one hole that contains at least two pigeons. Let $f : A \to B$ be a function. If $|A| > |B|$, then f is not injective.*

Bijections. A function is called *bijective* if it is both surjective and injective.

Proposition 6. *Let A and B be finite sets. Then*

1. *there exists a bijection $f : A \to B$ if and only if $|A| = |B|$; and*
2. *if $|A| = |B|$, $f : A \to B$ then f is injective if and only if f is surjective.*

Invertibility. A function $f : A \to B$ is called *invertible* if there exists a function $g : B \to A$ such that

1. for all $a \in A$, $g(f(a)) = a$; and
2. for all $b \in B$, $f(g(b)) = b$.

If such a function g exists, it is unique, and we call it the *inverse* of f, denoted by f^{-1}.

Proposition 7. *Let $f : A \to B$ be a function. Then f is invertible if and only if f is bijective.*

Proof. First we assume that f is invertible. So there exists an inverse g of f. For each $b \in B$, setting $a = g(b)$ we have

$$f(a) = f(g(b)) = b.$$

This prove that f is surjective. To show that f is injective, suppose that $f(a_1) = f(a_2)$. By applying g to both sides, we get $g(f(a_1)) = g(f(a_2))$, whence $a_1 = a_2$, by definition of g.

Noew assume that f is bijective. We shall construct an inverse g of f. Given any $b \in B$, there is some $a \in A$ such that $f(a) = b$, from surjectivity of f, and this a is unique, by injectivity of f. So set $g(b) = a$ (and repeat this process for every $b \in B$). We have $f(g(b)) = f(a) = b$, and for every $a \in A$, by definition of g the element $g(f(a))$ is the unique element in A that gets brought to $f(a)$ by f, so has to be a itself. □

Sometimes to prove that two sets have the same cardinality, it is easier to prove that there exists a bijection between them. Here is an example:

Proposition 8. *Let X be a finite nonempty set. Let E be the set of all subsets of X with even cardinality, and let D be the set of all subsets of X with odd cardinality. Then $|E| = |D|$.*

1.6 Cardinality

We say that A and B are *equipotent* or have the same *cardinality*, if and only if there exists a bijection $f : A \to B$. We write $|A| = |B|$.

Theorem 3. *We have $|N| = |Z|$.*

It is possible to remove an infinite number of elements from $\mathbb{N}$ and end up with something still equipotent with $\mathbb{N}$.

We say that a set A is countably infinite if if there exists a bijection $f : N \to A$, that is if $|N| = |A|$. A set is countable if it is finite or countably infinite. Otherwise it is called uncountable.

Sometimes it is difficult to come up with a bijection between two sets. Instead, we would like to find an injection from B to A.

Theorem 4 (Schroder-Bernstein Theorem). *If there exists an injective function $f : A \to B$ and another injective function $g : B \to A$, then there exists a bijection $h : A \to B$.*

Theorem 5 (Fundamental Theorem of Arithmetic). *Every positive integer $n \geq 2$ can be factored into a product*

$$n = p_1^{v_1} p_2^{v_2} \cdots p_k^{v_k},$$

where the p_i are distinct prime numbers and the v_i are positive integers. This factorization is unique up to the order of the factors; that is, apart from the sequence in which the primes appear, there is only one way to factor n into primes.

Theorem 6. *If A and B are countably infinite sets, then $A \times B$ is also countably infinite.*

Corollary 2. *We have $|\mathbb{Z} \times \mathbb{Z}| = |\mathbb{N}|$.*

Theorem 7. *The set $\mathbb{Q}$ of rational numbers is countable.*

Theorem 8. *The set A of all infinite binary sequences is uncountable.*

1.7 Relations

Definition 4. *A relation on a set X is a subset $R \subseteq X \times X$. If $(a, b) \in R$, we write aRb and say that a is related to b.*

Properties of relations. Let R be a relation on a set X. We say that R is

1. *reflexive* if aRa for all $a \in X$.
2. *symmetric* if aRb implies bRa for all $a, b \in X$.
3. *transitive* if aRb and bRc implies aRc for all $a, b, c \in X$.

Definition 5 (Equivalence relations). *If R is reflexive, symmetric, and transitive, we say that R is an equivalence relation.*

If R is an equivalence relation, often we shall write $a \sim b$ to mean aRb. Sometimes we might even just say that $\sim$ is the equivalence relation.

Definition 6 (Equivalence class)**.** *Let $R \subseteq X \times X$ be an equivalence relation on a set X. Define the equivalence class of an element $a \in X$ to be the set*

$$[a] = \{b \in X : a \sim b\}.$$

Proposition 9. *Let R be an equivalence relation on A. Then*

1. *for all $x \in A$, $x \in [x]$;*
2. *for all $x, y \in A$, $x \sim y$ if and only if $[x] = [y]$;*
3. *for all $x, y \in A$, $x \not\sim y$ if and only if $[x] \cap [y] = \emptyset$.*

Definition 7 (Quotient set)**.** *Let A be a set and let $\sim$ to be an equivalence relation on A. We define the quotient of A by $\sim$ as the set of all equivalence classes of A under $\sim$:*

$$A/\sim = \{[x] : x \in A\}.$$

Proposition 10. *Let A be a set and let $\sim$ to be an equivalence relation on A. Then $A/\sim$ is a partition of A.*

2 Number Theory

2.1 Division

Proposition 11. *For all $a, b, c \in \mathbb{Z}$,*

1. *if $a|b$ then $a|bc$;*
2. *if $a|b$ and $a|c$ then $a|(b + c)$;*
3. *if $a|b$ and $b|c$ then $a|c$;*
4. *if $a|b$ and $b \neq 0$, then $|a| \leq |b|$; and*
5. *if $a|b$ and $b|a$, then $|a| = |b|$.*

Theorem 9 (Division algorithm)**.** *Let $a, b \in \mathbb{Z}$ with $b > 0$. Then there exist unique integers $q, r \in \mathbb{Z}$ such that*

$$a = bq + r \text{ and } 0 \leq r < |b|\,.$$

Euclid's algorithm. Given two nonnegative integers a and b, not both zero, this algorithm outputs $\gcd(a, b)$.

E1. If $b = 0$, output a and stop.

E2. Since $b \neq 0$, by the division algorithm we may write $a = bq + r$ with $0 \leq r < b$. Set $a \leftarrow b$ and $b \leftarrow r$, and return to step E1.

Why does this algorithm work? The following lemma clarifies the situation.

Lemma 1. *Let $a, b, q, r \in \mathbb{Z}$ be integers such that $a = bq + r$. Then* $\gcd(a, b) = \gcd(b, r)$.

Theorem 10 (Bezout's identity). *Let $a, b \in \mathbb{Z}$ be nonzero integers with greatest common divisor* $\gcd(a, b)$. *Then there exist integers s, t such that*

$$\gcd(a, b) = sa + tb.$$

Moreover, $\gcd(a, b)$ *is the smallest positive integer that can be written in an integer linear combination of a and b.*

The proof of Bezout's identity is a good example of a minimality proof.

Proposition 12. *Let a and b be nonzero integers. The set*

$$X = \{s'a + t'b : s', t' \in \mathbb{Z}\}$$

is exactly the set of multiples of $d = \gcd(a, b)$.

