Contents

1

Graph Structures

1.1 Multigraphs

1.1.1 Multigraphs

Intuitively, we would describe a *graph* as a collection of nodes (or vertices) plus some lines (or edges) joining some nodes together. This can be rigorously defined as follows:

> **Definition 1.1.1 ▸ Multigraph**
>
> A **multigraph** G consists of a non-empty finite set of vertices denoted by $V(G)$ and a finite set of edges denoted by $E(G)$. $|V(G)|$ is known as the **order** of G, denoted by $v(G)$ and $|E(G)|$ is known as the **size** of G, denoted by $e(G)$.
>
> In particular, if $v(G) = m$ and $e(G) = n$, we say that G is an (m, n)-**graph**.
>
> G is said to be **trivial** if $v(G) = 1$ and **non-trivial** otherwise.

Note that in a multigraph, by default there can be a plural number of edges between any two vertices. We would define the notion of *simple graph* as a special multigraph.

> **Definition 1.1.2 ▸ Simple Graph**
>
> A multigraph G is said to be **simple** if there is at most one edge between any two distinct vertices.

> *Remark.* Note that if G is simple and undirected, then
>
> $$E(G) \subseteq \big\{(v_j, v_i) : v_j, v_i \in V(G), j \geq i\big\}.$$

Notice that if (v_i, v_i) is an edge, it connects a vertex to itself. This is known as a *loop*.

In a graph, it is important to know the layout of the vertices and edges. For this purpose, we define the notion of *adjacency* in a multigraph.

Definition 1.1.3 ▶ Adjacency and Neighbourhood

Let $v_i, v_j \in V(G)$, we say that they are **adjacent** if $v_i v_j \in E(G)$. Alternatively, we say that v_i and v_j are **neighbours** to each other. The edge $v_i v_j$ is said to be **incident** with v_i and v_j. Two edges e and f are said to be **adjacent** if there exists some $v \in V(G)$ such that both e and f are incident with v.

The set of all neighbours to some $v_i \in V(G)$ is called the **neighbourhood set** of v_i, denoted by $N_G(v_i)$. In particular, the set $N_G[v_i] := N_G(v_i) \cup \{v_i\}$ is known as the **closed neighbourhood set** of v_i.

Alternatively, one may write $v \sim u$ if v and u are adjacent vertices. We can further add on to the definition by discussing the size of the neighbourhood of a vertex.

Definition 1.1.4 ▶ Degree

The **degree** of v, denoted by $d_G(v)$ is defined as the number of edges incident to v. If $d_G(v)$ is even (repectively, odd), then we say that v is an even (respectively, odd) vertex. If $d_G(v) = 0$, we say that v is **isolated**; if $d_G(v) = 1$, we say that v is an **end** vertex.

In particular, we define

$$\Delta(G) = \max_{v \in G} d_G(v), \qquad \delta(G) = \min_{v \in G} d_G(v).$$

In particular, we can denote the average degree of a graph G by

$$\bar{d}(G) = \frac{\sum_{v \in V(G)} d_G(v)}{v(G)}.$$

Remark. Note that a high maximal degree does not imply a high average degree. A classic counter example is a complete graph plus an isolated vertex.

With the notion of degree established, we can now define a *regular graph*.

Definition 1.1.5 ▶ Regular Graph

A **regular graph** is a graph in which every vertex has the same degree. In particular, if $d_G(v) = k$ for all $v \in V(G)$, G is known as a k-**regular graph**.

Remark. A graph G is regular if and only if $\Delta(G) = \delta(G)$.

Since a graph essentially consists of two sets, it is natural to consider the notion of graph

complementation. We now proceed to introducing the notion of *complement*.

Definition 1.1.6 ▶ Complement

Let G be a graph of order n, the **complement** of G, denoted by $\overline{G}$, is the graph of order n where
$$V\left(\overline{G}\right) = V(G), \qquad E\left(\overline{G}\right) = \{(u,v) : (u,v) \notin E(G)\}.$$

We discuss the following interesting question:

> 15 students went to a party. During the party some of them shook hands with each other. At the end of the party, the number of handshakes made by each student was recorded and it was reported that the sum was 39. Was this possible?

Note that if we represent each student as a vertex, then we can use $V(G) = \{v_1, v_2, \cdots, v_{15}\}$ to construct a graph, in which an edge $v_i v_j \in E(G)$ if and only if students i and j shook hands. As such, $d_G(v_i)$ is the number of persons student i shook hands with.

Note that if we sum up $d_G(v_i)$ for all the vertices, every edge will be counted exactly twice! This means that
$$\sum_{i=1}^{15} d_G(v_i) = 39$$

is impossible since the left-hand side must be even. In fact, by the above reasoning, we see that in any graph, the sum of degrees of its vertices must be even.

Lemma 1.1.7 ▶ Handshaking Lemma

If G is a graph of order n and size m, then

$$\sum_{i=1}^{n} d_G(v_i) = 2m.$$

Remark. It can be easily deduced from the above lemma that in any graph, the number of vertices with odd degrees must be even.

Relating to Definition 1.1.4, we see that the average degree of a graph G can be computed as
$$\bar{d}(G) = \frac{2e(G)}{v(G)}.$$

As an extension of Lemma 1.1.7, the minimum size of any graph is obviously 0, and the maximum size of a simple graph occurs when there is an edge between any two vertices.

> **Definition 1.1.8 ▶ Empty and Complete Graphs**
>
> Let G be a simple graph of order n. G is said to be an **empty graph** or **null graph**, denoted by 0_n, if $e(G) = 0$, and a **complete graph**, denoted by K_n if for all $u, v \in V(G)$, we have $(u, v) \in E(G)$.

1.1.2 Subgraphs

Since a multigraph is just two sets, we can define a "subset" relation between graphs.

> **Definition 1.1.9 ▶ Subgraph**
>
> Let G, H be graphs, then H is a **subgraph** of G if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. In particular, a subgraph H is a **proper subgraph** of G if $V(H) \neq V(G)$ or $E(H) \neq E(G)$.

Note that we a graph can possibly be reproduced by connecting its vertices correctly, so a subgraph containing all vertices of a graph "spans" the original graph.

> **Definition 1.1.10 ▶ Spanning Subgraph**
>
> Let H be a subgraph of G. H is called a **spanning subgraph** of G if $V(H) = V(G)$.

Note that by definition, a spanning subgraph retains all vertices of the original graph. Therefore, a way to quickly generate a subgraph of a given graph G is to keep the vertex set and delete some edges from the edge set. We denote such a graph by $H = G - F$ for some $F \subset E(G)$. Observe that

$$V(H) = V(G), \qquad E(H) = E(G) - F.$$

On the other hand, by deleting some vertices together with edges incident to them, we can produce a subgraph from any given graph.

> **Definition 1.1.11 ▶ Induced Subgraph**
>
> Let H be a subgraph of G. H is called a **induced subgraph** of G if
>
> $$E(H) = \{uv \in E(G) : u, v \in V(H)\}.$$
>
> Let $S \subseteq V(G)$, the subgraph of G induced by S is denoted by $[S]$.

Alternatively, let $F \subseteq E(G)$. Define

$$V' := \bigcup \{\{u, v\} : uv \in F\},$$

then (V', F) is the subgraph of G induced by F, denoted by $G[F]$.

Intuitively, an induced subgraph pf G consists of a selected subset of the vertices of G together with all edges in G connecting any vertices in this subset.

Using the idea of deletion, we can quickly generate an induced subgraph H of G as follows: first, set $V(H) = V(G) - A$ for some $A \subseteq V(G)$, i.e., remove some vertices; then, we will remove all edges from $E(G)$ which are incident to some vertices in A, i.e., set

$$E(H) = \{e \in E(G) : e \text{ is not incident to any } v \in A\}.$$

Then, $H = \big(V(H), E(H)\big)$ is an induced subgraph of G. In fact, we can denote H as

$$H = G - A = [V(G) - A].$$

This leads to the following proposition:

> **Proposition 1.1.12**
>
> *Let G be a graph. If $A \subseteq V(G)$, then $G - A = [V(G) - A]$. Suppose H is a subgraph of G, then H is an induced subgraph of G if and only if $H = G - \big(V(G) - V(H)\big)$.*

We would like to consider the relations between the properties of a graph to those of its subgraphs. In particular, it follows from intuition that if a graph has a high average degree, then we can naturally find a subgraph with a high minimal degree.

This seems trivial from intuition, as we can always keep deleting vertices with the minimal degree from a graph until we cannot find any vertex whose degree is less than $\frac{1}{2}\bar{d}(G)$, but there is more to that — for instance, how do we know if we would not end up deleting all vertices from the original graph? Thus, to prove our claim is essentially asking to prove the correctness of this greedy deletion, which we shall do in the following proposition.

> **Proposition 1.1.13 ▶ Subgraphs with High Minimum Degree**
>
> *Every non-empty graph G has a subgraph H such that $\delta(H) \geq \frac{1}{2}\bar{d}(G)$.*
>
> *Proof.* Define $H_0 = G$ and $H_{i+1} = H_i - v_i$ where v_i is a vertex in H_i whose degree is the smallest. Note that we can always find this v_i for any H_i with $V(H_i) \neq \varnothing$

because $V(H_i)$ is a finite well-ordered set. We will repeatedly perform the deletion until we obtain some H_k with $d_{H_k}(v_k) \geq \frac{1}{2}\bar{d}(G)$. We shall proceed to proving that this algorithm always terminates with $V(H_k) \neq \emptyset$.

Suppose on contrary that $V(H_k) = \emptyset$, then we have performed the deletion for $v(G)$ times. Suppose we have deleted N edges in total, then clearly,

$$N < v(G)\frac{1}{2}\bar{d}(G) = v(G)\frac{e(G)}{v(G)} = e(G),$$

which means $e(H_k) = e(G) - N > 0$. However, this is impossible since $v(H_k) = 0$, which is a contradiction. $\qquad\square$

1.1.3   Graph Isomorphism

Given two graphs G and H, are they the same graph? This seemingly innocent question proves to be extremely hard to answer. Two graphs can look drastically different but be structurally identical in reality. For example, we could shift around the vertices of a graph without changing any edge to alter the shape of the graph dramatically. Therefore, to compare the structures of graphs, we require some rigorous definition.

Definition 1.1.14 ▸ Graph Isomorphism

Two graphs G and H are said to be **isomorphic**, denoted by $G \cong H$, if there exists a bijection $f : V(G) \to V(H)$ such that

$$uv \in E(G) \quad \text{if and only if } f(u)f(v) \in E(H).$$

Remark. One may check that graph isomorphism is obviously an equivalence relation.

Such a bijection f is said to **preserve adjacency**, i.e., if u and v are neighbours in G, then their images are also neighbours in H. In particular, it is possible to map a graph to its own complement via an isomorphism.

Definition 1.1.15 ▸ Self-Complementary Graph

A graph G is said to be **self-complementary** if $G \cong \overline{G}$.

It turns out that a self-complementary graph satisfies some special properties.

> **Proposition 1.1.16 ▶ Order of Self-Complementray Graphs**
>
> *If G is a self-complementary graph of order n, then $n = 4k$ or $n = 4k+1$ for some $k \in \mathbb{Z}^+$.*
>
> ---
>
> *Proof.* Since G is self-complementary, $G \cong \overline{G}$ and so $e(G) = e\left(\overline{G}\right)$. Note that
>
> $$e(G) + e\left(\overline{G}\right) = \binom{n}{2} = \frac{n(n-1)}{2},$$
>
> so we have $e(G) = \frac{n(n-1)}{4}$. Since $e(G) \in \mathbb{Z}^+$, either n or $n-1$ is divisible by 4, so $n = 4k$ or $n = 4k + 1$ for some $k \in \mathbb{Z}^+$. $\qquad\square$

It is very hard to determine whether two specific graphs are isomorphic, but there are some considerations in the general case. For example, some trivial conclusions include:

- Two graphs with different orders cannot be isomorphic.

- Two graphs with different sizes cannot be isomorphic.

- Two graphs with different numbers of components cannot be isomorphic.

- If the numbers of vertices with degree k are different in two graphs, they cannot be isomorphic.

Notice that by Definition 1.1.14, we can define a function g such that $uv \notin E(G)$ if and only if $g(u)g(v) \notin E(H)$ and relate this function to the complement graphs of G and H. Here we introduce a way to determine isomorphism by considering complement graphs, the proof of which is left to the reader as an exercise.

> **Theorem 1.1.17 ▶ Complementation Preserves Isomorphism**
>
> *Let G and H be two graphs of the same order. $G \cong H$ if and only if $\overline{G} \cong \overline{H}$.*

Now let us think the reverse: if it is not easy to prove isomorphism, can we find a way to quickly determine that two graphs are not isomorphic? Here we present the necessary conditions for isomorphism:

> **Theorem 1.1.18 ▶ Necessary Conditions for Isomorphism**
>
> *If $G \cong H$, then*
> 1. *G and H must have the same order and size.*
> 2. *$\delta(G) = \delta(H)$ and $\Delta(G) = \Delta(H)$.*
> 3. *The number of vertices with degree i in G and H is the same for all $i \in \mathbb{N}$.*

The first two conditions are easy to observe. For the third condition, we introduce a tool known as *degree sequences*.

> **Definition 1.1.19 ▸ Degree Sequence**
>
> Let G be a graph of order n. If we label its vertices by $v_1, v_2, \cdots, v_n$ such that
>
> $$d(v_1) \geq d(v_2) \geq \cdots \geq d(v_n),$$
>
> then the non-increasing sequence $\big(d(v_n)\big)$ is known as the **degree sequence** of G.

Now we consider the following question:

> Let (d_n) be a non-increasing sequence, is there some graph whose degree sequence is (d_n)?

We will make use of the following definition:

> **Definition 1.1.20 ▸ Graphic Sequence**
>
> Let (d_n) be a sequence of non-negative integers at most $n-1$. (d_n) is said to be **graphic** if there exists a graph G whose degree sequence is (d_n).

