McGill University - Fall 2022

Math 240: Discrete Structures

Ritchie Yu
December 20, 2022

Condensed material given in lectures by Prof. Jérôme Fortier and Prof. Adrian Vetta.

As an introduction to discrete mathematics, this course covered set theory, logic, proofs, number theory, combinatorics, and graph theory.

1 Set Theory

Definition 1.1. A set is a collection of objects of the same nature.

Definition 1.2. If x is an element of a set A, then we say $x \in A$.

Definition 1.3. If all elements of set A are elements of set B, then A is a subset of B, and we say $A \subseteq B$.

Definition 1.4. If all elements of set A are elements of set B, then A is a subset of B, and we say $A \subseteq B$.

2 Logic

3 Proofs

4 Number Theory

4.1 Divisibility

Definition 4.1 (Divisibility)**.** If $\exists k \in \mathbb{Z}$ such that $b = ka$ for $a, b \in \mathbb{Z}$, then $a|b$. Otherwise, $a \nmid b$.

Definition 4.2 (Greatest Common Divisor)**.** $d = \gcd(a, b)$ for $a, b, d \in \mathbb{Z}$ if $d|a$ and $d|b$ and any other common divisor is less than or equal to d.

Remark. $\gcd(x, 0) = x$ for any $x \in \mathbb{Z}$ and $x \neq 0$.

Definition 4.3 (Coprime)**.** Two numbers are coprime if $\gcd(a, b) = 1$.

Theorem 4.1. A few results about divisibility:

(1) If $a|b$ then $a|bc \ \forall c \in \mathbb{Z}$

(2) If $a|b$ and $a|c$ then $a|b \pm c$

(3) If $a|b$ and $b|c$ then $a|c$

Proof. Apply definition of divisibility, where $a, b, c, k, l \in \mathbb{Z}$

(1)	(2)	(3)
$b = ak$	$b = ak$	$b = ak$
$bc = (ak)c$	$c = al$	$c = bl$
$bc = a(kc)$	$b \pm c = ak \pm al$	$= (ak)l$
$= a(k \pm l)$	$= a(k \pm l)$	$= a(kl)$

■

Lemma 4.1. If $a = qb + r$, then $\gcd(a, b) = \gcd(b, r)$.

Proof. Let $x = \gcd(a, b)$ and let $y = \gcd(b, r)$.

$$\begin{aligned}
x = \gcd(a,b) &\implies (x|a) \wedge (x|b) \\
x|b &\implies x|qb \\
x|a &\implies x|qb + r \\
(x|qb) \wedge (x|qb + r) &\implies x|r \\
(x|b) \wedge (x|r) &\implies x \leq y
\end{aligned}$$

Similarly,

$$\begin{aligned}
y = \gcd(b,r) &\implies (y|b) \wedge (y|r) \\
y|b &\implies y|qb \\
y|r &\implies y|a - qb \\
(y|qb) \wedge (y|a - qb) &\implies y|a \\
(y|a) \wedge (y|b) &\implies y \leq x
\end{aligned}$$

Since $x \leq y$ and $y \leq x$, then $x = y$, so $\gcd(a, b) = \gcd(b, r)$.

■

Theorem 4.2 (Division algorithm)**.** Given $a, b \in \mathbb{Z}, b \neq 0$, there exist $q, r \in \mathbb{Z}$ with $0 \leq r < b$ such that $a = qb + r$. Furthermore, q and r are *unique*.

Euclidean Algorithm
To find $\gcd(a, b)$, apply the division algorithm to write $a = qb + r$ and find $\gcd(b, r)$. Terminate at $r = 0$ (guaranteed to happen). This algorithm relies on Lemma 4.1.

Theorem 4.3 (Bézout's Theorem)**.** If $d = \gcd(a, b)$, then there exist integers $s, t \in \mathbb{Z}$ such that:

$$d = sa + tb$$

Proof. Let $r_0 = a$ and let $r_1 = b$. Assume that Euclid's algorithm gives

$$\begin{aligned} r_0 &= r_1 q_1 + r_2 \\ r_1 &= r_2 q_2 + r_3 \\ r_2 &= r_3 q_3 + r_4 \\ &\vdots \\ r_{n-1} &= q_n r_n + r_{n+1} \end{aligned}$$

We claim that for some $s_n, t_n \in \mathbb{Z}$ and $n \in \mathbb{N} \cup \{0\}$

$$r_n = s_n a + t_n b$$

We prove this by strong induction on $n \in \mathbb{N} \cup \{0\}$.

Base case: When $n = 0$, $r_0 = a = a + 0 \cdot b$, so $s_0 = 1, t_0 = 0$.

Inductive step: For $n \in \mathbb{N}$, assume that

$$\begin{aligned} r_{n-1} &= s_{n-1} a + t_{n-1} b \\ r_n &= s_n a + t_n b \end{aligned}$$

From Euclid's algorithm and by the I.H., we have that

$$\begin{aligned} r_{n+1} &= -q_n r_n + r_{n-1} \\ &= -q_n(s_n a + t_n b) + (s_{n-1} a + t_{n-1} b) \\ &= a\underbrace{(s_{n-1} - s_n q_n)}_{s_{n+1}} + b\underbrace{(t_{n-1} - t_n q_n)}_{t_{n+1}} \end{aligned}$$

Euclid's algorithm says there exists $r_n = d = \gcd(a, b)$.
We have shown $r_n = s_n a + t_n b$ for some $s_n, t_n \in \mathbb{Z}$.

■

Theorem 4.4. $1 = sa + tb \iff \gcd(a, b) = 1$ $(a, b, s, t \in \mathbb{Z})$.

Proof.

- (⇒) Let $d|a$ and $d|b$. Then $a = dx$ and $b = dy$ for some $x, y \in \mathbb{Z}$ and we can write $1 = sdx + sdy = d(sx + ty)$. This implies $d|1$ so $d = \pm 1$. Hence $\gcd(a, b) = 1$.
- (⇐) This direction follows by Bézout's Theorem.

■

Theorem 4.5. If $c|a$ and $c|b$, then $c| \gcd(a, b)$.

Proof. Let $d = \gcd(a, b)$.

Then $\exists s, t \in \mathbb{Z}$ such that $d = sa + tb$ (Bézout's Theorem).
Since $c|a$, then $a = cx$ for some $x \in \mathbb{Z}$.
Since $c|b$, then $b = cy$ for some $y \in \mathbb{Z}$.
We have $d = sa + tb = scx + tcy = c(sx + ty)$.

■

4.2 Prime Numbers

Definition 4.4 (Prime)**.** If $p \geq 2, p \in \mathbb{Z}$ and its only divisors are 1 and p, then p is prime.

Definition 4.5 (Composite)**.** If $n \in \mathbb{Z}$ is not prime, then it is composite. In other words, $\exists a, b \in \mathbb{Z}$ such that $n = ab$ where $2 \leq a, b < n$.

Remark. Goldbach's Conjecture (open since 1742) asks if every even number is the sum of 2 primes.

Theorem 4.6. Twin primes p, q are time primes if $q = p \pm 2$.

Remark. The Twin Prime Conjecture (open since 1846) asks if there are infinitely many pairs of twin primes.

Lemma 4.2. If p is prime and $p|ab$, then $(p|a) \vee (p|b)$.

Proof. Assume $p \nmid a$ (otherwise we are done).

Since $p|ab$, then $ab = pk$ for some $k \in \mathbb{Z}$.
Since $p \nmid a$ then p, a are coprime so $\exists s, t \in \mathbb{Z}$ such that $1 = sa + tp$.
Then we have $b = sab + tpb = spk + tpb = p(sk + tb)$.