Proof. By Bezout's identity, there exist integers s, t such that $d = sa + tb$. First let $n \in \mathbb{Z}$ be a multiple of d, so $n = kd$ for some $k \in \mathbb{Z}$. Then we have

$$n = kd = d(sa + tb) = (ds)a + (dt)b,$$

which means that $n \in \mathbf{X}$, since $ds, dt \in \mathbb{Z}$.

Conversely, suppose that $n \in \mathbf{X}$. Then $n = s'a + t'b$ for some $s', t' \in \mathbb{Z}$. Then since d divides a and b, we can write $a = ld$ and $b = md$ for some integers $l, m \in \mathbb{Z}$. So we have

$$n = s'a + t'b = s'ld + t'md = (s'l + t'm)d,$$

which shows that $d|n$, since $s'l + t'm \in \mathbb{Z}$. □

Coprime We say that integers a and b are relatively prime or coprime if $\gcd(a, b) = 1$.

Proposition 13. *For all integers $n > 1$, n and $n + 1$ are relatively prime.*

2.2 Primes

Theorem 11. *An integer p with $p \geq 2$ is prime if and only if the only if for all $a, b \in \mathbb{N}$, $p|ab$ implies that $p|a$ or $p|b$.*

To illustrate that this theorem, we can also consider $p = 6, a = 2$ and $b = 15$. We have $p|ab$ since $6|30$, but 6 does not divide 2 or 15. By induction, the theorem can be extended to arbitrary finite products.

This relates to the definition of a prime number as a number p such that $\gcd(p, n) = 1$ for all integers n with $1 < n < p$. Since p shares no common divisors with any number less than itself (other than 1), it cannot divide a product ab unless it divides at least one of a or b. This emphasizes the fundamental property of prime numbers in relation to divisibility and greatest common divisors.

Corollary 3. *Let p be prime and n be a positive integer. If $a_1, a_2, \ldots, a_n$ are integers such that $p|a_1a_2 \cdots a_n$, then $p|a_i$ for some $1 \leq i \leq n$.*

Theorem 12 (Fundamental Theorem of Arithmetic, again). *Every integer $n \geq 2$ can be expressedc as a product*

$$n = p_1 p_2 \cdots p_k$$

where $p_1 \leq p_2 \leq \cdots \leq p_k$ are prime numbers. Furthermore, this factorization is unique.

The Fundamental Theorem of Arithmetic can be used to prove the following theorem.

Theorem 13. *Let k and n be positive integers. Then either $\sqrt[k]{n}$ is an integer or it is irrational.*

Theorem 14. *There are infinitely many prime numbers.*

Although the set of prime numbers is infinite, it does sort of get "sparser" as one heads off towards infinity. This is quantified by the following theorem.

Theorem 15 (Prime number theorem). *For $x \in \mathbb{R}$, let*

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

Then $\pi(x) \sim x/\ln x$ in the sense that

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\ln x} = 1.$$

Corollary 4. *Let n be a positive integer and let m be chosen uniformly at random from the set $\{1, 2, \ldots, n\}$. Then the probability that m is primes satisfies*

$$(\ln n)\mathbb{P}\{m \text{ prime}\} \to 1$$

as $n \to \infty$. In other word, the probability that that m is prime is asymptotically $1/\ln n$.

2.3 Modular Arithmetic

Definition 8 (Congruence). *Fix $n \geq 1$ and let $a, b \in \mathbb{Z}$. We say that a is congruent to b modulo n, if $n|a - b$. i.e., if $a - b = kn$ for some $k \in \mathbb{Z}$. We write this as $a \equiv_n b$ or $a \equiv b \pmod{n}$.*

For any fixed n, the set of all $(a, b) \in \mathbb{Z}$ with $a \equiv_n b$ is a relation on $\mathbb{Z}$.

Proposition 14. *For all fixed n, the relation $a \equiv_n b$ is an equivalence relation on the set $\mathbb{Z}$.*

Proposition 15. *Fix an integer $n \geq 2$. Let $a, b \in \mathbb{Z}$. Then $a \equiv_n b$ if and only if*

$$a \,\%\, n = b \,\%\, n.$$

This proposition is useful in practice if we want to know whether two numbers are congruent modulo n.

The proposition also implies that for all $a \in \mathbb{Z}$ and $n \geq 2$, one has $[a]_n = [a \,\%\, n]$. From the division algorithm, we know that $a \,\%\, n$ is an element in the $[0, n)$. It is equal to the integer r in that range such that we may write $a = qn + r$ for some $q \in \mathbb{Z}$. We may choose to denote the whole equivalence class by this element r. The set

$$\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}/\equiv_n = \{[0], [1], \dots, [n-1]\}$$

is called the *ring of integers modulo n*. or the *cyclic group on n elements*.

Computation in modular arithmetic

An element $a \in \mathbb{Z}$ with $a \not\equiv_n 0$ is said to be a zero divisor if there exists $b \in \mathbb{Z}$ with $b \not\equiv_n 0$ such that $ab \equiv_n 0$.

Proposition 16. *Let $a \equiv_n c$ and $b \equiv_n d$. Then*

1. $a + b \equiv_n c + d$;
2. $ab \equiv_n cd$;
3. $a^k \equiv_n c^k$ *for all* $k \in \mathbb{N}$;

Definition 9 (Inverses modulo n). *An element $a \in \mathbb{Z}$ is said to be invertible modulo n if there exists $b \in \mathbb{Z}$ such that $ab \equiv_n 1$. The element b is called the inverse of a. In fact, inverses are unique (in $\mathbb{Z}/n\mathbb{Z}$)*

Proposition 17. *Let $a, n \in \mathbb{Z}$ with $n \geq 2$. Then if $ab \equiv_n 1$, and $ac \equiv_n 1$, then $b \equiv_n c$.*

Theorem 16. *Let $a, n \in \mathbb{Z}$ with $n \geq 2$. Then:*

1. *a is invertible modulo n if and only if* $\gcd(a, n) = 1$;
2. *if a is invertible modulo n, then there is a unique $b \in \mathbb{Z}$ such that $ab \equiv_n 1$. Namely, if*
$$1 = sa + tn,$$
then we can set $b = s \,\%\, n$.

We can use this theorem to systematically find inverses of integers modulo other integers.

Proposition 18. *Let p be prime. Then*

1. *every $x \in \mathbb{Z}$ with $x \not\equiv_p 0$ is invertible modulo p;*
2. *for all $a, b \in \mathbb{Z}$ with $ab \equiv_p 0$ one has $a \equiv_p 0$ or $b \equiv_p 0$.*

Solving congruences modulo n

Proposition 19. *Let p be prime. Then $a^2 \equiv a \pmod{p}$ if and only if a is either congruent to 0 or 1 modulo p.*

Proposition 20. *Let p be prime and let $a \not\equiv 0 \pmod{p}$ (so that a is invertible modulo p). Then $a \equiv a^{-1} \pmod{p}$ if and only if a is either congruent to 1 or -1 modulo p.*

Theorem 17 (Fermat's little theorem)**.** *Let p be prime and let a be an integer. If $a \not\equiv -p0$, then*

$$a^{p-1} \equiv 1 \pmod{p}.$$

Lemma 2. *For all prime number p, the integer $(p-1)!$ is congruent to -1 modulo p.*

Theorem 18 (Wilson's theorem)**.** *For all integers $n \geq 2$, the integer n is prime if and only if*

$$(n-1)! \equiv -1 \pmod{n}.$$

3 Graph Theory

3.1 Definitions

A graph is a pair $G = (V, E)$, where V is a nonempty set and

$$E \subseteq \{u, v : u, v \in V, u \neq v\}.$$

Degrees and k-regularity The *neighbors* of a vertex v are all $u \in V$ such that $uv \in E$. The *degree* of a vertex v is the number of neighbors of v, denoted by $deg(v)$. A graph is said to be $k - regular$ for some $k \in \mathbb{N}$ if every $v \in V$ has degree k.

