Chapter 1: Introduction to Graph Theory

The Königsberg Bridge Problem

The city of Königsberg (now Kaliningrad, Russia) is divided by a river, creating two large islands and two mainland portions, all connected by seven bridges as shown in the next figure.

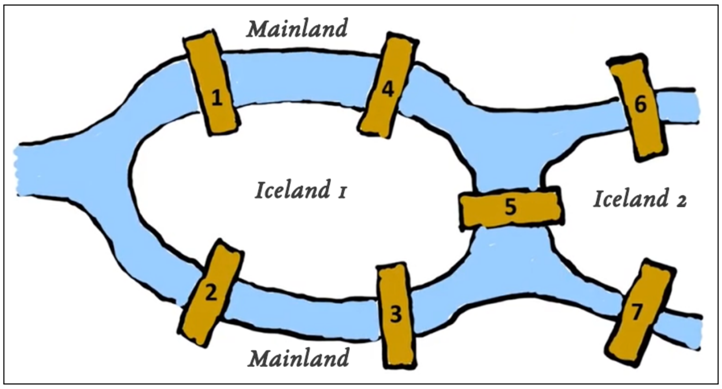


Figure 1: Seven Bridges of Königsberg

The problem asks if it's possible to walk through the city and cross each bridge exactly once, starting and ending at the same point. We can represent the problem using the following structure,

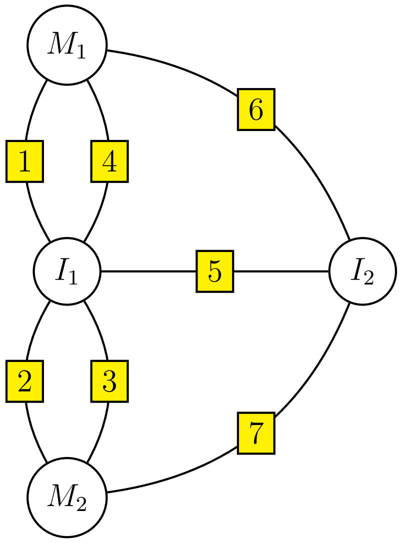


Figure 2: Graph representation of the Königsberg bridges

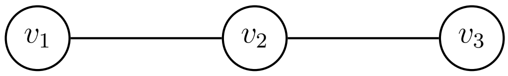
where the land areas are represented as points and the bridges as lines. This structure is known as **a graph**.

1. What's a graph?

Definition 1.1.
A graph $G = (V, E)$ is defined by two finite sets: V and E, where $V = \{v_1, v_2, \dots, v_n\}$ is the set of vertices (with each v_i representing a vertex), and $E = \{e_1, e_2, \dots, e_m\}$ is the set of edges, which are the connections between vertices. The vertices are typically represented as points, and the edges as lines connecting the corresponding vertices.

Remark.
Algebraically, a graph represents a binary relation E between the elements of a finite set V of vertices. The relation $R \subseteq V \times V$ consists of ordered pairs (v_i, v_j), indicating an edge between vertices v_i and v_j.

Example.
Consider a graph with vertices $V = \{v_1, v_2, v_3\}$ and edges $E = \{(v_1, v_2), (v_2, v_3)\}$. The corresponding graph is:



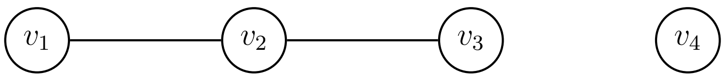
Definition 1.2 (Adjacent, Incident, and Isolated)**.**

- Two vertices v_i and v_j in a graph are said to be ***adjacent*** if there exists an edge that connects v_i and v_j.
- Similarly, two edges e_i and e_j are said to be ***adjacent*** if they share a common vertex.
- An edge is said to be ***incident*** to a vertex if the vertex is one of the endpoints of the edge.
- A vertex v_i is said to be ***isolated*** if it is not incident to any edge in the graph.

Example.
Consider the following graph $G = (V, E)$, where the set of vertices is $V = \{v_1, v_2, v_3, v_4\}$ and the set of edges is $E = \{(v_1, v_2), (v_2, v_3)\}$.