■

Remark. This works generally, if $p|abc\dots$ then $(p|a) \vee (p|b) \vee (p|c) \vee \dots$ by induction. The strategy is to absorb the first $n-1$ numbers as a so that we have the form ab. Then apply the I.H. on the first $n-1$ numbers and apply the lemma above on ab.

Theorem 4.7 (Fundamental Theorem of Arithmetic)**.** Let $n \geq 2$ be an integer. Then

(1) There *exists* prime numbers $p_1 \leq p_2 \leq \cdots \leq p_k$ such that

$$n = p_1 p_2 \cdots p_k$$

(2) This prime decomposition is *unique*.

Proof. We will employ strong induction on $n \geq 2$ ($n \in \mathbb{N}$).

Base Case:
$n = 2$ is a prime decomposition of $k = 1$ and $p_1 = 2$ (existence).
Assume there is another decomposition $2 = p_1 \cdots p_k$. Then $p_i|2 \ \forall i \in [k]$, but 2 is prime, so $p_i = 2$ and we have $2 = 2^k$ which requires $k = 1$ (uniqueness).

Inductive step:
Assume FTA for all $m \in \mathbb{Z}$ where $2 \leq m < n$. We will show FTA for n.

If n is prime, then this is trivial (same argument as the base case).

If n is composite, then $n = ab$ for $2 \leq a < n$, $2 \leq b < n$. By the I.H., $a = p_1 p_2 \dots p_k$ and $b = q_1 q_2 \dots q_l$. Then we have $ab = p_1 p_2 \dots p_k q_1 q_2 \dots q_l$. Rearrange the primes in increasing order and thus a factorization *exists*. To show uniqueness, suppose we have two factorizations of n

$$n = p_1 p_2 \cdots p_k$$
$$n = q_1 q_2 \cdots q_l$$

Since $p_1|n = q_1 q_2 \dots q_l$, then by Lemma 4.2

$$(p_1|q_1) \vee (p_1|q_2) \vee \cdots \vee (p_1|q_l)$$

Since q_i are prime for all $i \in [l]$, we have

$$(p_1 = q_1) \vee (p_1 = q_2) \vee \cdots \vee (p_1 = q_l)$$

Since the primes q_i are in increasing order, then $q_1 \leq p_1$. By symmetry, repeating the above argument shows that $p_1 \leq q_1$. Therefore $p_1 = q_1$. Now let $m = \frac{n}{p_1}$ and note that $2 \leq m < n$, as n is composite. We have

$$\begin{aligned} m &= p_2 \dots p_k \\ &= q_2 \dots q_l \end{aligned}$$

By the I.H., these two decompositions are the same. Therefore, we have shown the factorization of n is unique, since $k = l$ and $p_1 = q_1, p_2 = q_2, \dots, p_k = p_l$.

■

Remark. If we regroup repeated primes and let $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$, then $p_1 < p_2 < \cdots < p_k$ where $\alpha_i \geq 1 \ \forall i \in [k]$.

Lemma 4.3. Let $a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$ and let $b = p_1^{\beta_1} p_2^{\beta_2} \dots p_k^{\beta_k}$ where $\alpha_i, \beta_i \geq 0 \ \forall i \in [k]$. Then $a | b \iff \alpha_i \leq \beta_i \ \forall i \in [k]$.

Proof.

($\Rightarrow$) Assume $a|b$ so $b = ac$ where $c = p_1^{\gamma_1} p_2^{\gamma_2} \dots p_k^{\gamma_k}$ by FTA (c cannot have new primes since it divides b). Then we have

$$\begin{aligned} b &= ac \\ &= (p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k})(p_1^{\gamma_1} p_2^{\gamma_2} \dots p_k^{\gamma_k}) \\ &= p_1^{\alpha_1+\gamma_1} p_2^{\alpha_2+\gamma_2} \dots p_k^{\alpha_k+\gamma_k} \end{aligned}$$

Since decomposition is unique, $\beta_i = \alpha_i + \gamma_i$ so $\beta_i \geq \alpha_i$ as $\gamma_i \geq 0$ ($\forall i \in [k]$).

($\Leftarrow$) Assume $\alpha_i \leq \beta_i \ \forall i \in [k]$. Let $\gamma_i = \beta_i - \alpha_i$. Then we have

$$\begin{aligned} c &= p_1^{\beta_1-\alpha_1} p_2^{\beta_2-\alpha_2} \dots p_k^{\beta_k-\alpha_k} \\ &= \frac{b}{a} \end{aligned}$$

Then $b = ac$ so $a|b$.

■

Lemma 4.4. Let $a = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$ and $b = p_1^{\beta_1} p_2^{\beta_2} \cdots p_k^{\beta_k}$. Then

$$\gcd(a, b) = p_1^{\min(\alpha_1, \beta_1)} p_2^{\min(\alpha_2, \beta_2)} \cdots p_k^{\min(\alpha_k, \beta_k)}$$

Proof. Let $\delta_i = \min(\alpha_i, \beta_i)$ and let $d = p_1^{\delta_1} p_2^{\delta_2} \cdots p_k^{\delta_k}$. By Lemma 4.3, this implies $d|a$ and $d|b$, so d is a common divisor. Now let's show it is the greatest. Suppose we have another common divisor $c = p_1^{\gamma_1} p_2^{\gamma_2} \cdots p_k^{\gamma_k}$. Since $c|a$ and $c|b$, then by Lemma 4.3, $\gamma_i \leq \alpha_i$ and $\gamma_i \leq \beta_i$ for all $i \in [k]$. That means $\gamma_i \leq \min(\alpha_i, \beta_i) = \delta_i$. So $c|d$ by Lemma 4.3 and hence $c \leq d$.

■

Theorem 4.8. There are infinitely many prime numbers.

Proof. Assume not. Then there are finitely many primes so we can list them as

$$p_1, p_2, p_3, \ldots, p_n$$

Multiply these primes together and add one to obtain

$$m = p_1 p_2 p_3 \cdots p_n + 1$$

m is larger than any prime p_i where $i \in [n]$ so m is composite. This means $p_k|m$ for some $k \in [n]$. But also $p_k|p_1 p_2 \cdots p_n$ so $p_k|m - p_1 p_2 \cdots p_n$ and hence $p_k|1$. Contradiction arises.

■

Theorem 4.9. For any $n \in \mathbb{Z}$, there are consecutive primes at least n apart. Equivalently, there are consecutive sequences of composite numbers of any length $n \in \mathbb{Z}$.

Proof. Consider the $n - 1$ numbers $n! + 2, n! + 3, \ldots n! + n$. Then for any k such that $2 \leq k \leq n$, we have $k|n!$ and $k|k$ so $k|n! + k$.

■

Remark. This argument needs to begin at $n! + 2$ and $k = 2$ since divisibility by 1 cannot show that a number is composite.

Theorem 4.10 (Prime Number Theorem)**.** Let $\Pi(n)$ be the number of primes p where $p \leq n$ and $n \geq 2$. Then $\Pi(n) \sim \frac{n}{\ln n}$, which means $\lim_{n \to \infty} \frac{\Pi(n)}{n/\ln n}$ is a constant.

4.3 Modular Arithmetic

Definition 4.6 (Congruence modulo)**.** Let $a, b \in \mathbb{Z}$ and $n \in \mathbb{N}$. We say a, b are congruent modulo n ($a \equiv b \pmod{n}$) if $a = kn + b$ for some $k \in \mathbb{Z}$. Equivalently, $a \equiv b \pmod{n}$ if $n | a - b$.

Definition 4.7 (Modulo operator)**.** For $c, d \in \mathbb{Z}$, $c\%d = r$ where $c = qd + r$ by the division algorithm.