The graph G is said to be *finite* if both V are finite sets. A simple graph is a graph that does not have more than one edge between any two vertices and no edge starts and ends at the same vertex.

The following theorem relates vertex degree to the number of edges.

Theorem 19. *Let G = (V,E) be a finite graph. Then*

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

Corollary 5 (Handshaking Lemma)**.** *In every finite simple graph, the number of vertices having odd degree is even.*

From Theorem above, we can derive a corollary that counts the number of edges in k-regular graphs.

Corollary 6. *Let G = (V,E) be k-regular. Then*

$$|E| = \frac{k\,|V|}{2}.$$

Walks, paths, and cycles. A *walk* in a graph $G = (V, E)$ is a sequence of vertices $\sigma = (v_0, v_1, \ldots, v_k)$ such that $v_i v_{i+1} \in E$ for all i with $0 \leq i < k$. The *endpoints* of the walk are v_0 and v_k, and the *lenght* of the walk σ is $|\sigma|$ The walk is said to be *closed* if $v_0 = v_k$ and *open* otherwise.

A walk is a *path* if no vertices are repeated.

Theorem 20. *Let G = (V,E) be a graph. If u and v are vertices such that there exists a walk from u to v, then there exists a path from u to v.*

There we proform a minimality argument. Note that this theorem is about existence of path, a walk is not necessary a path.

Proof. Let $\sigma = (v_0, v_1, \ldots, v_n)$ be a walk from u to v of shortest length. We claim that σ is a path. Indeed, suppose for a contradiction that it is not a path; then there is some repeated vertex, so there exist $i, j \in 0, 1, \ldots, n$ such that $i < j$ and $v_i = v_j$. Hence there is no need to visit any of the vertices between v_i and v_j in the walk, since $v_i = v_j$ is connected to v_{j+1}. Concretely, consider the walk

$$\sigma' = (v_0, v_1, \ldots, v_i, v_{j+1}, \ldots, v_n).$$

Note that $|\sigma'| = |\sigma| - (j - i)$, and $j - i > 0$, so σ' is a shorter walk from u to v. But this contradicts our choice of σ as a walk of shortest length. We conclude that σ is a path. □

A *cycle* is a walk of length at least 3 and no vertices repeated except for $v_0 = v_k$.

Proposition 21. *Let G = (V,E). If G contains a closed walk of odd length, then G contains a cycle of odd length.*

The idea is similar to the proof of the previous theorem.

Proof. Let $\sigma = (v_0, v_1, \ldots, v_n)$ be an odd-length closed walk in G, and choose this walk to have minimal odd length (i.e, any shorter closed walk has even length). We shall prove that σ is a cycle.

For a contradiction, suppose that σ is not a cycle, so that there exist indices i, j with $0 \leq i < j \leq n$ such that $v_i = v_j$. Consider the two closed walks

$$\sigma_1 = (v_0, v_1, \ldots, v_i, v_{j+1}, \ldots, v_n)$$

and

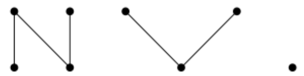
$$\sigma_2 = (v_i, v_{j+1}, \ldots, v_j).$$

Both are shorter than σ, so by the minimality of σ, they must have even length. But this implies that $|\sigma| = |\sigma_1| + |\sigma_2|$ is even. This leads to a contradiction.

We conclude that σ is a cycle. □

Connectedness. We say a graph $G = (V, E)$ is *connected* if for all $u, v \in V$, there exists a walk from u to v.

This is a disconnected graph:



Proposition 22. *For all $n \geq 1$, the graph K_n and Q_n are connected.*

Proof. Any two vertices u and v in K_n are connected by an edge, so we have a path (u, v) of length 1 between u and v. This shows that K_n is connected.

Now let u and v be any two vertices in Q_n. Suppose there are m bits that differ between u and v, Then we can flip them one by one to change u to v. This gives us a walk of length m between u and v, since there is an edge in Q_n between any two strings in Q_n that differ at exactly one bit. □

Here is an example:

Example 2 (*Disconnected Graph and Modular Arithmetic*)

Let $G = (V, E)$, where for $i, j \in \mathbb{Z}$ with $i < j$, we have $ij \in E$ if and only if $j - i \in 9, 15$. Then G is disconnected.

Proof. Starting at $n \in \mathbb{Z}$, we can reach any vertex m that is of the form

$$m = n + 15s + 9t$$

for some $s, t \in \mathbb{Z}$. By proposition in division section, the integers representable as $n + 15s + 9t$ are exactly the multiples of $\gcd(15, 9) = 3$. So in fact, from n one can reach any integer m of the form $n + 3k$ for some $k \in \mathbb{Z}$. That is, one can reach any integer m with $m \equiv n \mod 3$. Hence the three connected components of G are $[0]_3$, $[1]_3$, $[2]_3$, the equivalence classes of integers modulo 3. □

3.2 Triangles and bipartite graphs

A *subgraph* of a graph $G = (V, E)$ is a graph $G' = (V', E')$ such that $V' \subseteq V$ and $E' \subseteq E$ where for all $e = uv \in E'$, we have $u, v \in V'$.

The extremal question An extremal question asks for the extermal (maximum or minimum) number of objects we can have, subject to some restrictions.

We now want to derive the maximum number of edges in a triangle-free graph on n vertices.

Theorem 21 (Cauchy-Schwarz Inequality).
For all $u_1, \ldots, u_n, v_1, \ldots, v_n \in \mathbb{R}$, we have

$$\left(\sum_{i=1}^{n} u_i v_i\right)^2 \leq \left(\sum_{i=1}^{n} u_i^2\right)\left(\sum_{i=1}^{n} v_i^2\right).$$

Theorem 22 (Mantel's theorem). *Let $G = (V, E)$ be a graph not containing a triangle as a subgraph. Then*

$$|E| \leq \lfloor \frac{|V|^2}{4} \rfloor.$$

Proof. Consider the sum

$$\sum_{uv\in E} (\deg(u) + \deg(v)) .$$

The term $\deg(u)$ appears in the sum exactly once for every edge incident on u; that is, it appears $\deg(u)$ times. This is true for all $u \in V$, so we conclude that

$$\sum_{uv\in E} (\deg(u) + \deg(v)) = \sum_{u\in V} \deg(u)^2.$$

On the other hand, since G contains no triangle, for every pair of vertices u and v, the set of neighbors of u is disjoint from the set of neighbors of v. So $\deg(u) + \deg(v) \leq |V|$, and we have

$$\sum_{u\in V} \deg(u)^2 = \sum_{uv\in E} (\deg(u) + \deg(v)) \leq |E| \, |V| .$$

By the Cauchy-Schwarz inequality, we have

$$(2\,|E|)^2 \leq |V| \left(\sum_{u\in V} \deg(u)^2 \right)$$

Derivation using Cauchy-Schwarz inequality:

$$\begin{aligned} (2\,|E|)^2 &= \left(\sum_{u\in V} (\deg(u)) \right)^2 && (1)\\ &= \left(\sum_{u\in V} \deg(u) \cdot 1 \right)^2 && (2)\\ &\leq \left(\sum_{u\in V} \deg(u)^2 \right) \left(\sum_{u\in V} 1^2 \right) && (3)\\ &= |V| \left(\sum_{u\in V} \deg(u)^2 \right) . && (4) \end{aligned}$$

Hence

$$4\,|E|^2 \leq |V| \left(\sum_{u\in V} \deg(u)^2 \right) \leq |V|^2 \, |E| .$$

This implies that $|E| \leq \frac{|V|^2}{4}$. We can take the floor function on the R.H.S, since $|E|$ must be an integer.□

So if a graph has $|E| > \lfloor |V|^2 /4 \rfloor$, there must be a triangle subgraph in G.