To determine whether a sequence is graphic by eye power is difficult. Fortunately, we have the following recursive algorithm:

> **Theorem 1.1.21 ▸ Havel-Hakimi Algorithm**
>
> *Let (d_n) be a non-increasing sequence of non-negative integers at most $n - 1$. Define*
>
> $$d_m^* = \begin{cases} d_{m+1} - 1 & \text{if } 1 \leq m \leq d_1 \\ d_{m+1} & \text{if } d_1 + 1 \leq m \leq n - 1 \end{cases}.$$
>
> *(d_n) is graphic if and only if (d_m^*) is graphic.*
>
> *Proof.* Suppose (d_n) is graphic, we consider the following lemma:

> **Lemma 1.1.22**
>
> *For any graphic sequence (d_n), there is some graph G of order n with $d_G(v_i) = d_i$ for $i = 1, 2, \cdots, n$ such that v_1 is adjacent to v_j for $j = 2, 3, \cdots, d_1 + 1$.*
>
> ---
>
> *Proof.* Suppose on contrary there is no such a graph G, then for any graph G with degree sequence (d_n), v_1 is adjacent to at most $d_1 - 1$ vertices in
>
> $$A := \{v_j : j = 2, 3, \cdots, d_1 + 1\}.$$
>
> Thus, there exists some $v_j \in A$ such that $v_1 v_j \notin E(G)$. However, $d_G(v_1) = d_1$, so there exists some $v_k \in V(G) - A$ with $v_k \neq v_1$ such that $v_1 v_k \in E(G)$.
>
> Notice that since $v_k \notin A \cup \{v_1\}$, $k > j$ and so $d_G(v_k) = d_k \leq d_j = d_G(v_j)$. Therefore, v_j has at least as many neighbours as v_k, which means there must exists some v_r with $v_j v_r \in E(G)$ such that $v_k v_r \notin E(G)$.
>
> Define a graph G' by
>
> $$V(G') = V(G), \qquad E(G') = E(G) - v_j v_r - v_1 v_k + v_j v_k + v_1 v_j.$$
>
> Note that $d_{G'}(v_i) = d_G(v_i)$ for all $i = 1, 2, \cdots, n$, so (d_n) is also a degree sequence for G'. However, now v_1 in G' is adjacent to all vertices in A, which is a contradiction. $\qquad\square$

By Lemma 1.1.22, we can choose some graph G whose degree sequence is (d_n) such that v_1 is adjacent to d_1 vertices $v_2, v_3, \cdots, v_{d_1+1}$. Let $H = G - v_1$, then for any $u_i \in H$,

$$d_H(u_i) = \begin{cases} d_G(v_{i+1}) - 1 & \text{if } 1 \leq i \leq d_1 \\ d_G(v_{i+1}) & \text{if } d_1 + 1 \leq i \leq n - 1 \end{cases}.$$

Therefore, (d_m^*) is graphic.

Conversely, suppose (d_m^*) is graphic, then there is some graph H of order $n - 1$ such that $d_H(u_i) = d_i^*$ for $i = 1, 2, \cdots, d_1$. Construct a graph $G = H + v_1$ such

that $v_i = u_{i-1}$ for $i = 2, 3, \cdots, n$ and v_1 is adjacent to $v_2, v_3, \cdots, v_{d_1+1}$. Therefore,

$$
d_G(v_i) = \begin{cases} d_1 & \text{if } i = 1 \\ d_{i-1}^* + 1 & \text{if } 2 \leq i \leq d_1 + 1 \\ d_i^* & \text{if } d_1 + 2 \leq i \leq n \end{cases} .
$$

Note that $d_G(v_i) = d_i$, so (d_n) is graphic.   □

By Theorem 1.1.21, we can reduce a sequence recursively until we reach some obvious case. This obvious case is graphic if and only if the original sequence is graphic. We can also build the graph corresponding to the original sequence from a simple graphic sequence.

Alternatively, the following is another algorithmic way to determine whether a sequence is graphic.

Theorem 1.1.23 ▶ Erdos-Gallai Algorithm

A non-increasing sequence $(d_1, d_2, \cdots, d_n)$ is graphic if and only if $\sum_{i=1}^{n} d_i$ is even and for all $k = 1, 2, \cdots, n$,

$$
\sum_{i=1}^{k} d_i \leq k(k-1) + \sum_{i=k+1}^{n} \min\{d_i, k\}.
$$

We will not give the full proof but only briefly justify the "only if" direction. Let G be any graph. Define $d_i = d_G(v_i)$ where $v_i \in V(G)$ for $i = 1, 2, \cdots, n$, then $(d_1, d_2, \cdots, d_n)$ is graphic. Note that $\sum_{i=1}^{n} d_i$ is obviously even by Lemma 1.1.7.

Take some $k \in \mathbb{Z}^+$ with $k \leq n$ and let $S = \{v_1, v_2, \cdots, v_k\}$, then by Lemma 1.1.7 again,

$$
\sum_{i=1}^{k} d_i = 2e(G[S]) + e_G(S, V(G) - S)
$$

$$
\leq k(k-1) + e_G(S, V(G) - S),
$$

where $e_G(S, V(G) - S)$ is the number of edges in G which connect a vertex in S to a vertex in $V(G) - S$. Notice that

$$
e_G(S, V(G) - S) = \sum_{i=k+1}^{n} |\{u \in S : uv_i \in E(G)\}|
$$

$$
\sum_{i=k+1}^{n} \leq \min\{k, d_i\},
$$

and so it follows that

$$\sum_{i=1}^{k} d_i \le k(k-1) + \sum_{i=k+1}^{n} \min\{d_i, k\}.$$

> *Remark.* Both Theorems 1.1.21 and 1.1.23 are $\mathcal{O}(n^2)$ algorithms.

1.2 Graph Traversal

1.2.1 Paths and Cycles

We are often interested in the "connectedness" in a graph. In particular, we would wish to study how to traverse a graph from some vertices to others.

Definition 1.2.1 ▶ Walk, Trail, Path

Let G be a graph and $x, y \in V(G)$, then an *x-y* **walk** is an alternating sequence

$$W := v_0 e_1 v_1 e_2 v_2 \cdots v_{n-1} e_n v_n$$

where $v_i \in V(G)$, $e_i = v_{i-1} v_i \in E(G)$, $x = v_0$ and $y = v_n$. In particular, W is **open** if $v_0 \ne v_n$ and **closed** otherwise. The **length** of a walk is defined as the number of edges in it.

If $e_i \ne e_j$ whenever $i \ne j$, then W is called a **trail**. A closed trail is called a **circuit**. If $v_i \ne v_j$ whenever $i \ne j$, then W is called a **path**. A closed path is called a **cycle**.

> *Remark.* All paths are trails and all trails are walks, but the converses are not true.

In more informal terms, a trail is a walk with no repeated edge and a path is a trail with no repeated vertex.

Note that a path is essentially a graph. We denote a path of order n by P_n and a cycle of order n by C_n. It is easy to see that all cycles are circuits and all circuits are all closed walks, but the converses are not true. However, if we have a circuit, we can guarantee that this circuit contains at least one cycle.

Intuitively, all walks can be reduced to a path by removing duplicated edges and vertices repeatedly.

> **Proposition 1.2.2 ▶ Any Walk Contains A Path**
>
> *If a graph G contains a u-v walk of length k, then G contains a u-v path of length at most k.*
>
> ---
>
> *Proof.* Let S be the set of all u-v walks in G, Note that $S \neq \varnothing$ since there is a u-v walk of length k. Let P be a walk of the shortest length. We claim that P must be a path.
>
> Suppose P is not a path, then there exists some vertices $w_1 = w_2$ in P. Suppose Q is obtained by removing the w_1-w_2 walk from P, then Q is a walk. However, Q is shorter than P which is not possible. Therefore, P must be a path.
>
> Note that the length of P is at most k, so G contains a u-v path of length at most k. □

We also have the following relevant definitions with regard to cycles:

> **Definition 1.2.3 ▶ Girth and Circumference**
>
> Let G be a graph. The **girth** of G is the size of the shortest non-trivial cycle in G and the **circumference** of G is the size of the longest non-trivial cycle in G.

If G is acyclic, we define the girth of G to be ∞ and the circumference to be 0. In some special graphs, we realise that we can traverse through every vertex and return to the starting point in a single traversal. Such graphs are said to be *Hamiltonian*.

> **Definition 1.2.4 ▶ Hamiltonian Graph**
>
> A graph G is **Hamiltonian** if it contains a cycle of size $v(G)$.

1.2.2 Connected Components

An important question we are interested in with graphs is the reachability of a vertex, i.e., given two vertices u and v, we want to know whether we can reach one vertex from another. Intuitively, we can traverse between two vertices if there is a path between them. From here we define the notion of a *connected graph*.

> **Definition 1.2.5 ▶ Connected Graph**
>
> Let G be a graph. Two vertices u and v are **connected** if there is a path between them. G is said to be **connected** if for any $u, v \in V(G)$, there exists a path from u to v. For any $u \in V(G)$, we denote the set of all vertices connected to u (inclusive of u) as $c(u)$.

Intuitively, we of course always can find a connected subgraph of a connected graph, but we can in fact produce a much stronger result.

> **Proposition 1.2.6 ▶ Labelling of Vertices of Connected Graphs**
>
> *For every connected graph G, $V(G)$ can be labelled as $\{v_1, v_2, \cdots, v_n\}$ such that the graph $G_i := [\{v_1, v_2, \cdots, v_i\}]$ is connected for every $i = 1, 2, \cdots, n$.*
>
> *Proof.* For $i = 1$, we can pick any vertex in $V(G)$ as v_1 and G_1 is always trivially connected.
>
> Suppose we have fixed up to k vertices for some $k = 1, 2, \cdots, n-1$, then $G_1, G_2, \cdots, G_k$ are connected subgraphs of G. Note that $v(G) = n > k$, so there exists some $v \in V(G) - V(G_k)$ such that v and v_1 are connected by a path P. Note that there exists some $u \in V(P)$ such that $u \sim v$ for some $v \in V(G_k)$, so taking $v_{k+1} = u$ guarantees that G_{k+1} is connected. $\qquad\square$

Intuitively, if a graph has many edges, we would believe that it is easy to find a long path in it. Note that the notion of "having many edges" can be related to the minimal degree of a graph. More formally, the following result is true:

> **Proposition 1.2.7 ▶ Finding Paths and Cycles in Graphs**
>
> *Every graph G with $\delta(G) \geq 2$ contains a path of length $\delta(G)$ and a cycle of length at least $\delta(G) + 1$.*
>
> *Proof.* Let $P = x_0 x_1 \cdots x_k$ be the longest path in G. Note that it suffices to prove that $k \geq \delta(G)$. Suppose that there exists some $v \in N(x_k)$ such that $v \notin V(P)$, then we can find a longer path which is not possible. Therefore, $N(x_k) \subseteq V(P)$. Note that $|N(x_k)| \geq \delta(G)$, so $k \geq \delta(G)$ and so G contains a $P_{\delta(G)} \subseteq P$.
>
> Let $i \in \mathbb{N}$ be such that $x_j x_k \notin E(G)$ for all $0 \leq j < i$. Take $P' = x_i x_{i+1} \cdots x_k$, then $e(P') \geq \delta(G)$ by the previous argument. However, $x_i \sim x_k$, so $P' + x_i x_k$ is a cycle of length at least $\delta(G) + 1$ in G. $\qquad\square$

A classic problem regarding graph traversal in connected graphs is the Könisburg Bridge Problem:

> Let A, B, C, D be 4 distinct vertices such that there are 2 edges between A and B, 2 edges between A and C, and 1 edge between A and D, B and D, C and D. Can all the edges be visited exactly once in a single traversal of the graph produced?

This can be generalised as: given a connected multigraph G, does there exist a circuit in G

which contains all the edges of G? Such multigraphs are said to be *Eulerian*.

> **Definition 1.2.8 ▶ Eulerian Circuit**
>
> Let G be a connected multigraph, an **Eulerian circuit** in G is a circuit which uses all edges of G exactly once.

Intuitively, if a graph is Eulerian, then following any Eulerian circuit will visit each of its vertices for an even number of times as there are equally many incoming edges as outgoing edges for every vertex in a circuit. This implies that an Eulerian graph will necessarily force all of its vertices to have even degrees. However, it turns out that actually this is also a sufficient condition for the existence of Eulerian circuits!

> **Theorem 1.2.9 ▶ Necessary and Sufficient Condition for Eulerian Graphs**
>
> *A connected multigraph G is Eulerian if and only if every vertex in G has an even degree.*
>
> *Proof.* The necessity is trivial, so we will only prove for sufficiency.
>
> Consider the case where $e(G) = 0$, then the only connected multigraph is the trivial graph consisting of a singleton vertex, which is Eulerian.
>
> Suppose that for all $k < m$ where $m \in \mathbb{Z}^+$, a connected multigraph H with $e(H) = k$ is Eulerian if every vertex in H has an even degree. Let G be a connected multigraph with $e(G) = m > 0$ such that every vertex in G has an even degree, then $\delta(G) \geq 2$. By Proposition 1.2.7, there exists a cycle $S \subseteq G$ where $e(S) \geq 3$.
>
> Let G' be the multigraph such that $V(G') = V(G)$ and $E(G') = E(G) - E(S)$, then clearly, $e(G') < m$ and every vertex in G' has an even degree. Therefore, each connected component of G' is Eulerian.
>
> Let C_1 be a connected component of G' and let $v_1 \in V(C_1) \cap V(S)$. Note that there exists an Eulerian circuit in C_1 which starts and ends with v_1. We can traverse along this circuit and then traverse in the clockwise orientation along S until we reach another vertex in S which is in some connected component of G' and repeat the process. Note that $v_1 \in V(S)$, so we will eventually traverse back to v_1, which gives an Eulerian circuit in G.   □

Intuitively, even if a graph G is not Eulerian, it is possible to traverse all edges in G with finitely many trails. This is known as a decomposition of G into trails.

Proposition 1.2.10 ▶ Minimal Number of Trails to Decompose A Multigraph

Let G be a graph with $2k$ vertices having odd degrees for some $k \in \mathbb{N}$, then the minimal number of trails needed to decompose G is $\max\{k, 1\}$.

Proof. The case where $k = 0$ is immediate by Theorem 1.2.9. For $k \geq 1$, suppose on contrary that G can be decomposed with $k - 1$ trails labelled as $T_1, T_2, \cdots, T_{k-1}$. Since a trail contains at most 2 vertices with odd degrees, then

$$\left\| \left\{ v \in \bigcup_{i=1}^{k-1} T_i : d_G(v) \text{ is odd} \right\} \right\| \leq 2(k - 1) < 2k,$$

which is a contradiction. Therefore, we need at least k trails to decompose G.