- The vertices v_1 and v_2 are ***adjacent*** because there is an edge (v_1, v_2) connecting them.
- The edges (v_1, v_2) and (v_2, v_3) are ***adjacent*** because they share the common vertex v_2.
- The edge (v_1, v_2) is ***incident*** to the vertices v_1 and v_2.
- The vertex v_4 is ***isolated*** because it is not incident to any edge.

The graph is illustrated below:



Definition 1.3 (Order and Size of a Graph)**.**
The **order** of a graph $G = (V, E)$, denoted $\operatorname{ord}(G)$, is the number of vertices in the graph. That is,

$$\operatorname{ord}(G) = |V|.$$

The **size** of a graph $G = (V, E)$, denoted $e(G)$, is the number of edges in the graph. That is,

$$e(G) = |E|.$$

Example.
In the previous example, $ord(G) = 4$ and $e(G) = 2$.

Definition 1.4 (Degree of a Vertex)**.**
Let $G = (V, E)$ be a graph and $v \in V$ be a vertex.
The ***degree*** of a vertex v, denoted $\deg(v)$, is the number of edges incident to v. Formally, the degree of v is given by

$$\deg(v) = |\{e \in E : v \text{ is an endpoint of } e\}|.$$

- The ***maximum degree*** of the graph, denoted $\Delta(G)$, is the highest degree of any vertex in the graph:

$$\Delta(G) = \max_{v \in V} \deg(v).$$

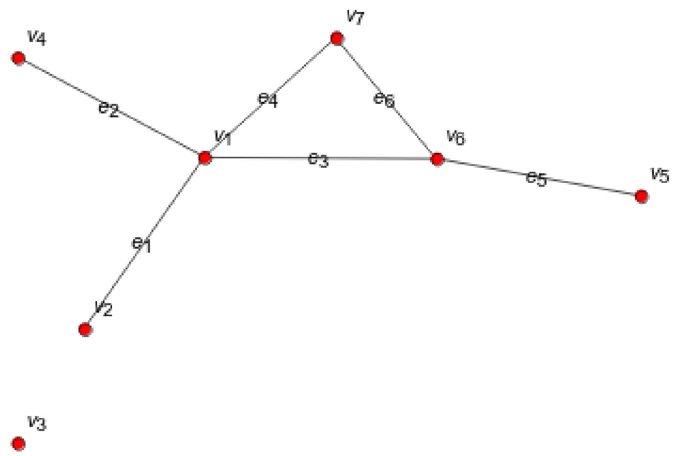
- The ***minimum degree*** of the graph, denoted $\delta(G)$, is the lowest degree of any vertex in the

graph:
$$\delta(G) = \min_{v \in V} \deg(v).$$

Definition 1.5 (Neighbor and Neighborhood).

- A vertex u is said to be a **neighbor** of vertex v if they are adjacent, i.e., there exists an edge $e = \{u, v\} \in E$.
- The **neighborhood** of v, denoted $N(v)$, is the set of vertices adjacent to v. Formally:
$$N(v) = \{u \in V : \exists e \in E \ (e = \{u, v\})\}$$
  - $N(v)$ is called the **open neighborhood** of v.
  - $N[v] = N(v) \cup \{v\}$ is called the **closed neighborhood** of v.

Example.
Consider the graph $G = (V, E)$, where
$$V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$$
and
$$E = \{e_1 = (v_1, v_2), e_2 = (v_1, v_4), e_3 = (v_1, v_6), e_4 = (v_1, v_7), e_5 = (v_5, v_6), e_6 = (v_6, v_7)\}.$$
The corresponding graph is illustrated below:



The degrees of the vertices are:

- $\deg(v_1) = 4$
- $\deg(v_2) = 1$
- $\deg(v_3) = 0$
- $\deg(v_7) = 2$

Thus, the maximum degree is $\Delta(G) = 4$ and the minimum degree is $\delta(G) = 0$.

Definition 1.6 (Loop).
A ***loop*** in a graph is an edge that connects a vertex to itself.

Definition 1.7 (Simple Graph).
A graph is called a ***simple graph*** if it does not contain any loops or multiple edges between the same pair of vertices.

Definition 1.8 (Empty Graph).
An ***empty graph*** is a graph with no edges. It can have any number of vertices, but there are no edges connecting them. In other words, $E(G) = \emptyset$.