How about the case a graph has exactly $|V|^2 /4$ edges? The Mantel's theorem does not assert that it must have triangles. This brings us to the definition of a bipartite graph.

This is the converse of the proposition we just proved. Pay attention here, because that means the theorem does not assert any graph has edges less or equal to $\lfloor |V|^2 /4 \rfloor$ is triangle-free.

Bipartite graphs A graph $G = (V, E)$ is *bipartite* if there exists a partition of $V = A \cup B$ of the vertex set ($A \cap B = \emptyset$) called the *bipartition* such that each edge has one endpoint in A and the other in B. For example, hypercubes Q_n are bipartite.

Proposition 23. *For all $n \geq 1$, the graph Q_n is bipartite.*

Proof. Let $Q_n = (V, E)$. Every elements $s \in V$ corresponds to a binary string of length n, $S = (s_1, s_2, \ldots, s_n)$ where each s_i is either 0 or 1. Define

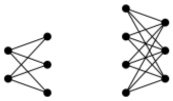
$$V_0 = \{ s \in V : s_1 + \cdots + s_n \equiv 0 \pmod 2 \}$$

and

$$V_1 = \{ s \in V : s_1 + \cdots + s_n \equiv 1 \pmod 2 \} .$$

It's clear that $V_0 \cup V_1 = V$ and $V_0 \cap V_1 = \emptyset$, so this is a bipartition of the vertex set. For every $e = s_1 s_2 \in E$, the strings s_1 and s_2 differ in exactly on bit, so if $s_1 \in V_0$, then $s_2 \in V_1$ and vice versa. Hence Q_n is bipartite. □

The *complete bipartite graph* $K_{m,n}$ is a bipartite graph with bipartition $V = V_m \cup V_n$, where $|V_m| = m$ and $|V_n| = n$, and E is the set $uv : u \in V_m, v \in V_n$ of all possible edges between the two sets.



The complete bipartite graphs $K_{2,3}$ and $K_{4,3}$.

When n is even, the graph $K_{n/2,n/2}$ has n vertices and exactly $n^2/4$ edges. When n is odd, the graph $K_{(n-1)/2,(n+1)/2}$ has n vertices and

$$\frac{n+1}{2} \cdot \frac{n-1}{2} = \frac{n^2-1}{4} = \frac{n^2}{4} - \frac{1}{4} = \lfloor \frac{n^2}{4} \rfloor$$

edges. (The last equality here follows from the fact that any odd n is congruent to either 1 or 3 modulo 4, which means its square is congruent to 1 modulo 4.)

$K_{n/2,n/2}$ has the largest possible number of edge shuch that the Mantel's theorem does not apply. But adding a single edge to $K_{n/2,n/2}$ results in a graph that must contain a triangle by Mantel's theorem.

Lemma 3. *A graph $G = (V, E)$ is bipartite if and only if its connected components are bipartite.*

Proof. Let $G_1 = (V_1, E_1), \ldots, G_k = (V_k, E_k)$ be the connected components of G.

If each connected component is bipartite, then we can bipartition each V_i into $A_i \cup B_i$. Then $A = \bigcup_{i=1}^n A_i$ and $B = \bigcup_{i=1}^n B_i$ is a bipartition of V. And every edge in G has one endpoint in A and the other in B, so G is bipartite.

Now suppose that there is some connected component G_i that is not bipartite. Now let $V = A \cup B$ be any partition of V into disjoint nonempty sets. Then $A_i = A \cap V$ and $B_i = B \cap V$ form a partition of V_i into disjoint nonempty sets. But since G_i is not bipartite, there must be an edge $uv \in E_i$ with both endpoints in A_i or both in B_i. This means that $A \cup B$ is not a bipartition of V, and since A and B were arbitrary, we conclude that G is not bipartite either. □

As a matter of fact, we shall prove something much stronger than just the fact that $K_{n/2,n/2}$ does not contain any triangles.

Theorem 23. *A graph is bipartite if and only if it does not contain any cycles of odd length.*

Proof. First, assume that G = (V,E) is bipartite with bipartition $V = A \cup B$. Let σ be a cycle in G. Each edge changes sides between A and B, so in order for starting and ending vertices of this cycle to be the same, $|\sigma|$ must be even.

Proof of this theorem requires the metric of distance imposed on a graph G. For all $u, v \in V$, the distance $dist(u, v)$ is the length of the shortest path from u to v. If no such path exists, then $dist(u, v) = \infty$.

Now suppose that G does not contain any cycles of odd length. To show that G is bipartite, it suffices to show that each of its connected components is bipartite by the lemma above. So without loss of generality we may assume that G is connected. That means $dist(u, v) < \infty$ for all $u, v \in V$.

Select on vertex $h \in V$ and set

$$V_0 = \{v \in V : dist(h,v) \equiv 0 \pmod 2\}$$

and

$$V_1 = \{v \in V : dist(h,v) \equiv 1 \pmod 2\}.$$

Let $e = uv \in E$. Consider a closed walk σ that follows a shortest path from u to h then a shortest path from h to v, and finally the edge uv. The length of σ is

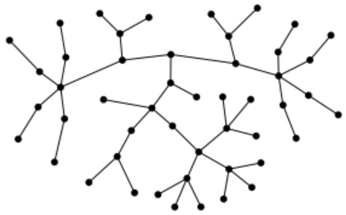
$$|\sigma| = dist(u,h) + dist(h,v) + 1$$

But since G contains no cycles of odd length, $|\sigma|$ must be even, by (the Contrapositive of) Proposition 21, any closed walk in G must have even length. This means that $dist(u,h), dist(h,v)$ are not the same modulo 2. That is, either $u \in V_0$ and $v \in V_1$ or vice versa. □

3.3 Trees

A graph G is called a *forest* if it contains no cycles. A *tree* is a connected forest. A vertex of degree 1 in a forest is called a *leaf*.

Proposition 24. *Let F be a forest with at least one edge (and hence at least two vertices). There are at least two leaves in F.*



A tree with 25 leaves.

Proof. Let $\sigma = (v_0, v_1, \ldots, v_k)$ be a path of maximum length in F. SInce there is at least one edge in F, we have $|\sigma| \geq 1$. The claim is that $\deg(v_0) = \deg(v_k) = 1$. Suppose for a contradiction that $\deg(v_0) \geq 2$. So there exists $u \in V$ such that uv_0 is an edge. If $u = v_i$ for some i, then there is a cycle $v_i, v_0, \ldots, v_i$ contradicting the fact that F is a forest. Hence for all $1 \leq i \leq k$, we have $u \neq v_i$. But this a path of length $k+1$ exists. This contradicts the maximality of σ. □

The intuition is that, a tree should have at least two ends, if not, either there is a cycle or there is an extending part (a new end) that can be added to the tree.