Let $v_1, v_2, \cdots, v_{2k}$ be the vertices in G with odd degrees. Let G' be the graph produced by adding the edges $v_1 v_2, v_3 v_4, \cdots, v_{2k-1} v_{2k}$ to G, then the vertices in G' all have even degrees. Therefore, there exists an Eulerian circuit in G' by Theorem 1.2.9.

Note that for all $i = 1, 2, \cdots, k - 1$, $v_{2i-1} v_{2i}$ and $v_{2i+1} v_{2i+2}$ are disjoint, so there exists a v_{2i}-v_{2i+1} trail. Note also that there exists a v_{2k}-v_1 trail. The union of these trails is exactly $G' - v_1 v_2 - v_3 v_4 - \cdots - v_{2k-1} v_{2k} = G$.   □

Note that connectedness is an equivalence relation. Let P be a partition formed by the equivalence classes of G under the connectedness relation, then for any $u, v \in V(G)$, u and v are in the same equivalence class if and only if u and v are connected. Any $c(u) \in P$ induces a subgraph $[c(u)]$.

Definition 1.2.11 ▶ Connected Component

Let G be a graph and R be a relation such that for any $u, v \in V(G)$, uRv if and only if u and v are connected. Let

$$C := V(G)/R$$

be the quotient set, then any $c(u) \in C$ is known as a (connected) **component** of G.

The number of components of G is denoted by $\omega(G)$.

Alternatively, a connected component is a maximally connected subgraph, that is, a connected component of G is a subgraph $H \subseteq G$ such that for all $e \in E(G)$, $H + e$ is not connected and for all $v \in V(G) - V(H)$, $[V(H) \cup \{v\}]$ is not connected.

It can be easily seen that G is connected if and only if $\omega(G) = 1$.

We will now establish a relationship between connectedness and complementation.

> **Theorem 1.2.12 ▸ Connectedness of Complement**
>
> *If G is disconnected, then $\overline{G}$ is connected.*
>
> *Proof.* Let $u, v \in V\left(\overline{G}\right)$ be two arbitrary vertices. If $uv \notin E(G)$, then $uv \in \left(\overline{G}\right)$ and so u and v are connected in $\overline{G}$.
>
> If $uv \in E(G)$, since G is disconnected, there exists some $w \in V(G)$ such that $uw, wv \notin E(G)$. Therefore, $uw, wv \in E\left(\overline{G}\right)$. This means that there is a u-v path in $\overline{G}$ and so u and v are connected in $\overline{G}$.
>
> Therefore, $\overline{G}$ is connected. □

Intuitively, we can transform a connected graph into a disconnected one by deleting some vertices or edges. It can be easily observed that in certain graphs, deleting one particular vertex or edge will immediately disconnect the graph.

> **Definition 1.2.13 ▸ Cut-Vertex, Bridge**
>
> Let G be a non-trivial graph. $v \in V(G)$ is called a **cut-vertex** of G if $\omega(G - v) > \omega(G)$. $e \in E(G)$ is called a **bridge** of G if $\omega(G - e) > \omega(G)$.

Note that a graph does not have to be connected to have cut vertices and bridges. Essentially, a cut-vertex (or bridge) is just a vertex (or edge) which disconnects a component upon removal.

Now the next question is how do we identify a cut-vertex or a bridge in a graph? Intuitively, a cut-vertex divides a graph into two portions such that traversal between the two portions has to pass through it.

> **Theorem 1.2.14 ▸ Cut-Vertex Characterisation**
>
> *Let G be a graph. $v \in V(G)$ is a cut-vertex if and only if there exists $a, b \in V(G)$ such that v is in every path between a and b.*
>
> *Proof.* Let $v \in V(G)$ be a cut-vertex and consider $H = G - v$, then we can find two non-empty disjoint connected components of H, say H_1 and H_2. Now, take $a \in V(H_1)$ and $b \in V(H_2)$, then they are disconnected. However, a and b are connected in G, which means v is in every a-b path.

We will prove the converse by contrapositive. Suppose v is not a cut-vertex of G, then $G - v$ is connected. Therefore, for any $u, w \in V(G)$, there is a path between them which does not contain v. $\qquad \square$

A similar argument can be established for bridges. Furthermore, since we have to pass through the bridge when traversing between the two portions, it can be easily seen that we have to reuse the bridge in order to traverse back.

> **Theorem 1.2.15 ▶ Bridge Characterisation**
>
> *Let G be a graph. $e \in E(G)$ is a bridge if and only if e is not contained by any cycle in G.*
>
> *Proof.* We will prove the forward direction by contrapositive. Suppose $e = xy$ is contained in some cycle in G, then there is some x-y path in $G - e$. Suppose $a, b \in V(G)$ are connected in G via a path containing e, this implies that there is some a-b path in $G-e$ and so $G-e$ is connected, which means that e is not a bridge.
>
> We will prove the converse also by contrapositive. Suppose $e = xy$ is not a bridge, the $G - e$ is connected and so there is some x-y path in $G - e$. Let this path be P, then $P \cup \{e\}$ is a cycle in G. $\qquad \square$

Combining the two characterisations, we can relate cut-vertices to bridges in the following proposition, the proof of which is left to the reader as an exercise:

> **Proposition 1.2.16**
>
> *Let G be a graph. If $uv \in E(G)$ is a bridge and u is not an end vertex, then u is a cut-vertex.*

A direct consequence of this is the following corollary:

> **Corollary 1.2.17**
>
> *If G is a graph with order at least 3 and G contains a bridge, then G contains a cut-vertex.*

1.2.3   Graphs as Metric Spaces

Given a graph G with u, v being two connected vertices, we are interested in the notion of distance between them. Intuitively, if there are multiple paths between u and v, we would take the length of the shortest one to represent their distance.

> ### Definition 1.2.18 ▶ Distance, Eccentricity, Diameter
>
> Let G be a connected graph and let $u, v \in V(G)$. The **distance** between u and v, de-noted by $d(u, v)$, is the length of the shortest path between u and v. The **eccentricity** of u is defined to be
>
> $$e(u) = \max_{v \in V(G)} \{d(u, v)\}.$$
>
> The **diameter** of G is defined by
>
> $$\text{diam}(G) = \max_{u \in V(G)} \{e(u)\}.$$
>
> The **radius** of G is defined by
>
> $$\text{rad}(G) = \min_{u \in V(G)} \{e(u)\}.$$
>
> A vertex v is called a **central** vertex if
>
> $$e(v) = \text{rad}(v).$$
>
> The subgraph induced by the set of central vertices of G is known as the **centre** of G.

This also justifies the use of words "diameter" and "radius" in a circle. We can view a circle as a graph consisting of a centre and infinitely many vertices in the circumference, with edges of length r connecting the centre to the circumference. Indeed, in this definition, the farthest vertices will be any two on the circumference with a distance of $2r$, and the nearest vertices will be the centre and any vertex on the circumference with a distance of r.

Intuitively, having a large diameter means that the graph is sparse, i.e., one has to take a very long path to traverse between two vertices. Now let us consider a connected but sparse graph, the above intuition means that we may expect to find that the cycles in this graph are very large.

> ### Proposition 1.2.19 ▶ Graphs with High Girth Have High Diameter
>
> *Let G be a graph with girth $g(G)$. If G contains a cycle, then $g(G) \leq 2\text{diam}(G) + 1$.*
>
> *Proof.* Suppose on contrary that $g(G) \geq 2\text{diam}(G) + 2$. Let $C \subseteq G$ be the shortest cycle in G, then we can take $x, y \in V(C)$ such that $d_C(x, y)$ is the greatest. Clearly, we have $d_C(x, y) \geq \text{diam}(G) + 1$.
>
> By Definition 1.2.18, there exists some path P between x and y with $e(P) \leq \text{diam}(G)$. Let $P' \subseteq C$ be the x-y path in C, then $e(P) < e(P')$. This means that P and P' are

distinct paths and so there is some $q \in V(P') - V(P)$.

Since P and P' are both x-y walks, $|P \cap P'| \geq 2$. Therefore, there are $p_1, p_2 \in P \cap P'$ which are connected to q, which implies that there is a p_1-p_2 path in P'. Since p_1 and p_2 are connected in P, we have found 2 different p_1-p_2 paths and so there is some cycle $Q \subseteq P \cup P'$. Therefore,

$$
\begin{aligned}
e(Q) &\leq e(P \cup P') \\
&\leq e(P) + e(P') \\
&\leq \mathrm{diam}(G) + d_C(x, y) \\
&\leq 2d_C(x, y) - 1 \\
&= 2\left\lfloor \frac{e(C)}{2} \right\rfloor - 1 \\
&< e(C),
\end{aligned}
$$

which is a contradiction.   □

For readers with knowledge in real analysis or topology, it is easy to see that an undirected unweighted connected graph is a metric space with distance between vertices as its metric.

Theorem 1.2.20 ▶ Connected Graph as A Metric Space

If G is connected, then $(V(G), d_G)$ is a metric space.

Proof. Clearly, for all $x, y \in V(G)$, $d_G(x, y) \geq 0 \in V(G)$ with $d_G(x, y) = 0$ if and only if $x = y$ and $d_G(x, y) = d_G(y, x)$.
Let U, V be the shortest u-w path and shortest w-v path respectively, then $U + V$ is a u-v walk with length $d(u, w) + d(w, v)$. Since the shortest u-v path has length $d(u, v)$, we have

$$d(u, v) \leq d(u, w) + d(w, v).$$

Therefore, $(V(G), d_G)$ is a metric space.   □

An application of Theorem 1.2.20 allows us to establish the following:

Theorem 1.2.21 ▶ Boundedness of Diameter

If G is connected, then

$$\mathrm{rad}(G) \leq \mathrm{diam}(G) \leq 2\mathrm{rad}(G).$$

Proof. $\mathrm{rad}(G) \leq \mathrm{diam}(G)$ is immediate from Definition 1.2.18. Take $u, v \in V(G)$ with $d(u, v) = \mathrm{diam}(G)$ and take $w \in V(G)$ such that $e(w) = \mathrm{rad}(G)$. Notice that this implies

$$d(u, w) \leq e(w) = \mathrm{rad}(G)$$
$$d(w, v) \leq e(w) = \mathrm{rad}(G).$$

By Theorem **??**,

$$\mathrm{diam}(G) = d(u, v) \leq d(u, w) + d(w, v) = 2\mathrm{rad}(G).$$

$\square$

1.3   Graphs and Matrices

So far we have been using hand-drawn illustrations to aid us in graph visualisation. However, for large graphs with many vertices and edges, drawing out the graph structures becomes nearly impossible. Therefore, we need to use some algebraic tools to help us represent a large graph.

Definition 1.3.1 ▶ Adjacency Matrix

Let G be a multigraph with $v(G) = n$, then the **adjacency matrix** denoted by $\boldsymbol{A}(G)$ is an $n \times n$ matrix such that a_{ij} is the number of edges incident to both $v_i, v_j \in V(G)$.

Remark. For an undirected multigraph, $\boldsymbol{A}$ is always symmetric.

Note that an adjacency matrix does not tell us which edges are incident to a particular vertex, so we may consider the following alternative:

Definition 1.3.2 ▶ Incidence Matrix

Let G be a multigraph with $v(G) = n$ and $e(G) = m$, then the **incidence matrix** denoted by $\boldsymbol{M}(G)$ is an $n \times m$ matrix such that

$$m_{ij} = \begin{cases} 1, & \text{if } e_j \text{ is incident with } v_i \\ 0, & \text{otherwise} \end{cases}.$$

> *Remark.* For an undirected multigraph, consider an arbitrary k-th column of M. Note that any column will have exactly 2 non-zero entries. Suppose $m_{ik} = m_{jk} = 1$, it means that $e_k = v_i v_j$.

For readers who have learnt about basic data structures in computer science, an incidence matrix is just an adjacency list with each list of neighbours converted to a (non-compact) boolean array.

Note that the sum of the i-th row for both A and M is the degree of vertex i.

Since A and M carries information about adjacencies between vertices, we can use them to draw some conclusions regarding graph traversal.

Proposition 1.3.3 ▶ Counting the Number of Walks between Two Vertices

Let G be a simple graph of order n with adjacency matrix A, then the (i, j) entry of A^k is the number of v_i-v_j walks of length k.

Proof. The case where $k = 1$ is trivially true since there is a v_i-v_j walk of length 1 if and only if v_i and v_j are neighbours.

Let the (i, j) entry of A^m be m_{ij}. Suppose m_{ij} is the number of v_i-v_j walks of length m. Then the (i, j) entry of A^{m+1} is given by

$$\sum_{r=1}^{n} m_{ir} a_{rj}.$$

Notice that m_{ir} is the number of v_i-v_r walks of length m and a_{rj} is the number of v_r-v_j walks of length 1, so by Theorem ??, $m_{ir} a_{rj}$ is the number of v_i-v_j walks of length n that passes through v_r. Therefore, summing up $m_{ir} a_{rj}$ for $r = 1, 2, \cdots, n$ gives the total number of v_i-v_j walks of length $n + 1$. □

Additionally, we have the following corollary:

Corollary 1.3.4 ▶ Connectedness Test

Let G be a graph with $v(G) \geq 2$ and adjacency matrix A. Let

$$B = \sum_{i=1}^{n-1} A^i,$$

---

then G is connected if and only if $b_{ij} \neq 0$ for all $i \neq j$.

As it turns out, the incidence and adjacency matrices are in fact related by a simple equation.

Proposition 1.3.5

Let $\boldsymbol{M}$ and $\boldsymbol{A}$ be the incidence matrix and adjacency matrix of a simple graph G respectively, then

$$
\boldsymbol{M}\boldsymbol{M}^{\mathrm{T}} = \begin{bmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ 0 & 0 & d_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_n \end{bmatrix} + \boldsymbol{A}.
$$

2

Trees and Bipartite Graphs

2.1 Tree

In computer science and data science, it is often useful to use a tree diagram. Colloquially, we view a tree as a graph consisting of nodes, each of which has exactly one parent node and some children nodes. Here, we shall introduce the rigorous mathematical definition for a tree.

Definition 2.1.1 ▶ Tree and Forest

A **forest** is an acyclic graph. A **tree** is a connected forest.