Definition 1.9 (Complete Graph).
A ***complete graph*** is a simple graph in which every pair of distinct vertices is connected by a unique edge. A complete graph on n vertices is denoted by K_n.

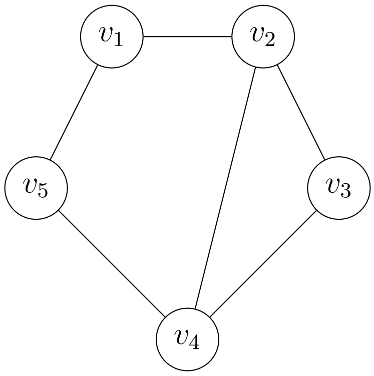
2. Directed and Undirected Graphs

Definition 2.1 (Undirected Graph).
An ***undirected graph*** is a graph in which the edges have no direction. That is, if there is an edge between vertices v_i and v_j, it can be traversed in both directions, and we denote it by $\{v_i, v_j\}$.

Example.
Consider an undirected graph $G = (V, E)$, where:

$$V = \{v_1, v_2, v_3, v_4, v_5\}, \quad E = \{\{v_1, v_2\}, \{v_2, v_3\}, \{v_2, v_4\}, \{v_3, v_4\}, \{v_4, v_5\}, \{v_5, v_1\}\}.$$

Here, the edges connect the vertices, and each edge can be traversed in both directions.



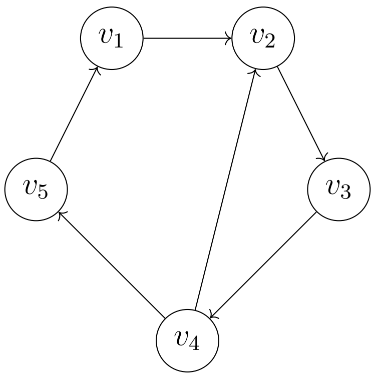
—

Definition 2.2 (Directed Graph).
A ***directed graph*** (or ***digraph***) is a graph in which the edges have a direction. Each edge is represented as an ordered pair of vertices, indicating a directed edge from one vertex to another. If there is an edge from v_i to v_j, we write it as (v_i, v_j).

Example.
Consider a directed graph $G = (V, E)$, where:

$$V = \{v_1, v_2, v_3, v_4, v_5\}, \quad E = \{(v_1, v_2), (v_2, v_3), (v_3, v_4), (v_4, v_2), (v_4, v_5), (v_5, v_1)\}.$$

Here, each edge can only be traversed in the specified direction.



Note: In this chapter, we will focus only on *undirected graphs*.

Lemma 2.3 (Handshaking Lemma).

For any graph $G = (V, E)$, the sum of the degrees of all vertices is twice the number of edges, i. e.,

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

Proof. Each edge contributes 2 to the sum of degrees (1 for each of its two endpoints). Therefore, the sum of degrees counts each edge exactly twice. Hence

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

□

Theorem 2.4.
In any graph, the number of vertices with odd degree is even.

Proof. By the Handshaking Lemma, $\sum_{v \in V} \deg(v) = 2|E|$, which is even. Split the sum into vertices with odd degrees (V_{odd}) and even degrees (V_{even}):

$$\sum_{v \in V} \deg(v) = \sum_{v \in V_{\text{odd}}} \deg(v) + \sum_{v \in V_{\text{even}}} \deg(v).$$

Since $\sum_{v \in V_{\text{even}}} \deg(v)$ is even, $\sum_{v \in V_{\text{odd}}} \deg(v)$ must also be even. The sum of an odd number of odd integers is odd, so $|V_{\text{odd}}|$ must be even. □

3. Walk, Trail, Path, Circuit and Cycle

Definition 3.1 (Walk, Trail, Path, Circuit, Cycle)**.**
Let $G = (V, E)$ be a graph.

- A ***walk*** is a sequence of vertices $v_0, v_1, v_2, \ldots, v_k$ such that $\{v_i, v_{i+1}\} \in E$ for all $0 \leq i < k$. The length of the walk is the number of edges, k.
- A ***trail*** is a walk in which no edge is repeated.
- A ***path*** is a walk in which no vertex is repeated (and thus no edge is repeated).
- A ***circuit*** is a closed walk (i.e., $v_0 = v_k$) in which no edge is repeated.
- A ***cycle*** is a closed walk (i.e., $v_0 = v_k$) in which no vertex is repeated except for the starting and ending vertex, and no edge is repeated.