Proposition 25 (A characterization of trees). *A graph G = (V,E) is a tree if and only if for all $u, v \in V$, there exist a unique path from u to v.*

Proof. First, suppose that G is a tree. Let $u, v \in V$, since G is a tree, it is connected, so there must exist some path σ from u to v. Now we prove that this path is unique. For a contradiction, let σ' be another path from u to v. Both σ and σ' starts at u, but eventually they must diverge. Let v_1 be the last vertex where σ and σ' have in common (this should be u). Eventually they must meet again, since they both end at v. Let v_2 be the first vertex where they meet again. Now let σ'' be the path obtained by following σ from v_1 to v_2, then following σ' from v_2 to v. This is a cycle, by choice of v_1 and v_2. But this give us a contradiction, since G is a tree and contains no cycles.

Now suppose that for all $u, v \in V$, there exists a unique path from u to v. The fact that there exists a path at all between every two vertices immediately implies that G is connected. It remains to prove that there is no cycle in G. But this is clear, since if σ is a cycle in G, then picking any two distinct vertices u, v on σ, we have two distinct paths from u to v (clockwise and counterclockwise along σ). □

We can prove another characterization of trees, but we need a lemma first.

Lemma 4 (Detour Lemma)**.** *Let G = (V,E) be a connected graph and let σ be a cycle in G, If G' is the graph obtained by deleting on edge from σ, then G' is still connected.*

Lemma 5. *A graph G = (V,E) is a tree if and only if it is connected and* $|E| = |V| - 1$.

Proof. We prove the forward implication by induction on $n = |V|$. Concretely, the statement we shall prove is, for all $n \geq 1$, for all $G = (V, E)$ with $|V| = n$, if G is a tree, then $|E| = n - 1$ and G is connected.

The base case in $n = 1$, where the only possible graph G is simply a single vertex with no edges. This is a tree, since it is connected and concains no cycles. We have $|E| = 0 = 1 - 1$.

For the inductive step, let $n \geq 1$ and suppose that the statement holds n. Let $G = (V, E)$ be a tree with $|V| = n + 1$. Since $n \geq 1$, $|V| \geq 2$, there are at least two vertices, and the connectedness of G means that there is at least one edge in G. Hence G is a firest with at least one edge, and by Proposition 24, there are at least two leaves in G. Pick one of these leaves and call it u. Form a new graph $G' = (V', E')$ by removing the leaf u and the edge incident on it. We have $|V'| = n$ and $|E'| = |E| - 1$. The graph G' is a tree, because for each pair of vertices $v, w \in V'$, there is a unique path from v to w in G, and this path does not pass through u (since u is a leaf). Hence

$$|E| = \left|E'\right| + 1 = \left|V'\right| - 1 + 1 = |V| - 1,$$

where in the second equality we have to used the inductive hypothesis.

Now assume that G is connected and $|E| = |V| - 1$. We want to show that G has no cycles. For a contradiction, suppose that σ is a cycle in G. Remove an edge from σ, by detour lemma, the resulting graph is still connected. If this graph still has a cycle, remove an edge from it again, and repeat this process until now cycles remain. This graph G' still has $|V|$ vertices, but it has a new edge set E' with $|E'| < |E|$. But G' is now a tree, since it contains no cycles but is still connected. So by the previous paragraph, we have $|E'| = |V| - 1 = |E|$. The contradiction completes the proof. □

Now we are ablt the answer the following extremal question: How many edges can a graph on n vertices have if it does not contain a cycle (i.e., is a forest)?

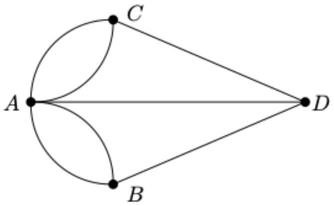
Corollary 7. *Let F = (V,E) be a forest, then* $|E| \leq |V| - 1$.

Proof. Let $k \geq 1$ denote the number of connected components of F, and number the connected components as $C_1, \ldots, C_k$. Join C_1 to C_2 by an edge, then join C_2 to C_3 by an edge, and so on. This creates a connected graph G = (V,E′) with $|E'| = |E| + k - 1$, and G must be a tree, since adding these edges does not introduce a cycle. By the previous theorem, $|E'| = |V| - 1$, so $|E| \leq |E'| = |V| - 1$. □

3.4 *Eulerian trails and circuits*

Recall that a wak that des not repeat any vertex is called a path. Now we introduce the concept of *trail*: this is a walk that does not repeat any *edge*. A walk is called an *Eulerian trail* if it uses every edge of G exactly once, and an *Eulerian circuit* if it is an Eulerian trail and it is closed.

Theorem 24. *Let G = (V,E) be a connected graph with a finite number of vertices and edges, where multiple edges between the same pair of vertices are allowed. Then G has an Eulerian circuit if and only if every vertex in G has even degree.*



The graph representing the seven bridges of Konigsberg

Proof. Suppose G has an Eulerian circuit $\sigma = (v_0, v_1, \ldots, v_k)$ where $v_0 = v_n$ For each $v_i \neq v_0$, each visit uses two edges, one edge to enter v_i and one edge to leave v_i. So $\deg(v_i)$ is even for all $0 \leq i < n$. For $v_0 = n$, there is one edge that leaves it at the beginning, two edges for each visit during the circuit, and then one returning to it at the end. So $\deg(v_0)$ is even as well.

We prove the reverse implication by induction on $m = |E|$. That is, we prove that for all integers $m \geq 0$, if G = (V,E) with $|E| = m$ and all vertices in G have even degree, then G has an Eulerian circuit.

For the base case $m = 0$, the graph G is a single vertex with no edges, and it has an Eulerian circuit.

Now let $m \geq 1$ and suppose the statement holds for all integers less than or equal to m. Let G = (V,E) with $|E| = m + 1$. Since there is at least one edge in G, G cannot be a tree. This is because every tree with at least one edge contains at least two leaves, and leaves have (odd) degree 1. Let σ be a cycle in G and form a new graph G′ = (V,E′) by deleting every edge in σ. This graph may no longer be connected, but since we remove a cycle, every vertex in G′ either has the same degree as in G, or degree two less than in G. So all vertices in G′ have even degree as well. Let k be the number of connected components of G′; number them as $H_1, \ldots, H_k$. Each has at most m

Process of building an Eulerian circuit in G:

- Start at any vertex σ
- Follow σ
- Each time we reach a vertex of a connected component H_i of G′ for the first time, follow its Eulerian circuit. *(We know that each connected component has an Eulerian circuit by the inductive hypothesis.)*

edges, so by the inductive hypothesis, each has an Eulerian circuit.

We build an Eulerian circuit in G using the process to the right. Eventually, this gives us an Eulerian circuit in G. □

Theorem 25. *Let G = (V,E) be a connected graph (with multiple edges allowed). Then G contains an Eulerian trail that isn't an Eulerian circuit if and only if exactly two vertices of G have odd degree.*

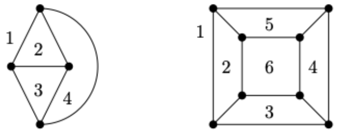
Proof. Suppose that G contains an Eulerian trail $\sigma = (v_0, v_1, \dots, v_k)$ in which $v_0 \neq v_k$. By a similar reasoning to the first paragraph of the previous proof, every vertex in the walk has even degree, except both v_0 and v_k must have odd degree, since starting at v_0 we must leave for the first time via an edge, and at the end we enter v_k one last time without exiting.