In other words, a tree is an **connected acyclic graph**. We also have the following alternative definition for a tree:

Theorem 2.1.2 ▶ Equivalent Definitions of a Tree

The followings are equivalent:
1. *T is a tree.*
2. *For any $u, v \in V(T)$, there exists a unique u-v path in T.*
3. *T is minimally connected, i.e., T is connected but $T - e$ is disconnected for any $e \in E(T)$.*
4. *T is maximally acyclic, i.e., T is acyclic but $T + xy$ for any $x, y \in V(T)$ with $xy \notin E(T)$ is cyclic.*

Proof. Let T be a tree. For any $u, v \in V(T)$, there exists a u-v path $P \subseteq T$. Suppose $P' \subseteq T$ is also a u-v path, then there exists a non-trivial closed walk between u and v which contains a cycle, leading to a contradiction.

Suppose T is such that any $u, v \in V(T)$ are connected by a unique path. Suppose on contrary $T - e$ is connected for some $e = xy \in E(T)$, then there exists some x-y path in T which does not contain xy. However, xy is a different x-y path, which is a contradiction.

Suppose T is minimally connected. Note that if T contains some cycle C, then

---

for any $e \in E(C)$, $T - e$ is connected. So T being minimally connected implies that T is acyclic. Take any $x, y \in V(T)$ with $xy \notin E(T)$, then there exists some x-y path $P \subseteq T$ which does not contain xy. Therefore, $P + xy$ is a cycle, which implies that T is maximally acyclic.

Suppose T is maximally acyclic. Suppose T is not connected, then there exists $u, v \in V(T)$ which are disconnected. Note that $T + uv$ is acyclic because uv is the unique u-v path. Therefore, T is not maximally acyclic. The contrapositive shows that T being maximally acyclic implies that T is a tree. $\qquad\square$

We first consider the following proposition:

Proposition 2.1.3 ▸ Degrees of Vertices in a Tree

Let T be a tree with $\Delta(T) = k$. Let n_i be the number of vertices in T of degree i, then

$$n_1 = 2 + \sum_{i=1}^{k-2} in_{i+2}.$$

Proof. Let the order of T be n, then $n = \sum_{i=1}^{k} n_i$. By Lemma 1.1.7, we have

$$\sum_{v \in V(T)} d(v) = 2(n-1).$$

However, since there are n_i vertices with degree i, we have

$$\sum_{v \in V(T)} d(v) = \sum_{i=1}^{k} in_i.$$

Therefore,

$$2\sum_{i=1}^{k} n_i - 2 = \sum_{i=1}^{k} in_i$$

$$n_1 = \sum_{i=3}^{k} in_i$$

$$n_1 = \sum_{i=1}^{k-2} in_{i+2}.$$

$\square$

The following corollary is an immediate consequence of Proposition 2.1.3:

> ### Corollary 2.1.4 ▶ Number of End Vertices in a Tree
>
> *A tree of order at least 2 contains at least 2 end vertices.*

In particular, a tree with exactly 2 end vertices is a doubly linked list. We will now give a name to these vertices in a tree whose degree is 1.

> ### Definition 2.1.5 ▶ Leaf
>
> Let G be a tree. A **leaf** of G is a vertex $v \in V(G)$ with $d_G(v) = 1$.

Given any tree, we can fix a vertex as the "root". We can carry out depth-first search to traverse the tree's vertices starting from the root. An interesting phenomenon here is that at any point of time, we notice that the vertex which we are currently at is adjacent to exactly one of the vertices which we have visited.

> ### Proposition 2.1.6 ▶ Re-labelling of Vertices in a Tree
>
> *Let T be a tree with order n, then there exists an ordering of $V(T)$ such that for any $v_k \in V(T)$, v_k is adjacent to exactly one of $v_1, v_2, \cdots, v_{k-1}$.*
>
> *Proof.* By Corollary 2.1.4, T has a leaf vertex. Let this vertex be v_n. Note that $T - v_n \subseteq T$, so it is acyclic. Let $T' = T - v_n$, then v_n has a unique neighbour in $V(T')$. Note that T' is also a tree, so we can repeat this process to take v_{n-1}. Continue doing so until we have taken all vertices in T, then for any $k = 1, 2, \cdots, n$, v_k has a unique neighbour in $v_1, v_2, \cdots, v_k$. $\square$

In fact, such an ordering of the tree is known as a *topological sort*.

Intuitively, if a connected graph contains cycles, we can always reduce it to a tree by deleting some edges. This leads to the following result:

> ### Corollary 2.1.7 ▶ Number of Egdes in a Tree
>
> *A connected graph T is a tree of order n if and only if $e(T) = n - 1$.*
>
> *Proof.* If T is a tree, then by Proposition 2.1.6 we can order $V(T)$ such that for all $i = 2, 3, \cdots, n$, v_i is adjacent to exactly one vertex from $v_1, v_2, \cdots, v_{i-1}$. Therefore, $e(T) = n - 1$.
>
> If $e(T) = n - 1$, note that T contains a tree $T' \subseteq T$ which has n vertices. Therefore, $e(T') = n - 1 = e(T)$. This implies that $T = T'$, so T is a tree. $\square$

Recall that every graph contains a tree. In fact, this result can be further strengthened.

Proposition 2.1.8 ▶ Existence of Subgraph Trees

Let T be a tree with k vertices. If G is a graph with $\delta(G) \geq k - 1$, then G contains a subgraph which is isomorphic to T.

Proof. Fix an ordering on $V(T)$ such that v_i has a unique neighbour in $v_1, v_2, \cdots, v_{i-1}$. Take any vertex in G to be v_1. Suppose we have chosen $v_1, v_2, \cdots, v_i$ in G for some $i < k$, for any v_j among these vertices, at most $(i - 1)$ of v_j's neighbours have been chosen, but

$$d_G(v_j) \geq \delta(G) = k - 1 > i - 1.$$

Therefore, there exists some $u \in N_G(v_j)$ which is not one of $v_1, v_2, \cdots, v_i$. Take u as v_{i+1} with the edge uv_j, then v_{i+1} is adjacent to exactly one of $v_1, v_2, \cdots, v_i$. This means we can choose all vertices of T from $V(G)$, which means $T \subseteq G$. □

2.2   Bipartite Graphs

Consider a tree T. Take some $r \in V(T)$ as the root. We define a vertex $v \in V(T)$ to be at depth k from the root if there is a r-v path of length k. Clearly, if we partition $V(T)$ into vertices of odd depth and even depth, we can easily see that all edges in T joins one vertex from the former to one from the latter.

In fact, there are many graphs which can be bi-partitioned in a similar manner, which are known as *bipartite* graphs.

Definition 2.2.1 ▶ Bipartite Graph

A graph G is **bipartite** if $V(G)$ can be partitioned into two disjointed subsets V_1 and V_2 such that every edge of G joins a vertex in V_1 to a vertex in V_2. (V_1, V_2) is known as a **bipartition** of G.

Note that the size of a bipartite graph G with partite sets V_1 and V_2 is just

$$e(G) = \sum_{x \in V_1} d_G(x) = \sum_{y \in V_2} d_G(y).$$

It can be deduced that the maximal size of a bipartite graph is achieved when every vertex in V_1 is connected to every vertex in V_2. To define this rigorously, we first introduce the notion of *join*.

> **Definition 2.2.2 ▶ Join**
>
> Let G_1 and G_2 be two disjoint graphs of order n_1 and n_2 respectively. The **join** of G_1 and G_2, deonoted by $G_1 + G_2$ is the graph such that
>
> $$V(G_1 + G_2) = V(G_1) \cup V(G_2)$$
> $$E(G_1 + G_2) = E(G_1) \cup E(G_2) \cup \{uv : u \in V(G_1), v \in V(G_2)\}.$$

Informally, we can describe the join of G_1 and G_2 as a graph which contains all the vertices from G_1 and G_2, all existing edges from G_1 and G_2, and new edges connecting every vertex in G_1 to every vertex in G_2. Thus, we can define a complete bipartite graph using a join.

> **Definition 2.2.3 ▶ Complete Bipartite Graph**
>
> A **complete bipartite graph** is defined as $K_{p,q} = 0_p + 0_q$.

Consider a k-regular bipartite graph. Intuitively, if every vertex in the first partite set must be adjacent to $k > 0$ vertices in the second partite set and vice versa, then it is only natural that the two partite sets have equal cardinalities.

> **Proposition 2.2.4 ▶ Regular Bipartite Graph**
>
> *For all $k \in \mathbb{Z}^+$, a k-regular bipartite graph with bipartition (U, V) is such that $|U| = |V|$.*
>
> ---
>
> *Proof.* Note that we have
> $$k|U| = e(G) = k|V|.$$
> Since $k > 0$, this implies $|U| = |V|$. □

Sometimes, it is not easy to identify the bipartition in a bipartite graph. Therefore, to quickly determine whether a graph is bipartite, we consider the following theorem:

2.2.1   Chromatic Number

Suppose G is bipartite with bipartition (U, V), then we can colour the vertices in G such that all vertices in U receives one colour and all those in V receives another.

> **Definition 2.2.5 ▶ k-Colourability**
>
> A graph G is said to be k-**colourable** if one can colour $V(G)$ with k distinct colours such that every pair of adjacent vertices receive distinct colours.

Clearly, if a graph is k-colourable, it must be $(k + i)$-colourable for all $i \in \mathbb{N}$. Therefore, we are interested in the smallest possible k such that a graph is k-colourable.

> **Definition 2.2.6 ▶ Chromatic Number**
>
> The **chromatic number** of a graph G is the smallest positive integer k such that G is k-colourable, denoted as $\chi(G)$.

Observe that if $\chi(G) \leq 2$, it is equivalent to saying that we can find a bipartition of $V(G)$, so a bipartite graph can be alternatively defined as a graph G with $\chi(G) \leq 2$.

Furthermore, observe that if $H \subseteq G$, then $\chi(H) \leq \chi(G)$. Consider a cycle C_{2n+1} for some $n \in \mathbb{N}$. It is clear that $\chi(C_{2n+1}) = 3$, so if $C_{2n+1} \subseteq G$, then G cannot be bipartite.

> **Theorem 2.2.7 ▶ Bipartite Graph Characterisation**
>
> *A graph is bipartite if and only if it contains no odd cycles.*
>
> *Proof.* The necessity following from the previous argument, so we only prove the sufficiency. Suppose G contains no odd cycles. Note that there exists some spanning tree $T \subseteq G$ with root u. For every $v \in V(T)$, v is coloured red if $d_T(u, v)$ is odd, and blue otherwise.
>
> Suppose that there exist some $w_1, w_2 \in V(G)$ such that both receive the same colour but $w_1 w_2 \in E(G)$. Note that this means that there exist a w_1-u path in T and a w_2-u path in T with the same parity. This implies that there exists an closed walk of odd length in G, and so G contains an odd cycle, which is a contradiction. Therefore, G is bi-colourable and so is bipartite. $\square$

Similar to bipartite graphs, we can define the notion of an r-partite graph with chromatic numbers.

> **Definition 2.2.8 ▶ r-partite Graph**
>
> A graph G is an r-**partite graph** if $\chi(G) \leq r$.

3

Matchings and Vertex Covers

3.1   Matching

In set theory, we define the notion of a function

$$f : X \to Y$$

as a mapping such that every $x \in X$ has a unique image in Y. In a graph-theoretical perspective, we can view X and Y as vertex sets and a "mapping" in this case can be seen as a set of disjoint edges joining a vertex from X to a vertex from Y.

Definition 3.1.1 ▶ Matching

A **matching** in a graph G is a set $M \subseteq E(G)$ of vertex-disjoint edges.

The term "vertex-disjoint edges" means that for any $e, f \in M$, they are not incident to the same vertex in G. Here we define two notions which sound similar:

Definition 3.1.2 ▶ Maximal and Maximum Matchings

A **maximal matching** is a matching $M \subseteq E(G)$ such that for all $e \in E(G) - M, M \cup \{e\}$ is not a matching.
A **maximum matching** is a matching $M \subseteq E(G)$ such that for any matching $N \subseteq E(G), |N| \leq |M|$.

It is clear that a maximum matching must be maximal but not the other way round.

Definition 3.1.3 ▶ Matching Number

The **matching number** of a graph G is the size of the maximum matching in G, denoted as $\nu(G)$.

Clearly, for any graph $G, \nu(G) \leq \frac{\nu(G)}{2}$ as the best we can do is to partition $V(G)$ into disjoint pairs of vertices with respect to $E(G)$.

> ### Definition 3.1.4 ▶ Perfect Matching
>
> Let G be a graph. We say that G has a **perfect matching** if $\nu(G) = \frac{\nu(G)}{2}$.

> *Remark.* In particular, if G is a bipartite graph with partite sets A and B, we say that a matching $M \subseteq G$ is perfect to A if for any $a \in A$, $a \in M$.

Intuitively, if G has a perfect matching, then we can find a bipartition (U, V) of $V(G)$ such that there exists a bijection between U and V.

> ### Definition 3.1.5 ▶ Vertex Covering Number
>
> Let G be a graph. A **vertex cover** of G is defined to be a set $S \subseteq V(G)$ such that for all $e \in E(G)$, there is some $v \in S$ such that $v \in e$. The **vertex covering number** is defined as the size of the smallest vertex cover of G, denoted as $\tau(G)$.

Now let us examine some special graphs:

- Complete graph K_n: $\nu(K_n) = \left\lfloor \frac{n}{2} \right\rfloor, \tau(K_n) = n - 1$.

- Cycle C_n: $\nu(C_n) = \left\lfloor \frac{n}{2} \right\rfloor, \tau(C_n) = \left\lceil \frac{n}{2} \right\rceil$.

- Path P_n: $\nu(P_n) = \left\lfloor \frac{n+1}{2} \right\rfloor, \tau(P_n) = \left\lceil \frac{n}{2} \right\rceil$.

- Complete bipartite graph $K_{m,n}$: $\nu(K_{m,n}) = \tau(K_{m,n}) = \min\{m, n\}$.

- Hypercube graph Q^d: $\nu(Q^d) = \tau(Q^d) = 2^{d-1}$.

- Petersen graph P: $\nu(P) = 5, \tau(P) = 6$.

We may have guessed based on the above observation that $\tau(G) \geq \nu(G)$ for all graphs.

> ### Proposition 3.1.6 ▶ Vertex Covering Number Is at Most Matching Number
>
> *For any graph G, $\tau(G) \geq \nu(G)$.*
>
> *Proof.* Fix $M := \{e_1, e_2, \cdots, e_{\nu(G)}\}$ to be a maximum matching in G. Let $S \subseteq V(G)$ be any vertex cover of G. For any $i = 1, 2, \cdots, \nu(G)$, we have $|e_i \cap S| \geq 1$, otherwise S does not cover e_i. Therefore,
>
> $$\tau(G) = |S| = \sum_{i=1}^{\nu(G)} |e_i \cap S| \geq \nu(G).$$
>
> $\square$

3.2   Clique and Independent Set

Given any graph G, then theoretically, if we were to find a subgraph of order n, the one with the greatest size is a complete graph and that with the smallest size is an empty graph.