Definition 4.8 (Congruence class)**.** $[a]_n$ is a congruence class such that $[a]_n = \{x \in \mathbb{Z} : x \equiv a \pmod{n}\}$. There are n congruence classes modulo n, denoted as

$$\mathbb{Z}_n = \mathbb{Z} \setminus n\mathbb{Z} = \{[0]_n, [1]_n, \dots, [n-1]_n\}$$

Definition 4.9 (Inverse)**.** Given $a \in \mathbb{Z}$, an inverse of $a \pmod{n}$ is a number b such that

$$ba \equiv ab \equiv 1 \pmod{n}$$

Theorem 4.11. $a \equiv b \pmod{n} \iff a\%n = b\%n$

Proof.

($\Rightarrow$) Assume $a \equiv b \pmod{n}$.
Then $a = kn + b$ for some $k \in \mathbb{Z}$.
By the division algorithm, $\exists q, r \in \mathbb{Z}$ s.t. $a = qn + r$ and $0 \le r < n$.
So $r = a\%n$.
We have $qn + r = kn + b$, so $b = (q - k)n + r$.
Since $0 \le r < n$ and r is unique, then $r = b\%n$ too.

($\Leftarrow$) Assume $a\%n = b\%n$.
Then $a = qn + r$ and $b = kn + r$ for some $q, k \in \mathbb{Z}$ and $0 \le r < n$.
Then $a - qn = b - kn$ so $a = (q - k)n + b$.
Then $a \equiv b \pmod{n}$ by definition.

■

Theorem 4.12. Congruence modulo is an equivalence relation on $\mathbb{Z}$, so it satisfies the following properties

(1) *Reflexivity*

$$x \equiv x \pmod{n}$$

(2) *Symmetry*

$$x \equiv y \pmod{n} \implies y \equiv x \pmod{n}$$

(3) *Transitivity*

$$(x \equiv y \pmod{n} \land y \equiv z \pmod{n}) \implies (x \equiv z \pmod{n})$$

Theorem 4.13. Congruence modulo is compatible with + and ×. If $x \equiv a \pmod{n}$ and $y \equiv b \pmod{n}$, then

$$(+) \qquad x + y \equiv a + b \pmod{n}$$

$$(\times) \qquad x \cdot y \equiv a \cdot b \pmod{n}$$

Remark. Congruence modulo is **not** compatible with exponentiation. This means

$$(a \equiv b \pmod{n} \land c \equiv d \pmod{n}) \not\Longrightarrow a^c \equiv b^d \pmod{n}$$

Theorem 4.14 (Unique inverse)**.** If $a \in \mathbb{Z}$ has an inverse, then it is unique modulo n.

Theorem 4.15 (Inverse properties)**.**

(1) $(a^{-1})^{-1} \equiv a \pmod{n}$

(2) $(ab)^{-1} \equiv b^{-1}a^{-1} \equiv a^{-1}b^{-1} \pmod{n}$

Theorem 4.16. a is invertible modulo n *if and only if* $\gcd(a, n) = 1$.

Remark. If a is invertible or if $\gcd(a, n) = 1$, then to find a^{-1}, we can reverse Euclid's algorithm to find the Bézout coefficient on a, which is a^{-1}.

4.4 Prime Modular Arithmetic

Lemma 4.5. If p is prime, then $\forall a \in \mathbb{Z}$, either $p|a$ or $\gcd(a, p) = 1$.

Theorem 4.17. If p is prime, then a is invertible modulo p *if and only if* $a \not\equiv 0 \pmod{p}$.

Proof.

$$\begin{aligned} a \not\equiv 0 \pmod{p} &\iff p \nmid a \\ &\iff \gcd(a,p) = 1 \\ &\iff a \text{ is invertible} \pmod{p} \end{aligned}$$

■

Corollary 4.17.1. Every congruence class of $\mathbb{Z}_p$, except 0, is invertible.

Remark. We call such a set $\mathbb{Z}_p$ a field.

Theorem 4.18 (Integrity Theorem)**.** If p is prime, then

$$ab \equiv 0 \pmod{p} \implies (a \equiv 0 \pmod{p}) \vee (b \equiv 0 \pmod{0})$$

Theorem 4.19 (Fermat's Little Theorem)**.** If p is prime, then

(1) If $a \not\equiv 0 \pmod{p}$ then $a^{p-1} \equiv 1 \pmod{p}$.

(2) $a^p \equiv a \pmod{p}$

Theorem 4.20 (Converse of FLT)**.** Suppose $a^{n-1} \equiv 1 \pmod{n}$ for all a s.t. $1 \leq a < n$. Then n is prime.

Remark. A Carmichael number is an odd composite n satisfying $a^{n-1} \equiv 1 \pmod{n}$ for every a where $\gcd(a,n) = 1$. A pseudoprime is a number a such that $a^{p-1} \equiv 1 \pmod{p}$ but a is composite.

4.5 Primality Testing

Lemma 4.6. If n is composite, then there exists a prime factor $p \leq \sqrt{n}$.

Proof. Assume not. Then there are at least 2 prime factors $p > \sqrt{n}$ and $q > \sqrt{n}$. Then $n \geq p \cdot q > \sqrt{n} \cdot \sqrt{n} = n$. Contradiction arises.

■

Lemma 4.7. If p is prime and $x^2 \equiv 1 \pmod{p}$, then $x \equiv \pm 1 \pmod{p}$.

Proof. Assume $x^2 \equiv 1 \pmod{p}$. Then $x^2 - 1 \equiv 0 \pmod{p}$ so $(x-1)(x+1) \equiv 0 \pmod{p}$. By the Integrity Theorem, $x + 1 \equiv 0 \pmod{p}$ or $x - 1 \equiv 0 \pmod{p}$ so $x \equiv \pm 1 \pmod{p}$.

■

4.6 Cryptography

Lemma 4.8. For primes p and q

$$a \equiv b \pmod{pq} \iff (a \equiv b \pmod{p}) \wedge (a \equiv b \pmod{q})$$

Proof.

($\Rightarrow$) Assume $a \equiv b \pmod{pq}$. Then $pq|a-b$. Since $p|pq$ and $q|pq$, then $p|a-b$ and $q|a-b$ by transitivity, and the result follows.

($\Leftarrow$) Assume $a \equiv b \pmod{p}$ and $a \equiv b \pmod{q}$. Then $p|a-b$ and $q|a-b$. Since p and q are both prime, then they are prime factors in the prime decomposition of $a-b$. Hence $pq|a-b$ so $a \equiv b \pmod{pq}$.

■

4.6.1 RSA Encryption

First generate the public and secret keys.

(1) Generate 2 large primes, p and q. Let $n = pq$.

(2) Generate the public key k such that $\gcd(k, (p-1)(q-1)) = 1$.

(3) Generate the secret key $s = k^{-1} \pmod{(p-1)(q-1)}$.

Then encrypt and decrypt messages.

(1) Encryption: $\hat{M} \equiv M^k \pmod{n}$

(2) Decryption: $M \equiv \hat{M}^s \pmod{n}$

An agent A would keep (s, p, q) private, and share (k, n) publicly.

4.6.2 Fast Exponentiation

If we want to evaluate $a^b \pmod{n}$, we can create a table for computing this faster.

	a	a^2	a^4	a^8	a^{16}	$\cdots$
mod n	r_1	r_2	r_3	r_4	r_5	$\cdots$

To determine r_i, square r_{i-1} modulo n. Multiply as many terms as needed in the top row to achieve a^b. Multiply the corresponding terms in the bottom row modulo n to achieve $a^b \pmod{n}$.

5 Combinatorics

5.1 Functions

Definition 5.1 (Function)**.** $f : A \to B$ is a function from a set A (domain) to a set B (codomain) if for each element $x \in A$ it assigns a specific element $y \in B$.