Now assume that G contains exactly two odd-degree vertices. By the previous theorem, there does not exist an Eulerian circuit in G. Let u and v be the two odd-degree vertices. Add a new edge uv, even if there was already an edge between u and v before. We get G′ in which every vertex has even degree. So G′ must have an Eulerian circuit σ'. It contains the edge uv, but we can remove it to get an Eulerian trail in G. □

3.5 Planar graphs

A graph G is called *planar* if it can be drawn in the plane without any edges crossing. If we can, such a drawing is called a *planar embedding* of G.

Planarity is a property that passes down to subgraphs.



Planar embeddings of K_4 and Q_3

Proposition 26. *If G is planar then any subgraph H of G is also planar.*

Proof. If G is drawn in the plane without edges crossing, and we erase some edges and vertices to create a drawing of H, we cannot introduce any edge crossings along the way. □

A planar embedding of a graph determins regions, or *faces* bounded by edges, including on region outside the graph. We define the *Euler characteristic* of a planar embedding of G to be the quantity

$$\chi = |V| - |E| + f,$$

where f is the number of faces in the embedding.

From the figure above, we notice that both K_4 and Q_3 have $\chi = 2$, this is not just a coincidence.

Theorem 26 (Euler's formula)**.** *Let G = (V,E) be a connected planar graph, let f be the number of faces in a planar embedding of G. Then* $|V| - |E| + f = 2$.

To prove this, we need the following theorem.

Theorem 27 (Jordan curve theorem). *Every closed curve in the plane $\mathbb{R}^2$ that does not intersect itself divides the plane into two regions.*

Proof of Theorem 26. We proceed by induction on $m = |E|$. In the case that $m = 0$, G can only be a single vertex, in which we have $\chi = 1 - 0 + 1 = 2$.

Now let $m \geq 0$ and assume that the theorem holds for all natural numbers at most m. Let G = (V,E) be a connected planar graph with $|E| = m + 1$. Draw G in the plane and let f be the number of faces in the embedding. There are two cases, according to whether G is a tree or not. If G is a tree, then $f = 1$ since G has no cycles. But we also have $|E| = |V| - 1$, so

$$\chi = |V| - |E| + f = |V| - |V| + 1 + 1 = 2.$$

If G is not a tree, then G has a cycle, call it σ. Form G′ by removing exactly one edge e from σ, so

$$\left|E'\right| = |E| - 1 = m + 1 - 1 = m.$$

By the detour lemma, G′ is still connected, and it is planar, since removing an edge doesn't introduce any crossings. By the inductive hypothesis, the Euler characteristic χ' of G′ satisfies

$$\chi' = |V| - \left|E'\right| + f' = 2,$$

where f' is the number of faces in a planar embedding of G′. Returning to the drawing of G, the cycle σ is closed curve without intersections in the plane, so it divides the plane into two regions (each of which may have multiple faces within them). In particular, the removed edge e forms a border between two distinct daces, and in G′, these two faces combine into one face. Hence $f = f' + 1$. Putting everything together, we calculate the Euler characteristic of G satisfies

$$\chi = |V| - |E| + f = |V| - (\left|E'\right| + 1) + (f' + 1) + 1 = 2. \quad \square$$

Euler's formula tells us that the number of faces in any drawing of G depends entirely on the number of vertices and edges in G. We can use this information to give conditions for a graph to be non-planar.

Theorem 28. *Let G = (V,E) be a connected planar graph with $|V| \geq 5$. The*

$$|E| \leq 3\,|V| - 6.$$

Under the further assumption that G contains no triangles, we have a better bound

$$|E| \leq 2\,|V| - 4.$$

Proof. The proof is by double counting. Let R be the set of all regions into which G divides the plane, so that $|R| = f$, the number of faces. Consider the set

$$S = \{(e, r) \subset E \times R : \text{the edge e touches the region r}\}.$$

We will count $|S|$ in two ways. First off, each edge $e \in E$ touches at most two regions, so $|S| \leq 2|E|$. On the other hand, every region is bounded by a cycle, and a cycle has at least three edges, so each region $r \in R$ touches at least 3 edges. In other words, $|S| \geq 3f$. Chaining these into two inequalities, we have

$$3f \leq |S| \leq 2|E|.$$

But by Euler's formula, we have $|V| - |E| + f = 2$, so $f = 2 + |E| - |V|$, and substituting this above, we get

$$3(2 + |E| - |V|) \leq 2|E|.$$

Distributing and rearranging terms yields the desired inequality $|E| \leq 3|V| - 6$.

If we have the further assumption that there are no triangles in G, then every region must touch at least 4 edges, allowing us to conclude the stronger inequality $4f \leq 2|E|$. Then we proceed as above to get

$$4(2 - |V| + |E|) \leq 2|E|,$$

which can be manipulated to get $|E| \leq 2|V| - 4$. □

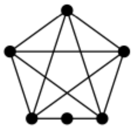
The condition that $|V| \geq 5$ is a complete non-issue, since a graph with fewer than 5 vertices is a subgraph of K_4, and we can apply Proposition 26, since we already know that K_4 is planar. As a corollary of the (contrapositive of the) previous theorem, we can give two important examples of nonplanar graphs.

Corollary 8. *The complete graph K_5 and the complete bipartite graph $K_{3,3}$ are nonplanar.*

Proof. If $(V, E) = K_5$ we have $|V| = 5$ so $3|V| - 6 = 9$, but $|E| = 10$. So by the contrapositive of the previous theorem, K_5 is nonplanar.

The graph $K_{3,3}$ does not contain any triangles, so if it were planar, then the stronger bound $|E| \leq 2|V| - 4$ must hold. But in $(V, E) = K_{3,3}$, we have $|V| = 6$ and $2|V| - 4 = 8$, but $|E| = 9$. So $K_{3,3}$ is nonplanar. □

This corollary, combined with the contrapositive of Proposition 26, shows that K_n is nonplanar for all $n \geq 5$, and that $K_{m,n}$ is nonplanar for all $m, n \geq 3$. More generally, any graph that contains K_5 or $K_{3,3}$ as a subgraph is nonplanar.



The graph K_5 with an extra vertex subdividing an edge.

However, subgraphs are not exactly the right notion to be considering when talking about planarity. To see why, consider the graph to the right.

It is nonplanar, since if it had a planar embedding, then by contracting on of the edges incident on the vertex, and bending other edges accordingly, we would obtain a planar embedding of K_5. On the other hand, this graph does not contain K_5 as a subgraph. We thus introduce the following notion.

We say that a graph H is a *graph minor* of a graph G if H can be obtained from G repeatedly by either one of the following operations to the right.

- deleting an edge;
- deleting a vertex; or
- contracting an edge uv by removing it and merging u and v into a single vertex (and also combining any resulting multiple edges into a single edge).

Any number of the first two operations simply creates a subgraph. It is the contraction operation that produces interesting examples of graph minors. It is easy to see pictorially that contracting edges preserves planarity, so any graph minor of a planar graph is planar. We can use this to show, for instance, that the Petersen graph, depicted to the right, is nonplanar. Since by contracting each of the five edges connecting the inner star pentagram to the other pentagon, we obtain a K_5. (Note that Theorem 28 does not apply to the Petersen graph.)