> **Definition 3.2.1 ▶ Clique**
>
> A **clique** in G is a complete subgraph of G. The **clique number** of G is the order of the largest clique in G, denoted as $\omega(G)$.

> **Definition 3.2.2 ▶ Independent Set**
>
> The **independent set** of G is a subset $I \subseteq V(G)$ such that for all $u, v \in I$, $uv \notin E(G)$. The **independence number** of G, denoted as $\alpha(G)$, is the size of the largest independent set in G.

Intuitively, if a set $S \subseteq V(G)$ induces an empty subgraph in G, it will induce a complete subgraph in $\overline{G}$. Therefore, $\alpha(G) = \omega(\overline{G})$.

We shall examine the independence numbers of some special graphs.

- Complete graph K_n: $\alpha(K_n) = 1$.

- Cycle C_n: $\alpha(C_n) = \left\lfloor \dfrac{n}{2} \right\rfloor$.

- Path P_n: $\alpha(P_n) = \left\lfloor \dfrac{n+1}{2} \right\rfloor$.

- Complete bipartite graph $K_{m,n}$: $\alpha(K_{m,n}) = \max\{m, n\}$.

- Hypercube graph Q^d: $\alpha(Q^d) = 2^{d-1}$.

- Petersen graph P: $\alpha(P) = 4$.

The following theorem discusses the numerical relationship between the independence number of vertex covering number:

> **Proposition 3.2.3 ▶ Sum of Independent Number and Vertex Covering Number**
>
> *For any graph G of order n, $\alpha(G) + \tau(G) = n$.*
>
> ---
>
> *Proof.* It suffices to prove that S is an independent set in G if and only if $V(G) - S$ is a vertex cover of G.
>
> Let $S \subseteq V(G)$ be an independent set, then for any edge $e = uv \in E(G)$, at least one of u and v is in $V(G) - S$ for otherwise there exists $u, v \in S$ such that $u \sim v$, which is not possible. Therefore, $V(G) - S$ is a vertex cover.

Suppose conversely that $S \subseteq V(G)$ is such that $V(G) - S$ is a vertex cover. $\qquad\square$

So far we have been talking a lot about the Petersen graph, but it turns out that the Petersen graph can be generalised into a *Kneser graph*.

Definition 3.2.4 ▶ Kneser Graph

The **Kneser graph** is a graph $KG(n, k)$ such that

$$V\big(KG(n, k)\big) = \{S \in \mathcal{P}([n]) : |S| = k\},$$
$$E\big(KG(n, k)\big) = \{(S, T) : S \cap T = \varnothing\}.$$

The vertex set of a Kneser graph is the set of all k-element subsets of $\{1, 2, \cdots, n\}$ and two vertices are adjacent if and only if their corresponding subsets are disjoint.

We introduce a theorem closely related to the Kneser graph:

Theorem 3.2.5 ▶ Erdos-Ko-Rado Theorem

Let $n \geq 2k$, then any family

$$\mathcal{F} \subseteq \{S \in \mathcal{P}([n]) : |S| = k\}$$

with $S \cap T \neq \varnothing$ for all $S, T \in \mathcal{F}$ is such that $|\mathcal{F}| \leq \binom{n-1}{k-1}$.

We can see that $\mathcal{F}$ is in fact a subset of $V(KG(n, k))$, so this is equivalent to saying that the biggest subset of $V(KG(n, k))$ such that no vertices are adjacent is C_{k-1}^{n-1}, which is just to prove that $\alpha(KG(n, k)) = C_{k-1}^{n-1}$.

Another property of the independence number is related to the notion of a *line graph*.

Definition 3.2.6 ▶ Line Graph

Let G be a graph. The **line graph** of G is defined as $L(G)$ with $V\big(L(G)\big) = E(G)$ and

$$E\big(L(G)\big) := \{(e, f) \in E(G)^2 : e \neq f, e \cap f \neq \varnothing\}.$$

In other words, $L(G)$ uses $E(G)$ as the vertex set and two vertices are adjacent if and only if the respective edges share at least 1 common vertex in G.

> **Proposition 3.2.7 ▶ Independence Number of Line Graphs**
>
> *For any graph G, $\nu(G) = \alpha(L(G))$.*
>
> *Proof.* Fix a maximum matching $M \subseteq G$, then for every $e, f \in E(M)$ with $e \neq f$, we have $ef \notin E(L(G))$, therefore, $E(M)$ is an independent set in $L(G)$ and so $\alpha(L(G)) \geq \nu(G)$.
>
> Suppose $\alpha(L(G)) > \nu(G)$, then there exists some $e_0 \in V(L(G)) = E(G)$ such that e_0 is vertex-disjoint with all $e \in E(M)$. However, this means that we can find a matching with bigger size in G, which is a contradiction. Therefore,
>
> $$\nu(G) = \alpha(L(G)).$$
>
> $\square$

3.3 König's Theorem

Let us take a slight detour and unpack the matching number from a different perspective. Note that while forming a matching in a graph G, we actually divide $E(G)$ into two subsets, one of which is included in the matching and the other excluded. Now, let us denote an included edge by letting it have a weight of 1, and assign 0 weight to all other edges which are unselected.

For every $e \in E(G)$, let x_e denote its weight. This way, to find a maximum matching is in fact equivalent to finding an optimal solution for the *integer program*

$$\max \sum_{e \in E(G)} x_e$$
$$\text{such that} \quad x_e \in \{0, 1\},$$
$$\sum_{v \in e} x_e \leq 1 \quad \text{for all } v \in V(G),$$

where the optimal value is just $\nu(G)$. Similarly, we can assign 0-1 weights, y_v, to every $v \in V(G)$, then a vertex cover is equivalent to an optimal solution for the integer program

$$\min \sum_{v \in V(G)} y_v$$
$$\text{such that} \quad y_v \in \{0, 1\},$$
$$\sum_{v \in e} y_v \geq 1 \quad \text{for all } e \in E(G),$$

where the optimal value is just $\tau(G)$. We can relax both programs into linear programs and it is clear that the 2 programs form a pair of primal and dual problems. Let the optimal values for the 2 linear programs be $\nu^*(G)$ and $\tau^*(G)$ respectively. By considering strong duality of linear programs, we can check that

$$\nu(G) \le \nu^*(G) = \tau^*(G) \le \tau(G).$$

Here $\nu^*(G)$ and $\tau^*(G)$ are known as *fractional matching number* and *fractional vertex covering number* respectively.

It turns out that for bipartite graphs, the above inequality can be further strengthened to strict equality. We shall first discuss a few useful concepts.

Fix some matching M in a bipartite graph G. Intuitively, we can find some paths in G which consisting of alternating edges from M and $G - M$.

> **Definition 3.3.1 ▶ Alternating Path**
>
> Let G be a graph and fix $M \subseteq G$ to be a matching. An **alternating path** in G with respect to M is a path $v_1 v_2 \cdots v_n$ such that $v_1 v_2 \in G - M$, $v_k v_{k+1} \in M$ for all even k and $v_k v_{k+1} \in G - M$ for all odd k.

Intuitively, if an alternating path contains more edges from $G - M$ than from M, we can utilise it to extend the matching M by doing a switch of edges.

> **Definition 3.3.2 ▶ Augmenting Path**
>
> An **augmenting path** with respect to a matching M in G is an alternating path which ends at some $v \in V(G) - V(M)$

In general, we can characterise a maximum matching by using augmenting paths.

> **Theorem 3.3.3 ▶ Maximum Matching Characterisation**
>
> *A matching $M \subseteq G$ is maximum if and only if there is no augmenting path in G with respect to M.*
>
> ---
>
> *Proof.* Suppose M is a maximum matching in G. We will prove that there is no augmenting path in G with respect to M by considering the contrapositive statement. Suppose we can find some augmenting path $P := v_1 v_2 \cdots v_n$ in G with respect to M, then P must contain an odd number of edges. Consider the subgraph
>
> $$M' := M - \{v_i v_{i+1} \in P : i \text{ is even}\} + \{v_j v_{j+1} \in P : j \text{ is odd}\}.$$

Note that M' is clearly a matching and $e(M') = e(M) + 1$, so M is not a maximum matching.

Suppose conversely that there is no augmenting path in G with respect to M. We will prove that M is maximum by considering the contrapositive statement. Suppose that M is not maximum, then there exists some matching M' with $e(M') > e(M)$. Consider the subgraph $H := M \triangle M'$. Note that $\Delta(H) = 2$ and that every connected component of H is either a path or a cycle whose edges alternate between $E(M)$ and $E(M')$. Note that H is bipartite, so every cycle in H must be even and so contain equal number of edges from M and M'. However, since $e(M') > e(M)$, we can find a path in H which begins and ends with edges in $M' - M$ and whose edges alternate between M and M'. This path is an augmenting path with respect to M. $\qquad\square$

With the above preliminary results established, we are now ready to prove the *König's Theorem*.

Theorem 3.3.4 ▸ König's Theorem

If G is bipartite, then $\nu(G) = \tau(G)$.

Proof. By Proposition 3.1.6, we already know that $\nu(G) \le \tau(G)$. Therefore, it suffices to prove that $\nu(G) \ge \tau(G)$. Fix $M \subseteq G$ to be a maximum matching. Let (A, B) be a bipartition of G, then for every $e \in E(M)$, we have $e = v_A v_B$ for some $v_A \in A$ and $v_B \in B$.

We select a subset of vertices from $V(M)$ with the following algorithm: for each $v_A v_B \in E(M)$, we take v_B if there exists some alternating path starting at a vertex in A ends at v_B. Otherwise, we take v_A. Let the vertex set produced this way be U. We claim that U is a vertex cover for G.

Take any edge $ab \in E(G)$. If $ab \in E(M)$, then $a \in U$ or $b \in U$. Otherwise, since M is a maximum matching in G which is bipartite, at least one of a and b is matched.

If $b \in V(M)$, consider a. If $a \notin V(M)$, then ab is an alternating path and so $b \in U$. Otherwise, $a \in V(M)$ which reduces the case to the other case.

If $a \in V(M)$ and $a \in U$ then we are done. Otherwise, $a \notin U$. Suppose $ab' \in E(M)$, then $b' \in U$. This means that we can find some alternating path P with

respect to M which ends at b'. If $b \in P$, then there is some alternating path with respect to M which ends at b. Otherwise, note that $a \notin P$ because a is matched, so $P + ab$ is an alternating path which ends at b with respect to M. Since M is maximum, by Theorem 3.3.3, this alternating path is not augmenting, which means that $b \in V(M)$. Therefore, $b \in U$.

Therefore, U is a vertex cover for G. Note that $|U| = \nu(G)$, so $\tau(G) \leq \nu(G)$. This means that $\nu(G) = \tau(G)$. $\qquad\square$

We can generalise Theorem 3.3.4 to *fractional vertex covers*.

> **Definition 3.3.5 ▶ Fractional Vertex Cover**
>
> The **fractional vertex cover** of a non-negatively weighted graph G is defined as the function
> $$c : V(G) \to \mathbb{R}_0^+$$
> such that $c(v) + c(u) \geq w(uv)$ for all $u, v \in V(G)$, where $w : E(G) \to \mathbb{R}_0^+$ is the weight function.

We shall state the generalised theorem without proof.

> **Theorem 3.3.6 ▶ Egerváry's Theorem**
>
> *For every bipartite G, the minimum weight of a fractional vertex cover equals the weight of the matching with the maximum weight.*

3.4   Hall's Marriage Theorem

Suppose G is a bipartite graph with bipartition (A, B). In the case where $|A| \neq |B|$ we cannot get any matching with size greater than $\min\{|A|, |B|\}$ for the obvious reason. Therefore, the best we can hope is to construct a matching **perfect to one of the partite sets**.

Suppose $S \subseteq A$ is a set of vertices in G. We denote the union of the neighbourhoods of all vertices in S by $N_G(S)$. Apparently, if we can find some S such that $|N_G(S)| < |S|$, it means that we cannot find a matching perfect to A because we simply cannot match every vertex in S.

This means that for G to contain a matching perfect to A, it necessarily needs to satisfy $|S| \leq |N_G(S)|$ for every $S \subseteq A$. However, it turns out that amazingly, this necessary condition alone is also sufficient to ascertain the existence of a matching perfect to A.

> ### Theorem 3.4.1 ▶ Hall's Theorem
>
> *Let G be a bipartite graph G with bipartition (A, B). G contains a matching perfect to A if and only if for every $S \subseteq A$,*
> $$|N_G(S)| \geq |S|.$$

> *Proof.* Suppose $M \subseteq G$ is a matching perfect to A. Let $S \subseteq A$ be any subset. Clearly, for each $a \in S$, there exists a unique $b \in B$ such that $ab \in E(M)$ and if $ab, a'b' \in E(M)$ for $a \neq a'$, we have $b \neq b'$. Therefore, $|N_G(S)| \geq |S|$.
>
> Suppose that for every $S \subseteq A$, $|N_G(S)| \geq |S|$, we take any matching $M \subseteq G$ such that $A - V(M) \neq \varnothing$. Take $(a_0, b_1, a_1, \cdots)$ to be a finite sequence of pair-wise distinct vertices with greatest length such that a_0 is unmatched, b_i is adjacent to some $a_{f(i)} \in \{a_0, a_1, \cdots, a_{i-1}\}$ in G and $b_i a_i \in M$ for all $i \geq 1$.
>
> We claim that this sequence does not end in A. Suppose on contrary that the sequence ends at some $a_{n-1} \in A$. Take $T = \{a_0, a_1, \cdots, a_{n-1}\}$, then $|N_G(T)| \geq n$. Therefore, there exists some $b_n \in N_G(T) - \{b_1, b_2, \cdots, b_{n-1}\}$, which is a contradiction because we can now extend the length of the sequence.
>
> Let b_k be the end of the sequence, then the path $P \coloneqq a_0 b_1 a_{f(1)} b_2 a_{f(2)} \cdots a_{f(k-1)} b_k$ is an alternating path. If b_k is matched to some $a \in A \cap V(M)$, we know that $a \notin \{a_1, a_2, \cdots, a_{k-1}\}$ because that would imply $b_k = b_i$ for some $i \in \{1, 2, \cdots, k-1\}$, which is a contradiction. Since a_0 is unmatched, we have find a longer sequence satisfying our construction, which is not possible. Therefore, b_k must be unmatched. This means that P is an augmenting path and so by Theorem 3.3.3 M is not a maximum matching. This means that the maximum matching in G must be perfect to A. □

Beside the above proof, there are several alternative ways to prove Theorem 3.4.1. Here we first present a proof by contradiction using Theorem 3.3.4.