Definition 5.2 (Injective function)**.** A function $f : A \to B$ is injective if $f(a) = f(b) \implies a = b$. In other words, $\forall y \in B$, $y = f(x)$ for at most one $x \in A$.

Definition 5.3 (Surjective function)**.** A function $f : A \to B$ is surjective if $\forall y \in B, \exists x \in A$ s.t. $f(x) = y$. In other words, $\forall y \in B$, $y = f(x)$ for at least one $x \in A$.

Definition 5.4 (Bijective function)**.** A function $f : A \to B$ is bijective if it is injective and surjective. In other words, $\forall y \in B$, $y = f(x)$ for exactly one $x \in A$.

Definition 5.5 (Function composition)**.** The composition of $f : X \to Y$ with $g : Y \to Z$ is a function $h : X \to Z$ where $h(x) = (g \circ f)(x) = g(f(x))$.

Definition 5.6 (Inverse function)**.** A function $f : A \to B$ has an inverse function $f^{-1} : B \to A$ if:

1. $(f^{-1} \circ f)(x) = x \quad \forall x \in A$
2. $(f \circ f^{-1})(y) = y \quad \forall y \in B$

Theorem 5.1. If $f : A \to B$ has an inverse $f^{-1} : B \to A$, then f^{-1} is unique.

Proof. Assume that $g : B \to A$ and $h : B \to A$ are both inverse functions of $f : A \to B$. Then:

1. $(g \circ f)(x) = x$ and $(h \circ f)(x) = x \quad \forall x \in A$
2. $(f \circ g)(y) = y$ and $(f \circ h)(y) = y \quad \forall y \in B$

Suppose $y \in B$. Then:

$$\begin{aligned} g(y) &= g(f \circ h)(y) \\ &= g(f(h(y))) \\ &= (g \circ f)(h(y)) \\ &= h(y) \end{aligned}$$

It follows that $g = h$.

■

Theorem 5.2. A function is invertible *if and only if* it is bijective.

Theorem 5.3. If $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ are bijective functions, then $f \circ g$ is also bijective.

5.2 Counting

Definition 5.7 (Set cardinality)**.** A set A has cardinality k if there exists a bijection $f : [k] \longrightarrow A$.

Definition 5.8 (k-permutation)**.** A k-permutation of a set A with $|A| = n$ is an ordered arrangement of k members A. An n-permutation is simply called a permutation. The number of k-permutations in A is:

$$P(n, k) = \frac{n!}{(n-k)!}$$

Definition 5.9 (k-combination)**.** A k-combination of a set A with $|A| = n$ is a subset of k. The number of k-combinations in A is:

$$C(n, k) = \frac{P(n, k)}{k!} = \frac{n!}{k!(n-k)!}$$

Remark. $C(n, k) = \frac{P(n,k)}{P(k,k)}$

Definition 5.10 (Derangement)**.** A derangement is a permutation $(x_1, x_2, \cdots, x_n)$ of the set $[n]$ such that $\forall i \in [n]\ x_i \neq i$.

Remark. For example, $(1, 3, 4, 5, 2)$ is not a derangement.

Theorem 5.4 (Bijection Principle)**.** $|A| = |B|$ *if and only if* there exists a bijection $f : A \longrightarrow B$.

Theorem 5.5 (Product Principle)**.**

$$|A \times B| = |A| \cdot |B|$$

Theorem 5.6 (Sum Principle)**.** If $A \cap B = \phi$, then

$$|A \cup B| = |A| + |B|$$

Theorem 5.7 (Complement principle)**.** If $B \subseteq A$, then

$$|A \setminus B| = |A| - |B|$$

Theorem 5.8 (Inclusion-Exclusion Principle)**.**

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Theorem 5.9 (General Inclusion-Exclusion Principle)**.**

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{k=1}^{n} (-1)^{k+1} \sum_{\substack{s \subseteq [n] \\ |s| = k}} \left| \bigcap_{i \in s} A_i \right|$$

Theorem 5.10 (Stars and Bars/Balls and Urns)**.** The number of ways of placing n *identical balls* into k *distinct urns* is:

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

Proof. The problem amounts to counting the number of binary strings with n stars and $k-1$ bars. Stars represent balls and bars separate balls (functioning as urns). Among $n+k-1$ locations, we choose where to put stars, and the remaining locations hold bars (or vice versa). For example, with $n = 10$ and $k = 5$, here is one possible configuration:

$$* * \mid * * * \mid * \mid\mid * * * \mid *$$

■

Theorem 5.11 (Pascal's identity)**.**

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

Proof. (Combinatorial proof) The LHS counts the number of k-subsets of $[n+1]$. On the RHS, $\binom{n}{k-1}$ counts the number of k-subsets of $[n+1]$ where $n+1$ is included. Meanwhile, $\binom{n}{k}$ counts the number of k-subsets of $[n+1]$ where $n+1$ is not included. When combined, these two disjoint cases count the same problem as the LHS. Therefore LHS = RHS.

■

Remark. Pascal's identity provides a recursive algorithm for computing binomial coefficients in Pascal's triangle. Using it, we can break down the computation of large factorials.

Theorem 5.12 (Binomial theorem)**.**

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

Proof. To compute each term of $(x+y)^n = (x+y)(x+y)\cdots(x+y)$, x needs to be selected k times, where $k \in [n]$. In every remaining bracket, we select y. There are $\binom{n}{k}$ ways to select x k times in this fashion.

■

Theorem 5.13 (Derangement formula)**.** The number of derangements of $[n] = [1, \dots, n]$ is:

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

Proof. Let X be the set of all permutations of $[n]$, and let A_i be the permutations of $[n]$ with i in position i. Then:

$$\begin{aligned}
D_n &= |X \setminus (A_1 \cup A_2 \cup \cdots \cup A_n)| \\
&= |X| - |A_1 \cup A_2 \cup \cdots \cup A_n| && \text{(complement principle)} \\
&= n! - \sum_{k=1}^{n} (-1)^{k+1} \sum_{\substack{s \subseteq [n] \\ |s| = k}} \left| \bigcap_{i \in s} A_i \right| && \text{(general I.E.)} \\
&= n! - \sum_{k=1}^{n} (-1)^{k+1} \sum_{\substack{s \subseteq [n] \\ |s| = k}} (n-k)! && \text{(see note)} \\
&= n! + \sum_{k=1}^{n} (-1)^{k} (n-k)! \sum_{\substack{s \subseteq [n] \\ |s| = k}} 1 \\
&= n! + \sum_{k=1}^{n} (-1)^{k} (n-k)! \frac{n!}{k!(n-k)!} \\
&= n! + n! \sum_{k=1}^{n} \frac{(-1)^k}{k!} \\
&= n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}
\end{aligned}$$

Note: $\bigcap_{i \in S} A_i$ fixes $|s| = k$ elements in their natural position, so there are $(n-k)!$ ways to order the other elements. Hence $\left|\bigcap_{i \in S} A_i\right| = (n-k)!$.

■

Remark. This is relevant to the Hat-Check Problem. Guests arriving at a large party leave their hat at the entrance. When they leave, they each take a hat randomly - what is the probability that no one leaves with their own hat?

Theorem 5.14 (Pigenhole Principle)**.** Suppose we wish to place pigeons in holes. If there are more pigeons than holes, then at least one hole contains at least 2 pigeons. More generally, for N pigeons and k holes, at least one hole contains at least $\lceil \frac{N}{k} \rceil$ pigeons.

Proof. Suppose we could fit less than $\lceil \frac{N}{k} \rceil$ pigeons per hole. Since $N = \sum_{holes} pigeons$, it follows that:

$$N = \sum_{holes} pigeons \leq k(\lceil \frac{N}{k} \rceil - 1) < k(\frac{N}{k}) = N$$

Contradiction arises.