Not only all graphs that have K_5 and $K_{3,3}$ as minors are nonplanar, it turns out that there are the only nonplanar graphs.

Theorem 29 (Wagner's theorem). *A graph G is nonplanar if and only if either K_5 or $K_{3,3}$ is a minor of G.*

It deals with the notion of a *subdivision* of a graph G, subdividing an edge into two edges each time. (We have shown an example of a subdivision of K_5)

Theorem 30 (Kuratowski's theorem). *A graph G is nonplanar if and only if it contains a subdivision of either K_5 or $K_{3,3}$ as a subgraph.*

It is easy to see that if H is a subdivision of G, then G is a minor of H, since we can reobtain G by contracting all the edges created by the subdivision operations. So the two theorems above are certainly very closely related, though it is not immediately obvious if one should imply the other. The truth is that they are equivalent, because it can be shown that any graph with either K_5 or $K_{3,3}$ as a minor also has a subgraph that is a subdivision of one of them.

4 Combinatorics

4.1 Counting

Theorem 31 (Principle of inclusion and exclusion)**.** *Let $n \geq 2$ be an integer and let $A_1, \ldots, A_n$ be finite sets. Then*

$$\left|\bigcup_{i}^{n} A_i\right| = \sum_{k=1}^{n} (-1)^{k+1} \left(\sum_{1 \leq i_1 < \ldots < i_k \leq n} \left| \bigcap_{j=1}^{k} A_{i_j} \right| \right).$$

Theorem 32 (Binomial theorem)**.** *For all $x, y \in \mathbb{R}$ and positive integers n,*

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

Proof. Expanding the left-hand side, we get

$$(x+y)^n = \underbrace{(x+y)(x+y)\cdots(x+y)}_{n \text{ times}}.$$

We see that there will be 2^n different terms after repeatedly using the distributive property. Each term will be a product of x to some power and y to some power, where the powers add up to n. In other words, each term will be of the form $x^{n-k}y^k$ for some k between 0 and n. The number of times that the term $x^{n-k}y^k$ appears is the number of ways to choose k of the n factors to be y, which is $\binom{n}{k}$. This gives exactly the right-hand side of the equation. □

Proof of Theorem 31. Let $n \geq 2$ and let $A_1, A_2, \ldots, A_n$ be finite sets. Recall that the identity we want to prove is

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{k=1}^{n} (-1)^{k+1} \left(\sum_{1 \leq i_1 < \ldots < i_k \leq n} \left| \bigcap_{j=1}^{k} A_{i_j} \right| \right).$$

Let x be an element that belongs to the union $\bigcup_{i=1}^{n} A_i$. Clearly, this contributes 1 to the left-hand side, so we must show that it contributes exactly 1 to the right-hand side as well. We know that x is a member of at least one of A_i; let $s \geq 1$ be the number of the sets A_i that contain x.

Each time x is a member of some A_i, it contributes +1 to the right-hand side. This happens $\binom{s}{1}$ times. Each time it is a member of A_i and A_j for some $i < j$, x contributes -1 to the right-hand side; this happens $\binom{s}{2}$ times. generally, if x is a member of $A_1 \cap A_2 \cap \ldots \cap A_r$, it contributes +1 if r is odd and -1 if r is even, and this term is repeated $\binom{s}{r}$. So the total contribution of x to the right-hand side is

$$\binom{s}{1} - \binom{s}{2} + \binom{s}{3} - \ldots + (-1)^{s+1}\binom{s}{s}$$

By the binomial theorem,

$$0 = (1-1)^s = \binom{s}{0} - \binom{s}{1} + \binom{s}{2} - \ldots + (-1)^s \binom{s}{s}.$$

We know that $\binom{s}{0} = 1$, so by rearranging the terms, we get

$$\binom{s}{1} - \binom{s}{2} + \binom{s}{3} - \ldots + (-1)^{s+1} \binom{s}{s} = \binom{s}{0} = 1.$$

This means that the contribution of x to the right-hand side is 1, as desired. The proof is finished since x is arbitrary.□

Example 3

Let $X = \{1, 2, \ldots, 100\}$ *and we want to count the number of* $n \in \mathbb{X}$ *with* $\gcd(n, 30) = 1$.

To accomplish this, it turns out to be easier to count the number of $n \in X$ with $\gcd(n, 30) \geq 2$ (and then we must subtract this number from $|X| = 100$). For any positive integer r let

$$A_r = \{n \in X : r|n\}.$$

Since $30 = 2 \cdot 3 \cdot 5$, an integer n has $\gcd(n, 30) \geq 2$ if and only if $n \in A_2, n \in A_3$, or $n \in A_5$. By the principle of inclusion and exclusion,

$$|A_2 \cup A_3 \cup A_5| = |A_2| + |A_3| + |A_5| - |A_2 \cap A_3| - |A_2 \cap A_5| - |A_3 \cap A_5| + |A_2 \cap A_3 \cap A_5|.$$

The number of even integers in X is $|A_2| = 50$, similarly the number of multiples of 3 is $|A_3| = \lfloor 100/3 \rfloor = 33$. In general, $|A_r| = \lfloor 100/4 \rfloor$. Furthermore, since $\gcd(2, 3) = 1$, the intersection $A_2 \cap A_3$ is simply A_6, and by analogous reasoning, we see that

$$\begin{aligned} |A_2 \cup A_3 \cup A_5| &= |A_2| + |A_3| + |A_5| - |A_6| - |A_{10}| - |A_{15}| + |A_{30}| \\ &= \lfloor \frac{100}{2} \rfloor + \lfloor \frac{100}{3} \rfloor + \lfloor \frac{100}{5} \rfloor - \lfloor \frac{100}{6} \rfloor - \lfloor \frac{100}{10} \rfloor - \lfloor \frac{100}{15} \rfloor + \lfloor \frac{100}{30} \rfloor \\ &= 50 + 33 + 20 - 16 - 10 - 6 + 3 \\ &= 74. \end{aligned}$$

We conclude that the number of $n \in X$ with $\gcd(n, 30) = 1$ is $100 - 74 = 26$.

4.2 Permutations and combinations

Proposition 27. *Let* $n \geq 0$ *be an integer. Then*

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

Proof. Let X be set of size 2n. The right-hand side counts the number of subsets of X of size n. We count this in a different way. Let A and B be such the $|A| = |B| = n$, and $A \cup B = X$ (so we must have $A \cap B = \emptyset$). Then choosing a subset of X is the same as choosing k elements of A, where $0 \leq k \leq n$, and then choosing the remaining $n - k$ elements from B. In other words,

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

□

Now let's explore more identities involving binomial coefficients.

$$\begin{array}{ccccccccccccccccc}
& & & & & & & & 1 & & & & & & & & \\
& & & & & & & 1 & & 1 & & & & & & & \\
& & & & & & 1 & & 2 & & 1 & & & & & & \\
& & & & & 1 & & 3 & & 3 & & 1 & & & & & \\
& & & & 1 & & 4 & & 6 & & 4 & & 1 & & & & \\
& & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & & \\
& & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & & \\
& 1 & & 7 & & 21 & & 35 & & 35 & & 21 & & 7 & & 1 & \\
1 & & 8 & & 28 & & 56 & & 70 & & 56 & & 28 & & 8 & & 1 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots
\end{array}$$

Pascal's triangle

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

Laying out the binomial coefficients in the triangle.