> ### Proof for Hall's Theorem by König's Theorem
>
> *Proof.* Observe that it suffices to show that $\nu(G) = |A|$. First, note that since G is bipartite, it is not possible to have $\nu(G) > |A|$ because otherwise there exists an edge between 2 distinct vertices in B. Suppose on contrary that $\nu(G) < |A|$, then by Theorem 3.3.4, $\tau(G) = \nu(G) < |A|$. This means that there exists a vertex cover with strictly less vertices than A.

Suppose $U = A' \cup B'$ is a vertex cover for G with $A' \subseteq A$ and $B' \subseteq B$, then $|A'| + |B'| < |A|$. Clearly, there is no edge between $A - A'$ and $B - B'$. Take $S := A - A'$, then $N_G(S) \subseteq B'$. Therefore,

$$|N_G(S)| \leq |B'| < |A| - |A'| = |S|,$$

which is a contradiction. □

Another proof uses an argument by strong induction:

Proof for Hall's Theorem by Strong Induction

Proof. We shall perform strong induction on $|A|$. The case where $|A| = 1$ is trivial. Suppose that for all $|A| < n$, the theorem holds. Let G be a bipartite graph with bipartition (A, B) such that $|A| = n$. We consider 2 cases.

Suppose that for any proper subset $S \subset A$, we have $|N_G(S)| \geq |S| + 1$. Consider an edge $ab \in E(G)$ and define $G' := G - \{a, b\}$ be the graph with the vertices a and b removed. Let (A', B') be the bipartition of G'. Notice that for any $S' \subseteq A'$, we have $|N_{G'}(S')| \geq |S'|$. By the inductive hypothesis, this means that G' contains a matching M' perfect to A'. Clearly, $M' + ab$ is a matching perfect to A in G.

Suppose otherwise that there exists some proper subset $A' \subset A$ with $|N_G(A')| = |A'|$. Let $B' := N_G(A')$ and define

$$G' := G[A' \cup B'], \qquad G'' := G[(A - A') \cup (B - B')].$$

Note that for any $S \subseteq A'$, $N_G(S) \subseteq B'$, and so $|N_{G'}(S)| = |N_G(S)| \geq |S|$. By the inductive hypothesis, this implies that G' contains a matching M' perfect to A'. For any $T \subseteq A - A'$, we claim that $|N_{G''}(T)| \geq |T|$.

Suppose on contrary that $|N_{G''}(T)| < |T|$. Note that $N_G(T \cup A') \subseteq N_{G''}(T) \cup B'$, so

$$|N_G(T \cup A')| \leq |N_{G''}(T)| + |B'| < |T| + |A'| = |T \cup A|,$$

which is a contradiction. Therefore, this implies that G'' contains a matching M'' perfect to $A - A'$. Note that $M' \cup M''$ is a matching in G perfect to A. □

The last proof we are going to present stems from the following observation: suppose we have a graph G satisfying the conditions in Hall's Theorem, we wish to delete "redundant" edges until we cannot remove any more edge without violating the condition. We claim

that the remaining graph is a matching perfect to A.

> ### Proof for Hall's Theorem by Edge-Minimal Subgraph
>
> *Proof.* Let $H \subseteq G$ be the subgraph with the smallest size such that for every $S \subseteq A$, $|N_H(S)| \geq |S|$. We shall prove that H is a matching perfect to A.
>
> First, observe that $d_H(a) \geq 1$ for all $a \in A$. Suppose on contrary that H is not a matching perfect to A, then there exists some $a \in A$ which has two distinct neighbours $b_1, b_2 \in B$ in H. Define $H_1 := H - ab_1$ and $H_2 := H - ab_2$, then by the minimality of H, there exist subsets $A_1, A_2 \subseteq A$ both containing a such that $|N_{H_1}(A_1)| < |A_1|$ and $|N_{H_2}(A_2)| < |A_2|$. By our construction of H_1 and H_2, we know that $b_1 \in N_{H_2}(A_2) - N_{H_1}(A_1)$ and $b_2 \in N_{H_1}(A_1) - N_{H_2}(A_2)$. Consider
>
> $$S := A_1 \cap A_2 - \{a\},$$
>
> then $N_H(S) \subseteq N_{H_1}(A_1) \cap N_{H_2}(A_2)$. Therefore,
>
> $$\begin{aligned}
> |N_H(S)| &\leq |N_{H_1}(A_1) \cap N_{H_2}(A_2)| \\
> &= |N_{H_1}(A_1)| + |N_{H_2}(A_2)| - |N_{H_1}(A_1) \cup N_{H_2}(A_2)| \\
> &= |N_{H_1}(A_1)| + |N_{H_2}(A_2)| - |N_H(A_1 \cup A_2)| \\
> &\leq |A_1| - 1 + |A_2| - 1 - |N_H(A_1 \cup A_2)| \\
> &= |A_1 \cap A_2 - \{a\}| - 1,
> \end{aligned}$$
>
> which is a contradiction. $\qquad\square$

Here we introduce a famous corollary of Hall's Theorem:

> ### Corollary 3.4.2 ▶ Regular Bipartite Graphs Have A Perfect Matching
>
> *If G is a d-regular bipartite graph with bipartition (A, B), then G has a perfect matching.*
>
> *Proof.* It suffices to prove that G contains a matching perfect to A. Since $e(G) = d|A| = d|B|$, we have $|A| = |B|$. Take any $X \subseteq A$ and consider the induced subgraph
>
> $$H := G[X \cup N_G(X)].$$
>
> Note that for every $v \in N_G(X)$, $d_H(v) \leq d_G(v) = d$, so $d|N_G(X)| \geq e(H) = d|X|$. Therefore, $|X| \leq |N_G(X)|$ and by Theorem 3.4.1, G contains a matching perfect to A. $\qquad\square$

The above result immediately leads to the following corollary:

Corollary 3.4.3 ▶ Decomposition of Regular Bipartite Graphs

Let G be a d-regular bipartite graph, then $E(G)$ can be decomposed into perfect matchings.

Proof. By Corollary 3.4.2, we can find some perfect matching M_d in G. Consider the graph G' with $V(G') = V(G)$ and $E(G') = E(G) - E(M_d)$, then G' is a $(d-1)$-regular bipartite graph and so it contains a perfect matching M_{d-1}. Note that $E(M_{d-1}) \cap E(M_d) = \varnothing$ and that M_{d-1} is also a perfect matching in G. Repeat this process until we have used all edges in G. The set of all perfect matchings obtained this way forms a partition of $E(G)$. □

3.5   Tutte's Theorem

In the previous section, we have discussed how Hall's Theorem characterises the existence of perfect matchings in bipartite graphs. Can we derive a similar result for any kind of graphs in general?

It turns out that this curious question is closely related to the number of connected components with odd orders in a graph. We denote this quantity by $o(G)$ for any graph G.

Theorem 3.5.1 ▶ Tutte's Theorem

A graph G has a perfect matching if and only if for every $U \subseteq V(G)$, $o(G - U) \leq |U|$.

Proof. Let $M \subseteq G$ be a perfect matching. Take any $U \subseteq V(G)$ and consider the induced subgraph $H := G - U$. Let $C \subseteq H$ be a connected component of odd order. Since M is a perfect matching, then clearly there is at least one vertex $v \in C$ which is matched to some vertex $u \in U$ in M. Since every $u \in U$ is matched to exactly one vertex in M, this implies that $o(H) \leq |U|$.

Suppose conversely that for every $U \subseteq V(G)$, $o\big(G[V(G) - U]\big) \leq |U|$. We shall prove that G contains a perfect matching by considering the contrapositive statement. Let $v(G) = n$. First, we consider the following lemma:

Lemma 3.5.2

If graphs $G \subseteq G'$ are such that $V(G) = V(G')$ but $E(G) \subseteq E(G')$, then if $U \subseteq V(G')$ is such that $o(G' - U) > |U|$, we have $o(G - U) > |U|$.

Proof. Note that G is produced by deleting some edges in G'. For each edge $e \in E(G') - E(G)$, if e is in an odd component $H_1 \subseteq G'$, then $H_1 - e$ contains at least one odd component in G. Otherwise, if e is in an even component $H_2 \subseteq G'$, then $H_2 - e$ may or may not contain any odd component. Therefore, we have

$$o(G - U) \geq o(G' - U) > |U|.$$

$\square$

By Lemma 3.5.2, without loss of generality we can take G to be the edge-maximal graph of order n such that G does not contain any perfect matching.

Consider $U := \{v : v \sim u \text{ for all } u \in V(G) - v\}$, i.e., U consists of vertices in G which are adjacent to every other vertex in G, then $G[U]$ is a clique. We claim that every connected component of $G - U$ is complete. Suppose on contrary that there is some connected component $C_i \subseteq G - U$ which is not complete, then there exists distinct vertices $a, a' \in V(C_i)$ such that $a \nsim a'$. Therefore, there exists an a-a' path in C_i of length at least 2, and so we can find vertices $a, b, c \in V(C_i)$ such that $a \sim b$, $b \sim c$ but $a \nsim c$. Note that $b \notin U$, so there exists some $d \in V(G)$ such that $b \nsim d$.

Define graphs $G_1 := G + ac$ and $G_2 := G + bd$. By the edge-maximality of G, G_1 contains a perfect matching M_1 and G_2 contains a perfect matching M_2. Note that $ac \in E(M_1)$, because otherwise M_1 is a perfect matching with vertex set $V(G)$ and edge set $E(G)$, which is not possible. Similarly, $bd \in E(M_2)$ Let $P \subseteq G$ be a maximal d-v path for some $v \in V(G)$ which is alternating between edges in M_1 and M_2.

Suppose P ends with an M_1-edge, then v must be matched to d in M_2 because otherwise we can find a longer M_1-M_2 alternating path than P, which is not possible. Since $bd \in E(M_2)$, this implies that $v = b$. Note that $M_1 \cap P \subseteq G$ is a perfect matching in P and $M_2 - P - bd$ is a perfect matching in $G - P$, then $(M_1 \cap P) \cup (M_2 - P - bd)$ is a perfect matching in G.

Similarly, suppose P ends with an M_2-edge, then v must be either a or c, because otherwise there exists an edge $uv \in E(M_1) \cap E(G)$ that can further extend

P, which is not possible. Without loss of generality, suppose $v = c$, then it is clear that $a \notin P$. We claim that $b \notin P$. Suppose on contrary that $b \in P$, then there must be an M_1 edge in P containing b because there is no M_2 edge in P containing b. However, this means that P ends with an M_2 edge, which is a contradiction. Note that $M_2 \cap P \subseteq G$ is a perfect matching in P and $M_1 - P - ac + ab$ is a perfect matching in $G - P$. Therefore, $(M_2 \cap P) \cup (M_1 - P - ac + ab)$ is a perfect matching in M.

Either way, we have reached a contradiction because M does not contain any perfect matching. Therefore, every connected component of $G - U$ is complete. If $U = \varnothing$, then $G - U = G$ is complete. Since G has no perfect matching, $v(G)$ must be odd and so $o(G - U) = 1 > 0 = |U|$. Otherwise, for each connected component $C_i \subseteq G$, if C_i has an even order, it contains a perfect matching. Otherwise, C_i has an odd order. Take some $v_i \in V(C_i)$ and match it to some $u_i \in U$. Note that $C_i - v_i$ is a complete graph of even order and so it contains a perfect matching. Ho $\qquad \square$

Tutte's Theorem has many interesting applications, one of which is related to the concept of an *cut edge* or *cut vertex*.

Theorem 3.5.3 ▶ Petersen's Theorem

Every connected 3-regular graph with no cut edge contains a perfect matching.

Proof. Let $U \subseteq V(G)$ be non-empty and consider $H := G - u$. If H contains no odd-order component, then we are done. Otherwise, fix some odd-order component $C \subseteq H$. Consider

$$\sum_{v \in V(C)} d_G(v) = 3|C| = 2e(G[V(C)]) + e_G(V(C), U).$$

Note that the above sum is odd, so $e_G(V(C), U)$ is odd. Since G contains no cut-edge, $e_G(V(C), U) \geq 3$. Let p be the number of odd components in $G - u$ and M be the number of edges between U and $V(G) - U$, then $M \geq 3p$. However, $M \leq 3|U|$ because G is 3-regular, so $p \leq |U|$. By Theorem 3.5.1, G contains a perfect matching. $\qquad \square$

Note that a graph $G = (V, E)$ can actually be viewed as a collection of 2-element subsets of V. Therefore, a natural generalisation is to consider a graph defined as the collection of subsets of any size of V.

Definition 3.5.4 ▶ Hypergraph

Let V be a vertex set, a **hypergraph** on V is a graph whose edge set is a family

$$E := \{X : X \in \mathcal{P}(V)\}.$$

If $|X| = k$ for all $X \in E$, we say that the hypergraph is k-uniform.

4

Connectivity

4.1 Vertex Connectivity

Recall that in previous chapters, we have already defined what a *connected graph* is (Definition 1.2.5). However, intuitively, we may feel that graphs may have varying degrees of "connectedness". For example, a complete graph seems more connected than a path, since one could easily disconnect a path by deleting $\mathcal{O}(1)$ vertices but needs to delete $\mathcal{O}(n)$ vertices to disconnect K_n.

In this chapter, we probe further into the notion of connectivity. In particular, we would like to discuss the strength of connectedness in a graph.

> **Definition 4.1.1 ▶ Separating Set**
>
> Let G be a connected graph. $S \subseteq V(G)$ is a **separating set** (or **vertex cut**) of G if $G - S$ is disconnected. If $S = \{s\}$ is a singleton, then s is known as a **cut vertex**.

Separating set provides an intuitive measure for how connected a graph really is. Obviously, any finite graph must have only finitely many different separating set. Naturally, we could consider the size of the smallest separating set as an indicator for the amount of deletion needed to disconnect the graph.