■

5.3 Recurrence relations

Definition 5.11 (Recurrence relation)**.** A recurrence relation of a sequence is defined in two steps:

(1) Base case(s): Define initial values $a_0, a_1, \dots$

(2) Recursive step: Define a_n in terms of previous values $a_{n-1}, a_{n-2}, \dots$ in the sequence.

Theorem 5.15. Given two particular solutions $(p_n)_{n\geq 0}$ and $(q_n)_{n\geq 0}$ to the recurrence relation $a_n = f(n)a_{n-1} + g(n)$, their difference is a solution to the homogeneous equation $a_n = f(n)a_{n-1}$.

Proof. Using the fact p_n and q_n are particular solutions:

$$\begin{aligned} p_n &= f(n)p_{n-1} + g(n) \\ q_n &= f(n)q_{n-1} + g(n) \\ \implies p_n - q_n &= f(n)(p_{n-1} - q_{n-1}) \end{aligned}$$

■

Remark. It follows that the general solution to a first order linear recurrence relation is $a_n = h_n + p_n$.

Theorem 5.16. The solutions to a homogeneous recurrence relation of order k form a k-dimensional subspace of $\mathbb{R}^n$.

Remark. It follows that multiplying any solution by a scalar provides a solution, and so does adding any two solutions.

Linear recurrence relations, First Order

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

(1) **Solve the homogeneous equation**
Set $g(n) = 0$ to determine a solution h_n.

(2) **Find a particular solution**
Apply the *Method of Undetermined Coefficients*:

Form of $g(n)$	Form of guess
Polynomial (degree d)	$p_n = A_d n^d + A_{d-1} n^{d-1} + \cdots + A_0$
Exponential (base t)	$p_n = At^n$
Sum/Product of above	Sum/Product of guesses

Note: If the guess solves the homogeneous equation, then multiply by n.

(3) **Determine the general solution**
The general solution is always $\boldsymbol{a_n = h_n + p_n}$.

(4) **Apply initial conditions**
Determine unknown constants.

Linear recurrence relations, Higher Order

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

(1) **Solve the homogeneous equation**
Set $g(n) = 0$. Assume $a_n = x^n$ to find the *characteristic polynomial*:

$$x^k = c_1 x^{k-1} + c_2 x^{k-2} + \cdots + c_k$$

If all roots $\{x_1, x_2, \ldots, x_k\}$ are distinct, the homogeneous solution is:

$$a_n = A_1 x_1^n + \cdots + A_k x_k^n$$

If roots in the above equation are not distinct, multiply each term containing a repeated root by an increasing power of n (each term will remain a solution).

The remaining procedure is the same as the first order case.

Remark. The policy for dealing with repeated roots generates linearly independent solutions that span the entire solution subspace. If we instead accept repeated roots, our solution would not span the full subspace, and hence would not be general.

6 Graph Theory

6.1 Introduction

Definition 6.1 (Graph)**.** A graph $G = (V, E)$ is a collection of sets where V is a nonempty set of vertices and E is a set of edges $\{A, B\}$ where $A, B \in V$.

Remark. In a *directed* graph, each edge's vertex pair is ordered, and in an *undirected* graph, the pairs are unordered. In this course, we studied undirected graphs.

Definition 6.2 (Multigraph)**.** A multigraph $G = (V, E)$ is defined similarly as a graph, except E is a multiset.

Definition 6.3 (Degree)**.** The degree of a vertex $A \in V$, denoted $\deg(A)$, is the number of edges $e \in E$ such that $A \in e$. Self-loops count twice.

Definition 6.4 (Subgraph)**.** $H = (W, F)$ is a subgraph of $G = (V, E)$ if $W \subseteq V$ and $F \subseteq E$. We write $H \subseteq G$.

Definition 6.5 (Neighbourhood)**.** Given $G = (V, E)$, and given $S \subseteq V$, the neighbourhood of S is the set:

$$N(S) = \{x \in V : \exists s \in S \text{ s.t. } \{x, s\} \in E\}$$

Definition 6.6 (Graph homomorphism)**.** Given $G = (V, E)$ and $H = (W, F)$, a graph homomorphism $\phi : G \to H$ is a function $\phi : V \to W$ such that:

$$\phi(G) = \{\phi(V), \phi(E)\} \subseteq H$$

Note that $\phi(V) = \{\phi(x) : x \in V\}$ and $\phi(E) = \{\{\phi(a), \phi(b)\} : a, b \in E\}$.

Definition 6.7 (Graph isomorphism)**.** Given $G = (V, E)$ and $H = (W, F)$, a graph isomorphism is a graph homomorphism where $\phi : V \to W$ is a bijection which induces a bijection $\phi : E \to F$, such that:

$$\{a, b\} \in E \iff \{\phi(a), \phi(b)\} \in F$$

We say G and H are isomorphic ($G \simeq H$) if there exists an ismorphism between them.

Definition 6.8 (Simple graph)**.** A graph is simple if it has no self loops. K_n, C_n, L_n defined below are all simple.

Definition 6.9 (Complete graph)**.** $K_n = (V, E)$ is complete if exactly one edge connects every pair of distinct vertices.

Remark. K_n has $\binom{n}{2}$ edges, since its edges are all the 2-element subsets of V.

Definition 6.10 (Cyclic graph)**.** $C_n = (V, E)$ for $V = [n]$ and $E = \{\{i, j\} : i, j \in [n], j = i + 1 \pmod n\}$.

Definition 6.11 (Linear graph)**.** $L_n = (V, E)$ for $V = [n]$ and $E = \{\{i, i+1\} : 1 \leq i < n\}$.

Definition 6.12 (Cycle)**.** A cycle in G is a subgraph $H \simeq C_n$.

Definition 6.13 (Path)**.** A path in G is a subgraph $H \simeq L_n$.

Definition 6.14 (Clique)**.** A clique in G is a subgraph $H \simeq K_n$.

Definition 6.15 (Walk)**.** A walk in $G = (V, E)$ is a list of vertices $\{v_0, v_1, v_2, \cdots, v_k\}$ such that $\{v_i, v_{i+1}\} \in E$ for all $0 \leq i < k$. It is a homomorphism $w : L_n \to G$.

Remark. A walk does not need to be bijective.

Definition 6.16 (Connected graph)**.** $G = (V, E)$ is connected if for every $a, b \in V$, there exists a path (or walk) between a and b.

Definition 6.17 (Connected component)**.** A connected component H of G is a maximal connected subgraph, which means no other connected subgraph contains H.

Theorem 6.1 (Handshaking lemma)**.** For $G = (V, E)$, we have $\sum_{x \in V} \deg(x) = 2|E|$ (even if multiple edges and loops are present).

Proof. We will count incident edges. Since each vertex x has $\deg(x)$ edges incident to it, the total number of incident edges to vertices is $\sum_{x \in V} \deg(x)$. Since each edge is incident to 2 vertices, the total number of incident edges to vertices is also $2|E|$.

■

Corollary 6.1.1. In G, the number of vertices with odd degree is even.

Proof. Assume that the number of vertices with odd degree is odd. Then $\sum_{a \in V} \deg(a) = 2n + 1$ where a has odd degree. We know that $\sum_{b \in V} \deg(b) = 2m$ where b has even degree. But since $\sum_{a \in V} \deg(a) + \sum_{b \in V} \deg(b) = \sum_{x \in V} \deg(x)$, this implies that $\sum_{x \in V} \deg(x)$ is odd. Contradiction arises.

■

Theorem 6.2. There exists a path in $G = (V, E)$ with end points $a, b \in V$ *if and only* if there exists a walk with end points a, b.