This triangle can easily be drawn by hand by writing 1s along its boundary and filling in the interior by adding the two numbers above each cell. Now we shall prove this observation.

Proposition 28 (Pascal's identity). *Let $n \geq 0$ and $k \geq 1$ be integers. Then*

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

The only number that appears infinitely many times is 1, since any $n \in Z$ can only appear in the first n + 1 rows of the triangle.
A fun conjecture to think about is the Singmaster's conjecture, but whether the statement is true is still a open problem.

Proof. Set $X = \{1, 2, \ldots, n\}$. We count the number of sets $Y \subseteq X$ there are of size k. Either $1 \in Y$ or $1 \notin Y$. If $1 \notin Y$, then there are $\binom{n-1}{k}$ possibilities for what Y can be. If $1 \in Y$, then there are still $k - 1$ elements to choose from $n - 1$ elements, so there are $\binom{n-1}{k-1}$ possibilities for what Y can be. Hence the right-hand side counts the number of sets $Y \subseteq X$ of size k. But this is exactly what the left-hand side counts as well. □

Another proof of Proposition 8 We re-establish the notation of the proposition statement. Let X be a finite nonempty set, let E be the set of subsets of X with even cardinality, and let O bet the set of subsets with odd cardinality. By the binomial theorem with $x = -1, y = 1$, we have

$$\begin{aligned} 0 = (-1+1)^n &= \sum_{k=0}^{n} \binom{n}{k} (-1)^k 1^{n-k} \\ &= \binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \ldots + (-1)^n \binom{n}{n} \end{aligned}$$

Hence

$$\binom{n}{0} + \binom{n}{2} + \ldots + \binom{n}{n} = \binom{n}{1} + \binom{n}{3} + \ldots + \binom{n}{n-1}.$$

but the left-hand side is just $|E|$ and the right-hand side is $|O|$. □

Proposition 29. *For all $n \in \mathbb{N}$,*

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

Proof. By the binomial theorem with $x = y = 1$, we have

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

The proposition can also be proven by noting that the left-hand side counts the number of subsets of $\{1,2,\ldots,n\}$ of size 0, plus the subsets of size 1, and so on up to all subsets of size n. But adding up all these numbers gives the number of all subsets of $\{1,2,\ldots,n\}$, which is 2^n. □

Theorem 33 (Freshman's dream)**.** *Let p be a prime number and let $x, y \in \mathbb{Z}$. Then*

$$(x+y)^p = x^p + y^p \pmod{p}.$$

Proof. By the binomial theorem, we have

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

Taking this equation modulo p, we are done if we can show that for all $1 \leq k \leq p-1$,

$$\binom{p}{k} \equiv 0 \pmod{p}.$$

Let $1 \leq k \leq p-1$ and expand the binomial coefficient:

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

Multiplying both side by $k!(p-k)!$, we get

$$k!(p-k)! \binom{p}{k} = p! = p(p-1)!.$$

This is an integer and the right-hand side is a multiple of p, so the left-hand side is as well. Then since p devides the product $k!(p-k)!$, either p divides $k!(p-k)!$ or it divides $\binom{p}{k}$. But the former is not possible, since

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

and all of the factors on the right-hand side are positive integers less than p. Hence p can not divide any of them, and not divide $k!(p-k)!$, so it must divide $\binom{p}{k}$. □

4.3 *Recurrence*

Proposition 30. *Suppose that (a_n) is a sequence with $a_0 = c$ for some constant $c \in \mathbb{R}$ and, for all $n \geq 0$, $a_n = b \cdot a_{n-1}$ for some $b \in \mathbb{R}$. Then for all $n \geq 0$, we have*

$$a_n = c \cdot b^n.$$

In general, we will consider recurrences of the form

$$a_n = f_1(n)a_{n-1} + f_2(n)a_{n-2} + \ldots + f_k(n)a_{n-k} + g(n),$$

where $f_1, \ldots, f_k, g : \mathbb{N} \to \mathbb{R}$ are functions, and $k \geq 1$ is an integer.

These recurrences are said to be *linear of degree k*. If $g(n) = 0$, the recurrence is said to be *homogeneous*; otherwise, it is *non-homogeneous*.

For *non-homogeneous* recurrences, the same recurrence without the $g(n)$ term is called the *associated homogeneous recurrence*. If, for all $1 \leq i \leq k$, the function f_i is a constant function, then the recurrence is said to have *constant coefficients*.

A sequence (p_n) that satisfies the recurrence is called a *particular solution*. A *general solution* is a formula describing all possible solution using some parameters. If we specify values for the first few terms, these are call *initial conditions* for the recurrence.

Theorem 34. *Consider the non-homogeneous recurrence*

$$a_n = f(n)a_{n-1} + g(n),$$

where $f, g : \mathbb{N} \to \mathbb{R}$. If (p_n) is any particular solution to the recurrence and (h_n) is a general solution to the associated homogenous recurrence, i.e.,

$$h_n = f(n)h_{n-1},$$

for all $n \geq 1$, then the general solution for recurrence is

$$a_n = h_n + p_n.$$

Theorem 35. *Consider the recurrence*

$$a_n = f_1(n)a_{n-1} + f_2(n)a_{n-2} + \cdots + f_k(n)a_{n-k} + g(n),$$

for some $k \geq 1$ and $f_1, \ldots, f_k, g : \mathbb{N} \to \mathbb{R}$. If (p_n) is any particular solution to the recurrence and (h_n) is a general solution to the associated homogeneous recurrence, i.e.,

$$h_n = f_1(n)h_{n-1} + f_2(n)h_{n-2} + \cdots + f_k(n)h_{n-k},$$

for all $n \geq 1$, then the general solution for the recurrence is given by $a_n = h_n + p_n$.

Theorem 36. *Consider the recurrence given by*

$$a_n = c_1 a_{n-1}.$$

If the characteristic polynomial of this recurrence has two distinct roots r_1 and r_2, then the general solution of the recurrence is

$$a_n = \alpha_1 r_1^n + \alpha_2 r_2^n.$$

Supposing we know the initial conditions a_0 and a_1, we have the identities

$$\alpha_1 = \frac{a_1 - r_2 a_0}{r_1 - r_2} \quad \text{and} \quad \alpha_2 = \frac{r_1 a_0 - a_1}{r_1 - r_2}.$$

Theorem 37. *Let $k \geq 2$ be an integer and consider the recurrence given by*

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

If the characteristic polynomial of this recurrence has k pairwise distinct roots $r_1, \ldots, r_k$, then the general solution of the recurrence is

$$a_n = \alpha_1 r_1^n + \alpha_2 r_2^n + \cdots + \alpha_k r_k^n.$$

Theorem 38. *Consider the homogeneous recurrence*

$$a_n = c_1 a_{n-1} + c_2 a_{n-2}.$$

If the characteristic polynomial of this recurrence has a repeated root r, then the general solution of the recurrence is

$$a_n = \alpha_1 r^n + \alpha_2 n r^n.$$

Suppose we know the initial conditions a_0 and a_1, then we have the identities

$$\alpha_1 = a_0 \quad \text{and} \quad \alpha_2 = \frac{a_1 - a_0 r}{r}.$$

References