> **Definition 4.1.2 ▶ Vertex Connectivity**
>
> The **vertex connectivity** of a graph G, denoted by $\kappa(G)$ is the cardinality of the smallest subset $S \subseteq V(G)$ such that $G - S$ is either disconnected or a trivial graph. G is said to be k-**connected** if $\kappa(G) \geq k$.

Obviously, $\kappa(G) \leq v(G) - 1$ for any graph G, and a k-connected graph must be $(k - 1)$-connected. Our "old-fashioned" notion of a connected graph is simply a 1-connected graph. As usual, let us examine some special graphs as examples.

1. Complete graph K_n: $\kappa(K_n) = n - 1$.

2. Path P_n: $\kappa(P_n) = 1$.

3. Non-trivial cycle C_n: $\kappa(C_n) = 2$.

4. Complete bipartite graph $K_{m,n}$: $\kappa(K_{m,n}) = \min\{m, n\}$.

We may also wish to consider the hypercube graph Q^d. Intuitively, Q^d is d-regular, so it suffices to fix any $v \in V(Q^d)$ and delete all of its d neighbours to isolate the vertex. Therefore, $\kappa(Q^d) \leq d$. However, can we force the inequality to an equality?

Proposition 4.1.3 ▶ The Hypercube Graph Q^d is d-connected

Every hypercube graph Q^d is d-connected.

Proof. The case where $d = 1$ is trivial. Suppose that for some $d \in \mathbb{N}^+$, Q^d is d-connected. Consider Q^{d+1}. Notice that $Q^{d+1} = Q^+ \cup Q^- \cup M$, where Q^+ and Q^- are both Q^d and M is a perfect matching joining vertices in Q^+ and those in Q^- with exactly 1 different digit. Let S be the smallest vertex cut in Q^{d+1}. We consider 2 cases.

If $Q^+ - S$ and $Q^- - S$ are both connected, then $V(M) \subseteq S$ because otherwise there is an edge between $Q^+ - S$ and $Q^- - S$ in $Q^{d-1} - S$ and so $Q^{d-1} - S$ is connected, which is a contradiction. Therefore,

$$\kappa(Q^{d+1}) = |S| \geq |M| = 2^d \geq d + 1.$$

Otherwise, without loss of generality assume that Q^+ is disconnected, then $S \cap Q^+$ is a vertex cut in Q^+. By the inductive hypothesis, $|S \cap Q^+| \geq d$. We claim that $|S \cap Q^-| \geq 1$. Suppose on contrary that $|S \cap Q^-| = 0$, then $S \cap V(Q^-) = \varnothing$. This means that $Q^- - S$ is connected. Note that for every $v \in V(Q^+ - S)$, there exists some $u \in V(Q^- - S)$ such that $uv \in E(Q^{d+1} - S)$, because otherwise $S \cap V(Q^-) \neq \varnothing$. Therefore, for any $v_1, v_2 \in V(Q^+ - S)$, we can find $u_1, u_2 \in V(Q^- - S)$ such that there is a v_1-v_2 path in Q^{d+1} containing $v_1 u_1$ and $v_2 u_2$. Therefore, $Q^{d+1} - S$ is connected, which is a contradiction. Therefore,

$$\kappa(Q^{d+1}) = |S| = |S \cap Q^+| + |S \cap Q^-| \geq d + 1.$$

This means that Q^{d+1} is $(d + 1)$-connected. $\square$

Next, we will focus on 2-connected graph, i.e., graphs which we need to delete at least 2 vertices to disconnect. We will start with a straight-forward proposition.

Proposition 4.1.4 ▶ 2-connected Graphs Contain A Cycle

Every 2-connected graph contains a cycle.

Proof. Suppose on contrary that there exists a 2-connected graph T which is acyclic. Note that T is connected, so T must be a tree. By Theorem 2.1.2, $T - uv$ is disconnected for any $uv \in E(T)$. If both u and v are leaves, then T is an edge, which implies that

deleting any vertex reduces T to a trivial graph. Otherwise, without loss of generality, assume u is not a leaf, and so $T - u$ is disconnected. Either way, this contradicts with the fact that $\kappa(T) \geq 2$. $\qquad\square$

It should be taken note of that the converse of Proposition 4.1.4 is not true in general. Let C_n be a cycle, then adding a vertex v into C_n by joining it with any $u \in V(C_n)$ produces a graph H containing a cycle. However, $H - u$ is disconnected so $\kappa(H) = 1$.

An intuition about connectivity is that, if we were to increase the connectivity of a graph, we should try to make each vertex to have as many incident edges as possible.

Proposition 4.1.5 ▶ High Connectivity Implies High Minimum Degree

For any graph G, $\kappa(G) \leq \delta(G)$.

Proof. Take some $v \in V(G)$ with $d_G(v) = \delta(G)$, then clearly $N_G(v)$ is a vertex cut because v is an isolated vertex in $G - N_G(v)$, and so $\kappa(G) \leq |N_G(v)| = \delta(G)$. $\qquad\square$

The above trivial proposition shows that a graph with high connectivity has high minimum degree. However, the converse is clearly not true. Consider the graph G formed by 2 disjoint copies of K_n, then clearly $\delta(G) = n - 1$ but $\kappa(G) = 0$, i.e., G is not even connected.

Recall that in Proposition 1.1.13, while we cannot ensure a graph with a high average degree to have a high minimum degree, we have proved that the graph will contain a subgraph of a high minimum degree. In the same spirit, given a graph with a high minimum degree, we can try to find a subgraph with a high connectivity.

Theorem 4.1.6 ▶ Mader's Theorem

For every graph G, if $\bar{d}(G) \geq 4k$, then G contains a k-connected subgraph.

Proof. For $k = 1$, we have $\bar{d}(G) \geq 4$. This means that G is non-empty and so it contains an edge which is a 1-connected subgraph.

For $k \geq 2$, we consider the following claims: first, suppose $v(G) < 2k - 1$, then $\bar{d}(G) \leq v(G) - 1 < 2k - 2$. Therefore, if $\bar{d}(G) \geq 4k > 2k - 2$, we have $v(G) \geq 2k - 1$. On the other hand, notice that

$$e(G) = \frac{v(G)\bar{d}(G)}{2} \geq (2k - 3)\big(v(G) - k + 1\big) + 1.$$

Therefore, it suffices to prove that whenever $k > 2$, if $v(G) \geq 2k - 1$ and $e(G) \geq (2k - 3)\big(v(G) - k + 1\big) + 1$, then G contains a k-connected subgraph.

For $v(G) = 2k - 1$, we have

$$e(G) \geq k(2k - 3) + 1 = (2k - 1)(k - 1) = \binom{2k - 1}{2}.$$

However, we know that $e(G) \leq \binom{v(G)}{2}$, so $G \cong K_{2k-1}$. Therefore, $\kappa(G) = 2k - 2 \geq k$ and so G is k-connected.

For $v(G) = n \geq 2k$, suppose that for all $v(G) = m \in [2k - 1, n - 1]$, G contains a k-connected subgraph. We consider 2 cases. If $\delta(G) \leq 2k - 3$, then we can take some $v \in V(G)$ with $d_G(v) = \delta(G)$. Consider

$$\begin{aligned}
e(G - v) &= e(G) - \delta(G) \\
&\geq (2k - 3)(n - k + 1) + 1 - (2k - 3) \\
&= (2k - 3)\big((n - 1) - k + 1\big) + 1.
\end{aligned}$$

By the inductive hypothesis, $G - v$ contains a k-connected subgraph H. Since $G - v \subseteq G$, $H \subseteq G$.

If $\delta(G) \geq 2k - 2 > 2k - 3$, we again consider 2 cases. If G is k-connected, then we are done. Otherwise, G contains a vertex cut with size at most $k - 1$. Therefore, we can consider $G_1, G_2 \subseteq G$ such that $G_1 \cup G_2 = G$ and $V(G_1) \cap V(G_2)$ is a vertex cut such that $|V(G_1) \cap V(G_2)| \leq k - 1$. Notice that this implies that there is no edge between $V(G_1) - V(G_2)$ and $V(G_2) - V(G_1)$. Consider any $v \in V(G_1)$. Note that $N_G(v) \subseteq V(G_1)$ and $d_G(v) \geq \delta(G) \geq 2k - 2$, so $|V(G_1)| \geq d_G(v) + 1 = 2k - 1$. Similarly, $|V(G_2)| \geq 2k - 1$. Suppose that

$$\begin{aligned}
e(G_1) &\leq (2k - 3)\big(v(G_1) - k + 1\big), \\
e(G_2) &\leq (2k - 3)\big(v(G_2) - k + 1\big),
\end{aligned}$$

then

$$\begin{aligned}
e(G) &\leq e(G_1) + e(G_2) \\
&\leq (2k - 3)\big(v(G_1) + v(G_2) - 2k + 2\big) \\
&= (2k - 3)\big(n + |V(G_1) \cap V(G_2)| - 2k + 2\big) \\
&\leq (2k - 3)\big(n + k - 1 - 2k + 2\big) \\
&= (2k - 3)(n - k + 1),
\end{aligned}$$

which is a contradiction. Therefore, without loss of generality, we have

$$e(G_1) \geq (2k-3)\big(v(G_1) - k + 1\big) + 1.$$

Therefore, by the inductive hypothesis, G_1 contains a k-connected subgraph H. Since $G_1 \subseteq G$, $H \subseteq G$. Therefore, any graph G with $\bar{d}(G) \geq 4k$ contains a k-connected subgraph. $\qquad\square$

Since a high minimum degree implies a high average average degree, Theorem 4.1.6 in fact guarantees that a graph with a high minimum degree will have a high connectivity.

4.2   Edge Connectivity

Since we can measure connectivity in terms of the minimum number of vertices which need to be deleted to disconnected a graph, it is only natural to consider from the perspective of edges.

> **Definition 4.2.1 ▶ Edge Connectivity**
>
> A graph G is said to be ℓ-**edge-connected** if for any $F \subseteq E(G)$ with $|F| < \ell$, $G - F$ is connected. The **edge connectivity** of G, denoted by $\kappa'(G)$, is defined as the size of the smallest subset $F \subseteq E(G)$ such that $G - F$ is disconnected.

Similar to separating sets, we can analogously define a set of edges which will disconnect the graph when removed.

> **Definition 4.2.2 ▶ Disconnecting Set**
>
> Let G be a graph, a set $F \subseteq E(G)$ is called a **disconnecting set** if $G - F$ is disconnected.

However, as opposed to vertex cuts, the notion of an *edge cut* is slightly different.

> **Definition 4.2.3 ▶ Edge Cut**
>
> Let G be a graph and $(S, \bar{S})$ be a partition for $V(G)$, an **edge cut** in G with respect to $(S, \bar{S})$ is defined as
>
> $$E_G\big(S, \bar{S}\big) := \big\{xy \in E(G) : x \in S, y \in \bar{S}\big\}.$$
>
> If an edge cut contains only a single edge, that edge is known as a **cut edge** or **bridge**.

Intuitively, any edge set is a disconnecting set, but not every disconnecting set is an edge

cut. However, it turns out that every **minimal disconnecting set** is an edge cut! Let us do the routine examination on the special graphs.

1. Complete graph K_n: $\kappa'(K_n) = n - 1$.

2. Path P_n: $\kappa'(P_n) = 1$.

3. Non-trivial cycle C_n: $\kappa'(C_n) = 2$.

4. Complete bipartite graph $K_{m,n}$: $\kappa'(K_{m,n}) = \min\{m, n\}$.

5. Hypercube graph Q^d: $\kappa'\left(Q^d\right) = d$.

Notice that in all of these examples, $\kappa = \kappa'$. If for any graph G, $\kappa(G) = \kappa'(G)$, things will not be very interesting. However, this is luckily not the case. Consider a graph formed by 2 complete graphs of order n with exactly one common vertex. One may check that we only need to remove the shared vertex to disconnect the graph but need to delete at least $(n-1)$ edges to do the same.

In fact, we can derive the following relation:

Proposition 4.2.4 ▸ Numerical Relationship between Connectivities

For every graph G, $\kappa(G) \le \kappa'(G) \le \delta(G)$.

Proof. If G is disconnected, the result is trivial.

Otherwise, let $v \in V(G)$ be such that $d_G(v) = \delta(G)$. Note that deleting all edges incident to v will isolate the vertex, so $\kappa'(G) \le \delta(G)$.

Let $E\left(S, \bar{S}\right)$ be an edge cut with $\left|E\left(S, \bar{S}\right)\right| = \kappa'(G)$. We consider 2 cases. If $E\left(S, \bar{S}\right)$ induces a complete graph, then

$$
\begin{aligned}
\kappa'(G) &= |S||\bar{S}| \\
&= |S|\left(v(G) - |\bar{S}|\right) \\
&\ge v(G) - 1 \\
&\ge \kappa(G).
\end{aligned}
$$

Otherwise, there exist some $x \in S$ and $y \in \bar{S}$ such that $xy \notin E\left(S, \bar{S}\right)$. Define

$$
U := N_G(x) \cap \bar{S} \cup \left\{v \in S - \{x\} : N_G(v) \cap \bar{S} \ne \varnothing\right\}.
$$

Consider $H := G - U$. Notice that there is no x-y path in H because $E_H\left(S - U, \bar{S}\right) = \varnothing$.

Therefore, H is disconnected and so U is a vertex cut. Therefore,

$$\kappa(G) \leq |U| \leq |S| \leq |E(S, \bar{S})| = \kappa'(G).$$

Therefore, $\kappa(G) \leq \kappa'(G) \leq \delta(G)$. $\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Remark. Note that there exists some graph G with $\kappa(G) < \kappa'(G) < \delta(G)$.

Lastly, we will state the following theorem without proof:

> **Theorem 4.2.5 ▶ Chartrad-Haray Theorem**
>
> *For all positive integers a, b, c with $a \leq b \leq c$, there exists a graph G with $\kappa(G) = a$, $\kappa'(G) = b$ and $\delta(G) = c$.*

4.3   Blocks

We know that every graph can be naturally decomposed into connected components. However, can we actually present a stronger definition? Here, we consider components of a graph that are 2-connected.

> **Definition 4.3.1 ▶ Block**
>
> For any graph G, a **block** in G is defined as a maximal connected subgraph $B \subseteq G$ without any cut vertex.

Remark. We define an edge K_2 and a trivial graph both as blocks.

Let us make some observations regarding the properties of a block.

First, observe that a block is either an isolated vertex, a cut edge or a 2-connected graph. Next, we shall characterise the vertices of blocks.