Proof.

- (⇒) A path is a walk since an isomorphism is a homomorphism.
- (⇐) Assume there exists a walk with end points $a, b \in V$. If a vertex is encountered more than once, we can "delete" vertices in between (e.g. ac~~dec~~b), and retain a smaller walk. Since walks are finite, this process will eventually halt, leaving us with a path.

■

6.2 Trees and Forests

Definition 6.18 (Forest)**.** A forest is a graph with no cycles.

Definition 6.19 (Tree)**.** A tree is a connected forest.

Definition 6.20 (Leaf)**.** A leaf in a tree T is a vertex of degree 1.

Definition 6.21 (Spanning tree)**.** Given a graph G, a spanning tree of $G = (V_1, E_1)$ is a subgraph $T \subseteq G$, where $T = (V_2, E_2)$ is a tree and $V_1 = V_2$.

Lemma 6.1. Every tree with at least 2 vertices has at least 2 leaves.

Proof. For a tree T_n with $n \geq 2$, take the longest path $V_1 V_2 V_3 \cdots V_{k-1} V_k$. Assume that V_k is not a leaf, so $\deg(V_k) \geq 2$, which implies V_k has another neighbour $x \neq V_{k-1}$. Since we have chosen the longest path, x must not be a new vertex. So $x = V_i$ for $1 \leq i < k - 1$. But then we have a cycle. Contradiction arises since trees have no cycles, so V_k is a leaf. This argument applies to V_1 as well, so we have two leaves.

■

Theorem 6.3. If a tree T has n vertices, then it has $n - 1$ edges. The converse follows as a corollary of Theorem 6.4.

Proof. We will employ proof by induction on $n \in \mathbb{N}$.
Base case: If $n = 1$, then T has 0 edges, which passes.
Inductive step: Assume true for n vertices. Let T be a tree with $n + 1$ vertices. By Lemma 6.1, T has at least 1 leaf. Suppose we delete a leaf in T and the attached edge. Then the new tree T' has n vertices and, by the induction hypothesis, T' has $n - 1$ vertices. If we add the deleted leaf back, then we have $(n - 1) + 1 = n$ edges.

■

Theorem 6.4. G is connected *if and only if* G has a spanning tree.

Proof.

($\Leftarrow$) Assume $G = (V_1, E_1)$ has a spanning tree $T = (V_2, E_2)$. Since T is connected, then for every $a, b \in V_2$, there exists a path between a and b. Since $V_1 = V_2$, we thus have a path between any two vertices in G, so G is connected.

($\Rightarrow$) Assume that G is connected.

Case 1: G is a tree. Then G has a spanning tree since it is its own spanning tree.

Case 2: G is not a tree. Then G contains a cycle. We can remove an edge in the cycle and preserve connectivity of the new graph G' (why?), which has fewer edges now. Continue this process for G'. Since the edge set of G is finite, this process will halt, and we will be left with a tree, since there are no more cycles. Since the vertex set was untouched, we thus found a spanning tree of G.

■

Corollary 6.4.1. If a connected graph has n vertices and $n - 1$ edges, then the graph is a tree (converse of Theorem 6.3).

Proof. Suppose we have a connected graph $G = (V_1, E_1)$ with n vertices and $n - 1$ edges. Since G is connected, a spanning tree $T = (V_2, E_2)$ exists. T has n vertices, so it has $n - 1$ edges by Theorem 6.3. Since $T \subseteq G$, we have $E_2 \subseteq E_1$. But also $|E_1| = |E_2|$, so it must be that $E_1 = E_2$. Therefore $G = T$.

■

6.3 Bipartite Graphs

Definition 6.22 (Bipartite graph)**.** $G = (L \cup R, E)$ for $L \cap R = \phi$ is bipartite if $\nexists \{x, y\} \in E$ such that $\{x, y\} \subseteq L$ or $\{x, y\} \subseteq R$. This means edges must only occur between L and R.
Equivalently, G is bipartite if we can colour its vertices in 2 colours, such that 2 adjacent vertices never have the same colour.

Definition 6.23 (Complete bipartite graph)**.** A bipartite graph $G = (L \cup R, E)$ is complete bipartite if every vertex in L is connected to every vertex in R.

Theorem 6.5. All trees are bipartite.

Proof. We use induction on $n \in \mathbb{N}$, the number of vertices in a tree T_n.
Base case: If $n = 1$, then T_1 only has 1 vertex so it is bipartite.
Inductive step: Assume that any tree T_n is bipartite, and denote this statement $P(n)$. Let T_{n+1} be any tree. Since T_{n+1} is a tree, it contains at least 1 leaf. If we delete any single leaf in T_{n+1}, then we have a smaller graph with n vertices. This smaller graph is a tree T_n since it has no cycles (we only deleted a leaf, and T_{n+1} did not have any cycle). By the I.H., T_n is bipartite. We can put this leaf back and colour it opposite to its lone neighbour.

■

Theorem 6.6. A cycle C_k is bipartite *if and only if* k is even.

Theorem 6.7. G is bipartite *if and only if* G contains no odd cycles.

Proof.

(⇒) Assume G is bipartite. Then no odd cycle can exist, because a valid 2-colouring of G induces a valid 2-colouring of any subgraph of G.

(⇐) Assume G has no odd cycles. Let $H \subseteq G$ be a connected component (since G may be disconnected) and take a spanning tree T of H. We can 2-colour T greedily (pick any vertex and recursively colour adjacent vertices). Now suppose the 2-colouring of T is invalid on G. That is, $\exists a, b \in V(T)$ with the same colour, but $\{a, b\} \in E(G)$. Since $\text{colour}(a) = \text{colour}(b)$, then for any $r \in V(T)$

$$\text{dist}(a, r) \equiv \text{dist}(b, r) \pmod 2$$

Now consider the cycle formed by adding the edge $\{a, b\}$ to T. The length is

$$\begin{aligned} \text{dist}(a, r) + 1 + \text{dist}(b, r) &\equiv 2\text{dist}(a, r) + 1 \pmod 2 \\ &\equiv 1 \pmod 2 \end{aligned}$$

So the cycle is odd, which is a contradiction.

■

Remark. The "dist(x, y)" function returns the number of edges between vertices x and y.

6.4 Matchings

Definition 6.24 (Matching)**.** A matching in $G = (V, E)$ is a set of edges $M \subseteq E$ such that $e, f \in M \implies e \cap f = \phi$.

Definition 6.25 (Perfect matching)**.** A perfect matching is a matching $M \subseteq E$ that covers all vertices.

Theorem 6.8 (Hall's Marriage Theorem)**.** *Suppose we have a bipartite graph $G = (L \cup R, E)$ and $|L| = |R|$. Then there exists a perfect matching in G if and only if $\forall S \subseteq L$, $|N(S)| \geq |S|$.*

6.5 Eulerian Graphs

Definition 6.26 (Closed walk)**.** A walk is closed if $v_0 = v_k$.

Definition 6.27 (Trail)**.** A trail in $G = (V, E)$ is a walk which never visits an edge more than it appears in E.

Definition 6.28 (Euler trail)**.** An Euler trail contains all edges from G (possibly repeated if G is a multigraph).

Definition 6.29 (Euler circuit)**.** An Euler circuit is a closed Euler trail.

Remark. We say a graph is Eulerian if it has an Euler circuit.

Lemma 6.2. *If G is a graph where all vertices have even degree then there exist edge-disjoint cycles (cycles with no common edges) such that*

$$G = C_1 \cup C_2 \cup \cdots \cup C_k$$

Proof. We employ induction on the number of edges.
Base case: If the graph has no edges, then this is true.
Inductive step: Assume true for all graphs with fewer edges than m, where vertices all have even degree. Consider a graph G with m edges and even degree vertices. Since no vertex of degree 1 exists, G must not be a tree and hence contains a cycle C_1. Remove all edges in C_1 to obtain $G \setminus C_1$, whose vertices are all of even degree. By the I.H. we have $G \setminus C_1 = C_2 \cup C_3 \cup \cdots \cup C_k$. Add the edges back to the cycle C_1 to obtain $G = C_1 \cup C_2 \cup \cdots \cup C_k$.

■

Theorem 6.9. A connected graph G has an Euler circuit *if and only if* all vertices in G have even degree.

Proof.

(⇒) Assume a connected graph G is Eulerian. Then there exists a closed walk visiting every edge exactly once. Every time a vertex is entered, it must also be exited. Thus the degree of every vertex must be even.

(⇐) We show this by strong induction on the number of edges. In the base case, we have 1 vertex and no edge, and the empty walk is Eulerian. For the inductive step, assume that if a graph H with fewer vertices than a graph G has all vertices of even degree and is connected, then H is Eulerian. Now consider G, which is also connected with all vertices of even degree. No degree 1 vertex exists so G is not a tree and so a cycle C exists. Obtain $G' = G \setminus C$ by removing the edges of C but preserving the vertices. Then $G' = H_1 \cup H_2 \cup \cdots \cup H_k$ where H_i are connected components of G'. By the I.H., each component H_i is Eulerian. Add back the edges to C and now walk around C. Each time a component H_i is encountered, traverse the entire Euler circuit of H_i before continuing on the walk around C. Once the walk is complete, we will have defined an Euler circuit on G.

■

Theorem 6.10. A connected graph G has a non-closed Euler trail *if and only if* exactly two vertices have odd degree (corresonding to the endpoints).

Proof.

(⇒) Assume a connected graph G has a non-closed Euler trail. By the contrapositive, if the endpoints have even degree, then we would have an Euler circuit (since all vertices are then even), which is closed.

(⇐) Assume exactly two vertices x and y have an odd degree. Connect x and y by a new edge such that all vertices are now of even degree. Then an Euler circuit exists. Remove the edge between x and y to obtain a non-closed Euler trail.

■

6.6 Hamiltonian Graphs

Definition 6.30 (Hamilton path)**.** A Hamilton path visits all vertices in G.

Definition 6.31 (Hamilton cycle)**.** A Hamilton cycle visits all vertices in G.

Definition 6.32 (Hamiltonian graph)**.** G is Hamiltonian if it contains a Hamilton cycle.

Remark. The Petersen graph is not Hamiltonian. It is relatively simple to check if a graph is Eulerian, since we have a necessary and sufficient condition in the form of Theorem 6.9. For Hamiltonian graphs, such a biconditional theorem does not exist. In fact, checking if a graph is Hamiltonian is an NP-complete problem.

Theorem 6.11. If G is Hamiltonian, then $\forall x \in V(G), \deg(x) \geq 2$.

Proof. To complete a Hamilton cycle, we must enter every vertex and leave it. Therefore each vertex needs at least 2 incident edges.

■

Remark. As such, all trees are not Hamiltonian.

Theorem 6.12. $K_{m,n}$ is Hamiltonian $\iff m = n$.

Proof. (Sketch)

($\Rightarrow$) Assume $K_{m,n} = (L \cup R, E)$ is Hamiltonian. Assume $m \neq n$ towards a contradiction. WLOG suppose L contains fewer vertices than R. Whether we begin on L or R, eventually all L vertices will have been visited, and we will be stuck on R unable to complete the cycle.

($\Leftarrow$) Assume $m = n$. Alternate between L and R and there will be an edge to complete the cycle at the end.

■

Corollary 6.12.1. If $G = (L \cup R, E)$ is bipartite and Hamiltonian, then $|L| = |R|$.

Proof. Let $m = |L|$ and $n = |R|$. Since G is Hamiltonian, then that induces a Hamiltonian cycle on $K_{m,n}$. By Lemma 6.3, $K_{m,n} \iff m = n$.

■

Theorem 6.13. If G is Hamiltonian, then G is connected.

Proof. Since G is Hamiltonian, there must exist a cycle visiting all vertices exactly once. But this is impossible if G contains components that are not connected to each other.

■

Theorem 6.14. If $G = (V, E)$ is Hamiltonian (hence connected), then for $S \subseteq V$, $G \setminus S$ contains at most $|S|$ connected components.

Proof. Let $k = |S|$ and suppose that $G \setminus S$ contains at least $k+1$ connected components (by definition they are maximally connected, so all disjoint)

$$H_1 \cup H_2 \cup \cdots \cup H_{k+1}$$

Let C be a Hamiltonian cycle on G. We will use the general PHP. Let pigeons be edges of C between S and $G \setminus S$. Let holes be vertices on S that are incident to such edges. There are k holes by definition (why?). There are at least $2(k+1)$ pigeons, since we need an edge to enter each component and a different edge to leave. By the general PHP, there exists a vertex on C with more than $\lceil \frac{2(k+1)}{k} \rceil = 3$ incident edges. This contradicts the fact that C is a cycle.

■

Remark. This condition is not sufficient, for example take the Petersen graph.

Theorem 6.15 (Dirac's Theorem)**.** If a simple connected graph $G = (V, E)$ has $|V| \geq 3$ and $\deg(x) \geq \frac{n}{2}$ $\forall x \in V$, then G is Hamiltonian.

Remark. This condition is not necessary, for example take the dodecahedral graph.

6.7 Planar Graphs

Definition 6.33 (Planar graph)**.** G is planar if it can be drawn on the plane without intersection of any pair of edges.

Definition 6.34 (Faces)**.** The faces of a planar graph are the regions delimited by edges. The "exterior" region is called the outer face.

Definition 6.35 (Dual graph)**.** Any planar drawing $G = (V, E)$ induces another planar drawing $G^* = (F, E^*)$, which is called the dual.
In the dual G^*:

(1) There is a vertex for each face of G

(2) There is an edge for each edge of G. Particularly, if $e \in G$ is on faces F_1 and F_2, then (F_1, F_2) is an edge of G^*.

Definition 6.36 (Graph Minor)**.** H is a minor of G (denoted $H \leq G$) if we can reach H from G by applying a sequence of operations:

(1) Vertex deletion: *Delete all incident edges as well.*

(2) Edge deletion

(3) Edge contraction: *Replace vertices u and v with a new vertex w adjacent to all neighbours of u and v.*

Remark. All of these operations preserve planarity of G, so if G is planar, then any $H \leq G$ is also planar.

Theorem 6.16 (Euler's Formula)**.** If G is connected and planar, then $e + 2 = n + f$ where $n = |V(G)|$, $e = |E(G)|$, and f is the number of faces.

Proof. We perform induction on f.

Base case: If $f = 1$, then the outer face is the only face. This means there is no cycle in G, since that would delimit at least one other face. Then G is a tree, so $e = n - 1$. The base case then passes:

$$\begin{aligned} e + 2 &= n + f \\ n - 1 + 2 &= n + f \\ n + 1 &= n + 1 \end{aligned}$$

Inductive step: Assume Euler's formula for all connected and planar graphs with $f - 1$ faces. Consider G with $f > 1$ faces. Since $f \geq 2$, then G is not a tree, since a cycle is needed to delimit a second face. Remove an edge e from a cycle in G. Now $G \setminus e$ has $e - 1$ edges and $f - 1$ faces. Furthermore, $G \setminus e$ is connected (why?) and planar. By the I.H., we have:

$$\begin{aligned} &(e - 1) + 2 = n + (f - 1) \\ \implies &e + 2 = n + f \end{aligned}$$

■

Remark. An edge e on a cycle in G will lie on two faces F_1 and F_2 by the Jordan Curve Theorem, which we did not formally discuss. This is why $G \setminus e$ has $f - 1$ faces.