> **Proposition 4.3.2 ▶ Blocks Share at Most One Common Vertex**
>
> *Let B_i and B_j be distinct blocks of G, then $|V(B_i) \cap V(B_j)| \leq 1$.*
>
> *Proof.* Suppose on contrary that $|V(B_i) \cap V(B_j)| \geq 2$. Consider $B := B_i \cup B_j$. Note that for any vertex $v \in V(B_i) \triangle V(B_j)$, v is not a cut vertex of B. Take any $u_1, u_2 \in V(B_i) \cap V(B_j)$ and consider $B - u_1$. Let $w_1, w_2 \in V(B - u_1)$ be any vertices. If $w_1, w_2 \in V(B_i)$ or $w_1, w_2 \in V(B_j)$, then they are connected in $B - u_1$. Otherwise, without loss of

generality, assume $w_1 \in V(B_i) - V(B_j)$ and $w_2 \in V(B_j) - V(B_i)$. Note that there exist a u_2-w_1 path and a u_2-w_2 path in $B - u_1$, so w_1 and w_2 are connected in $B - u_1$. Therefore, B does not have a cut vertex and so is a block, but $B_i, B_j \subset B$, which is a contradiction. □

The above result allows us to show the following corollary about the edges in a graph with respect to its blocks:

Corollary 4.3.3 ▸ Every Edge Is in A Unique Block

Let G be a graph. Every $e \in E(G)$ is in a unique block of G.

Proof. First, notice that every edge is a connected subgraph of G without any cut vertex, so any edge $e \in E(G)$ must be in some block of G. Suppose on contrary that there exist distinct blocks $B_i, B_j \subseteq G$ such that $e \in E(B_i) \cap E(B_j)$, then $|V(B_i) \cap V(B_j)| \geq 2$, which contradicts Proposition 4.3.2. □

By the above corollary, we know that no two different blocks can share common edges.

It is clear that every graph can be decomposed into blocks. We would like to further study block decomposition and hopefully derive a relation between the structure of blocks and that of the graph itself.

Definition 4.3.4 ▸ Block Graph

Let G be a graph. Define $\mathcal{A}$ to be the set of cut vertices of G and $\mathcal{B}$ to be the set of blocks in G. A **block graph** on G is a bipartite graph with bipartition $(\mathcal{A}, \mathcal{B})$ such that a cut vertex $a \in \mathcal{A}$ is adjacent to $B \in \mathcal{B}$ if $a \in V(B)$.

Block graph might be a bit abstract to visualise. To better understand its structure, let us consider the following result:

Proposition 4.3.5 ▸ A Block Graph on A Connected Graph Is A Tree

Let T be a block graph of a connected graph G, then T is a tree.

Proof. Take any $B_1, B_n \in \mathcal{B}$. Note that there exists a sequence $(B_1, B_2, \cdots, B_{n-1}, B_n)$ of blocks in G such that B_i and B_{i+1} share a common vertex. Denote this vertex as v_i, then clearly, v_i is adjacent to B_i and B_{i+1} in T. Therefore, $B_1 v_1 B_2 v_2 \cdots B_{n-1} v_{n-1} B_n$ is a B_1-B_n walk in T, and so B_1 and B_n are connected. Take any $a \in \mathcal{A}$ and any $B \in \mathcal{B}$. Note that there exists some $B' \in \mathcal{B}$ such that $a \sim B'$. Since B and B' are connected, a and B are connected. Similarly, one can check that any $a, a' \in \mathcal{A}$ are connected in T. Therefore, T is connected.

We will proceed to proving that T is acyclic by contradiction. Suppose on contrary that T contains a cycle. Without loss of generality, we can re-label the vertices such that the cycle is in the form of $a_1 B_1 a_2 B_2 \cdots a_n B_n a_{n+1}$ with $a_1 = a_{n+1}$. This means that for any integer $k \in [1, n]$, $a_k, a_{k+1} \in V(B_k)$ and that $V(B_{k-1}) \cap V(B_k) = a_k$ whenever $k \geq 2$. Consider $B = \bigcup_{i=1}^{n} B_i$. Notice that B is connected and does not have any cut vertex, so B is a block in G. However, for any of the B_i's, $B_i \subset B$, which contradicts the maximality of B_i. Therefore, T is acyclic, which implies that T is a tree. $\qquad\square$

To summarise, we see that for any connected graph G, its block graph is a tree such that every $B \in \mathcal{B}$ is either an edge or a 2-connected graph.

We introduce a useful definition which will help us investigate the structure of 2-connected graphs later.

Definition 4.3.6 ▶ H-path

Let G be a graph with a subgraph H. An **H-path** is a path P in G such that its 2 ends are the only vertices in $P \cap H$.

Recall that in Proposition 4.1.4, we have shown that every 2-connected graph contains a cycle, but that the converse is not true. However, can we try to extend from this necessary condition somehow to give a more precise characterisation of a 2-connected graph?

Theorem 4.3.7 ▶ Ear-Decomposition

A graph G is 2-connected if and only if it can be constructed from a cycle by successively augmenting the partially constructed graph H with H-paths.

Proof. Suppose that G is constructed from a cycle by successively augmenting H-paths to a partially constructed graph H. Note that a cycle is 2-connected. Assume that there exists some $k \in \mathbb{N}$ such that after k iterations of such augmentation, the resulted graph H is 2-connected. Consider $H' := H \cup P_H$ where P_H is any H-path, the clearly H' is also 2-connected. Therefore, G must be 2-connected.

Suppose conversely that G is 2-connected, then by Proposition 4.1.4 it contains a cycle C_n. Therefore, G contains a maximal subgraph H which can be constructed by augmenting H-paths successively to a partially constructed graph starting from C_n. Note that it suffices to show that $G = H$.

We claim that H must be an induced subgraph of G. Suppose on contrary that

there exists $u, v \in V(H)$ such that $uv \in E(G)$ but $uv \notin E(H)$, then clearly uv itself is an H-path in G. Therefore, $H + uv$ is a larger graph constructible from C_n by successively augmenting H-paths to the partially constructed graph, which is a contradiction because H is assumed to be maximal.

Suppose on contrary that $G \neq H$, then $V(G) - V(H) \neq \varnothing$. Note that G is connected, so there exist $v \in V(G) - V(H)$ and $w \in V(H)$ such that $vw \in E(G)$. Note that $C_n \subseteq H$, so there exists some $u \in V(H)$ with $u \neq w$. Since G is 2-connected, we know that $G - w$ is connected. This implies that there is a v-u path in G which does not contain w. Without loss of generality, let u' be the first vertex in the path which is in $V(H)$, the we have found a v-u' path which is an H-path. This means we can construct a larger graph H' by augmenting the H-path to H, which contradicts the maximality of H. $\qquad\square$

An H-path is vividly termed as an "ear" to the graph. Clearly, Theorem 4.3.7 shows that any 2-connected graph can be decomposed into a cycle and a series of H-paths. This is thus also known as the "ear-decomposition" of G.

We would like to take one step further and probe into 3-connected graphs. We will state the following theorem by Tutte without proof:

Theorem 4.3.8 ▶ Tutte's Characterisation of 3-connected Graphs

A graph G is 3-connected if and only if there is a sequence $(G_0, G_1, \cdots, G_n)$ with $G_n \cong G$ such that $G_0 = K_4$ and G_{i+1} contains an edge xy with $d_{G_{i+1}}(x), d_{G_{i+1}}(y) \geq 3$ and $G_i = G_{i+1} \circ xy$.

Ear-decomposition is a fairly nice construction for a 2-connected graph. However, as a characterisation it is not very intuitive, in a sense that one still needs to work towards the graph from a cycle to actually verify 2-connectivity. Can we do better and propose a more direct characterisation?

Theorem 4.3.9 ▶ Whitney's Characterisation of 2-connected Graphs

A graph G is 2-connected if and only if for every $u, v \in V(G)$, there exist 2 internally disjoint u-v paths in G.

Here, "internally disjoint" paths means that the paths share no other vertices except u and v. More generally, we define the following:

Definition 4.3.10 ▶ A-B Path

Let G be a graph with $A, B \subseteq V(G)$. An A-B **path** is a path in G which contains exactly one vertex in A and one vertex in B.

Note that in the above definition, we do not require A and B to be disjoint. Clearly, if $A \cap B \neq \emptyset$, then every vertex in $A \cap B$ is a trivial A-B path. Intuitively, to disconnect G, we necessarily need to disconnect A and B, i.e., we need to make sure that there is no A-B path in G.

Definition 4.3.11 ▶ Separating Set for (A, B)

Let G be a graph with $A, B \subseteq V(G)$. A **separating set** for (A, B) is a set X consisting of both edges and vertices such that there is no A-B path in $G - X$.

Notice that we have not given the proof for Theorem 4.3.9 yet. One may choose to use ear-decomposition to prove the theorem, but we can also see it as a corollary of another theorem. However, before proving the actual theorem, we first consider the following lemma:

Lemma 4.3.12 ▶ A-B Paths Extension

Let G be a graph with $A, B \subseteq V(G)$ such that the minimum separating set for (A, B) has size k, then for any set P of less than k vertex-disjoint A-B paths in G, there is a set Q of vertex-disjoint A-B paths in G whose set of end vertices contains the set of end vertices in P as a proper subset.

Proof. Fix any $A \subseteq V(G)$. We will perform induction on $|V(G)| - |B|$. Note that when $|V(G)| - |B| = 0$, we have $B = V(G)$ and so $k = |A|$. If $|P| < k$, then there is at least one vertex $a \in A$ which is not contained by any path in P. Take $Q := P \cup \{a\}$ then Q is a set of $|P| + 1$ vertex disjoint A-B paths in G whose set of end vertices contains the end vertices in P as a proper subset.

Take any $A, B \subseteq V(G)$ and fix $P := \{P_1, P_2, \cdots, P_n\}$ to be a set of $n < k$ vertex-disjoint A-B paths. We denote the set of end vertices in P by $P_A := \{a_1, a_2, \cdots, a_n\}$ and $P_B := \{b_1, b_2, \cdots, b_n\}$. Clearly, P_B is not a separating set for (A, B), so there is some A-B path which does not contain any vertex from P_B. Let this path be R with end vertex $b_R \in B$. If R is vertex-disjoint to every $P_i \in P$, then we are done.

Otherwise, $R \cap P_j \neq \emptyset$ for some $P_j \in P$. Without loss of generality, let $x \in V(R)$ be the vertex such that the x-b_R path in R contains no vertex from P, i.e., x is the last intersection between R and the paths in P. Without loss of generality, we

can re-label the paths in P such that $x \in P_n$. Denote by xR the x-b_R path in R and by xP_n the x-b_n path in P_n. Define $B' := B \cup V(xR) \cup V(xP_n)$, then $B \subsetneq B'$ and so $|V(G)| - |B'| < |V(G)| - |B|$. Let $P' := \{P_1, P_2, \cdots, P_{n-1}, P_n x\}$, then P' is a set of n vertex-disjoint A-B' paths in G. Now, take any X to be a minimum separating set for (A, B'), then clearly X separates (A, B), so $|X| \geq k > n$. By the inductive hypothesis, we can find a set Q' of vertex-disjoint A-B' paths in G such that the set of end vertices of Q' contains the end vertices of P as a proper subset.

Let $P'_{B'} := \{b_1, b_2, \cdots, b_{n-1}, x\}$ be the set of end vertices in B' for paths in P'. Let $Q' := \{Q'_1, Q'_2, \cdots, Q'_{n+1}\}$ and denote the end vertex of Q'_i in B by $B(Q'_i)$, then without loss of generality, we can re-label the vertices such that

$$B(Q'_i) = \begin{cases} b_i & \text{if } i = 1, 2, \cdots, n-1 \\ x & \text{if } i = n \\ y & \text{if } i = n+1 \end{cases} ,$$

where $y \notin P'_{B'}$. Note that $Q'_1, Q'_2, \cdots, Q'_{n-1}$ are already vertex-disjoint A-B paths. We will try to extend Q' to obtain Q. We consider 2 cases on y.

If $y \in xP_n$, then every vertex in $V(xR) - \{x\}$ is unused by all paths in Q', and so we can extend Q'_n by xR to obtain an A-B path, and extend Q'_{n+1} by yP_n to obtain another A-B path. Note that xR and yP_n are vertex-disjoint, so we have obtained a set Q of $(n+1)$ vertex-disjoint A-B paths as desired.

If $y \notin xP_n$, then we can extend Q'_n by xP_n to connect b_n. Note that either $y \in yR$ or $y \in B$, but either way we can extend Q'_{n+1} to a new vertex in B, so we still can obtain Q as desired. $\qquad\square$

Now we are ready to state the following theorem:

Theorem 4.3.13 ▶ Menger's Theorem

Let G be a graph with $A, B \subseteq V(G)$, then the minimum size of separating sets for (A, B) equals the maximum number of vertex-disjoint A-B paths in G.

Proof. Let X to be a minimum separating set for (A, B) and k to be the number of vertex-disjoint A-B paths in G. Clearly, to disconnect A and B, we need to at least delete either one edge or one vertex from each of the k A-B paths. Therefore, $|X| \geq k$.

Take any $n < |X|$ disjoint A-B paths, then by Lemma 4.3.12, we can extend this set of paths by another vertex-disjoint A-B paths until $n = |X|$. Therefore, $k \geq |X|$, which implies that $k = |X|$ as desired. $\qquad\square$

5

Graph Colouring

5.1 Vertex Colouring

Theorem 5.1.1 ▶ Brooks' Theorem

Let G be a connected graph not isomorphic to a complete graph or an odd cycle, then $\chi(G) \leq \Delta(G)$.

Definition 5.1.2 ▶ Edge Contraction

Let G be a graph. For any $e = xy \in E(G)$, the **edge contraction** of G at e, denoted by G/e, is a graph produced by deleting e and merging x and y such that $N_{G/e}(xy) = N_G(x) \cup N_G(y)$.

Definition 5.1.3 ▶ k-constructible Graph

A k-**constructible graph** G for some $k \in \mathbb{Z}^+$ is such that

1. $G \cong K_k$;
2. there is some k-constructible graph H with $x, y \in V(H)$ and $xy \in E(H)$ such that $G \cong (H + xy)/xy$;
3. there are k-constructible graphs H_1, H_2 with $V(H_1) \cap V(H_2) = \{x\}$ and $y_1 x \in E(H_1), y_2 x \in E(H_2)$ such that $G \cong H_1 \cup H_2 - y_1 x - y_2 x + y_1 y_2$.