Lemma 6.3. In a simple, connected, planar graph with $|V(G)| \geq 4$:

$$e \leq 3n - 6$$

Proof. Let p be the number of pairs (x, y) in G where x is an edge on the boundary of a face y. Every edge is on the boundary of at most 2 faces, so $p \leq 2e$. Every face is delimited by at least 3 edges, so $p \geq 3f$. We have $3f \leq 2e$. By Euler's formula, $e + 2 = n + f$. Then $3e + 6 = 3f + 3n \leq 2e + 3n$. So $e \leq 3n - 6$.

■

Lemma 6.4. In a simple, connected, planar, *bipartite* graph with $|V(G)| \geq 4$:

$$e \leq 2n - 4$$

Proof. Continuing the previous proof, we still have $p \leq 2e$, but since G is bipartite now, it cannot contain any odd cycles. So every face is delimited by at least 4 edges, and we have $p \geq 4f$. So $4f \leq 2e$. Using Euler's formula, $4e + 8 = 4f + 4n \leq 2e + 4n$. So $e \leq 2n - 4$.

■

Theorem 6.17 (Kuratowski's Theorem)**.** *G is not planar if and only if $K_5 \leq G$ or $K_{3,3} \leq G$*

Remark. This is an important tool for checking a graph's planarity.

Lemma 6.5. *If G is connected and planar, then $\exists x \in V(G)$ such that $\deg(x) \leq 5$.*

Proof. Assume that $\deg(x) > 6$ for every $x \in V(G)$. Then by the handshaking lemma:

$$\begin{aligned} 2e &= \sum_{x \in V(G)} \deg(x) \\ &\geq \sum_{x \in V(G)} 6 \\ &= 6n \\ \iff e &\geq 3n \end{aligned}$$

By Lemma 6.2:

$$\begin{aligned} e &\leq 3n - 6 \\ &\leq e - 6 \\ \iff 0 &\leq -6 \end{aligned}$$

Contradiction arises.

■

6.8 Graph Colouring

Definition 6.37 (Chromatic number)**.** The chromatic number of G, denoted $\chi(G)$, is the minimum number of colours needed to colour G, such that no two adjacent vertices have the same colour.

Remark. The problem of graph colouring is difficult, it is NP-complete.

Theorem 6.18 (4-Colour Theorem)**.** Every planar graph is 4-colourable ($\chi(G) \leq 4$)).

Remark. The proof of this theorem requires computer assistance - it involves proving several thousand cases. The theorem was first proven in 1976, and it was the first theorem a human could not read in its entirety.

Lemma 6.6. For any complete graph K_n, we have $\chi(K_n) = n$.

Proof. Since every vertex is connected to every other vertex, all vertices need different colours.

■

Theorem 6.19 (6-Colour Theorem)**.** Every planar graph is 6-colourable ($\chi(G) \leq 6$). Note this is weaker than the 4-Colour Theorem, but it is included since it has a comparatively understandable proof.

Proof. We employ induction on the number of vertices $n \in \mathbb{N}$.

Base cases: For $n = 1, 2, \cdots, 6$, any planar graph is trivially 6-colourable.

Inductive step: Assume that any planar graph with $n - 1$ vertices is 6-colourable. Consider an arbitrary planar graph G with n vertices. By Lemma 6.5, $\exists x \in V(G)$ s.t. $\deg(x) \leq 5$. Delete such a vertex x to obtain G' with $(n - 1)$ vertices, which is still planar. By the I.H., G' is 6-colourable. Note there are at most 5 vertices in G that were the original neighbours of x. Colour them in 5 different ways, put x back, and colour x with the 6th distinct colour.

■

Remark. This proof is a template that can be used elsewhere. Consider graphs of type M. Suppose graphs of type M always contain a vertex x where $\deg(x) \leq k - 1$. Suppose that graphs of type M are invariant to vertex deletion (they remain type M). By following an inductive proof similar to the above, we can show that all graphs G of type M are k-colourable, so $\chi(G) \leq k$.
The crucial step is showing $\exists x \in V(G)$ s.t. $\deg(x) \leq k - 1$. One approach is

to show that all graphs G of type M have average degree strictly less than k. For example, all trees have average degree

$$\frac{1}{n}\sum_{x\in V}\deg(x) = \frac{2|E|}{n} = \frac{2(n-1)}{n} < 2$$

So all trees contain a vertex x such that $\deg(x) \leq 2-1$ and it follows they are all 2-colourable.

Greedy Colouring Algorithm

This algorithm follows from the proof of the 6-Colour Theorem. For a planar graph G, it looks like this.

(1) Order vertices as $V_1, V_2, \ldots, V_n$. Let V_n be a vertex with $\deg(x) \leq 5$ in $G_n = G$. Let V_{n-1} be a vertex with $\deg(x) \leq 5$ in $G_{n-1} = G \setminus V_n$. And so on, until we have one vertex left.

(2) Greedily colour the vertices in G in the order $V_1, V_2, \ldots, V_n$.

Every time a vertex V_i is coloured, at most 5 of its neighbours are already coloured. Why is this true? Since we are colouring in this order, it is as if vertices later in the sequence are "deleted", so the current vertex has at most 5 neighbours. Therefore, one of the 6 colours is always available for V_i.

6.9 Platonic Solids

Definition 6.38 (Polyhedron)**.** A polyhedron is a 3D solid bounded by polygonal faces. Two faces meet at each edge and three or more faces at a vertex.

Definition 6.39 (Regular polyhedron)**.** A polyhedron is regular if all faces are identical regular polygons, and the same number of faces meet at each vertex.

Definition 6.40 (Platonic solid)**.** A platonic solid is a convex regular polydron.

Definition 6.41 (Platonic graph)**.** A platonic graph is a projection of a platonic solid onto a plane. Every vertex has the same degree (d-regular), and the number of edges surrounding each face is the same (r-regular).

Theorem 6.20. There are exactly 5 platonic solids.

Proof. For a platonic graph G, let d be the degree of each vertex, and let r be the number of edges delimiting each face.

Since any platonic graph is planar, then by Lemma 6.5, $\exists x \in V(G)$ s.t. $\deg(x) \leq 5$. This implies $d \leq 5$. But also, each vertex must have at least 3 incident edges, since we are working with platonic solids, so $d \geq 3$. Further notice that a polygon must have at least 3 sides, so $r \geq 3$. The dual graph of a planar graph is planar, so this implies $r \leq 5$ by Lemma 6.5 (r corresponds to vertices in the dual graph).

So far, $3 \leq r \leq 5$ and $3 \leq d \leq 5$, corresponding to 9 possible platonic solids. We need to do more.

Let p be the number of pairs (E, F) where F is a face delimited by E. We have $p = rf$. Since every edge is at the boundary of 2 faces, then we also have $2e = p$. Thus $2e = rf$. By the handshaking lemma, $2e = \sum_{x \in V(G)} \deg(x) = nd$. Then:

$$\begin{aligned} rf + nd &= 2e + 2e \\ &= 4(n + f - 2) && \text{(by Euler's formula)} \\ &= 4n + 4f - 8 \\ \iff (r-4)f + (d-4)n &= -8 \\ \iff (r-4)f + (d-4)n &< 0 \end{aligned}$$

By this inequality, $(r = 3) \vee (d = 3)$. Thus, the only possibilities for (r, d) are:

$$(r, d) \in \{(3,3), (4,3), (3,4), (3,5), (5,3)\}$$

So there are exactly 5 platonic graphs, each corresponding to a platonic solid.

■