Definition 3.2 (Length of a Walk, Path, or Trail)**.**
The length of a walk, path, or trail is defined as the total number of edges it contains, with repeated edges counted each time they appear in the sequence.

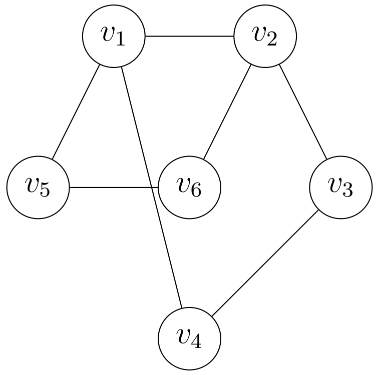
3.1 Key Differences

Term	Repeated Vertices	Repeated Edges	Closed ($v_0 = v_k$)
Walk	Allowed	Allowed	Not Required
Trail	Allowed	Not Allowed	Not Required
Path	Not Allowed	Not Allowed	Not Required
Circuit	Allowed	Not Allowed	Required
Cycle	Not Allowed (except start/end)	Not Allowed	Required

Table 1: Comparison of Walk, Trail, Path, Circuit, and Cycle

Example.
Consider the graph $G = (V, E)$, where:

$$V = \{v_1, v_2, v_3, v_4, v_5, v_6\}, \quad E = \{\{v_1, v_2\}, \{v_2, v_3\}, \{v_3, v_4\}, \{v_4, v_1\}, \{v_1, v_5\}, \{v_5, v_6\}, \{v_6, v_2\}\}.$$



- **Walk**: $v_1, v_2, v_3, v_4, v_1, v_5, v_6, v_2$ (vertices and edges repeat).
- **Trail**: $v_1, v_2, v_3, v_4, v_1, v_5, v_6$ (no repeated edges).
- **Path**: $v_1, v_5, v_6, v_2, v_3, v_4$ (no repeated vertices or edges).
- **Circuit**: v_1, v_2, v_6, v_5, v_1 (closed trail).
- **Cycle**: v_1, v_2, v_3, v_4, v_1 (closed path).

Remarks.

- The following inclusions hold:
$$\text{Paths} \subset \text{Trails} \subset \text{Walks}.$$
That is, all paths are trails, and all trails are walks.
- The length of a cycle must be at least 3. That is, a cycle must contain 3 or more edges.

Exercise.
Show that in any simple graph with n vertices, there are at most $\frac{n(n-1)}{2}$ edges.

Solution (Combinatorial Proof)**.**
Let $G = (V, E)$ be a simple graph with n vertices. Since G is simple, it has no loops or multiple edges. The graph has the most edges when every pair of vertices is connected by exactly one edge. This is equivalent to the number of ways to choose 2 distinct vertices from n, which is:
$$|E| \leq \binom{n}{2} = \frac{n(n-1)}{2}.$$
This result can also be proved by induction on n.

Exercise.
Determine the number of distinct simple graphs that can be formed with n vertices.

Solution.
A simple graph with n vertices is determined by its set of edges. The maximum number of edges in a simple graph with n vertices is $\binom{n}{2}$, since each edge is uniquely determined by a pair of distinct vertices.
For each pair of vertices, there are two choices:

- Include the edge between them.
- Exclude the edge between them.

Since there are $\binom{n}{2}$ pairs of vertices, the total number of distinct simple graphs with n vertices is:
$$2^{\binom{n}{2}}.$$

Lemma 3.3.
In any graph with at least two vertices, there are at least two vertices with the same degree.

Proof. In a graph with $n \geq 2$ vertices, the possible degrees range from 0 to $n-1$. However, a graph cannot simultaneously have a vertex of degree 0 (isolated) and a vertex of degree $n-1$ (connected to all others). Therefore, there are at most $n-1$ distinct possible degrees for n vertices. By the **Pigeonhole Principle**, at least two vertices must have the same degree. (See Worksheet 2, Exercise 6.) □

Theorem 3.4.
Let $G = (V, E)$ be a graph, and let $u, v \in V$ be two distinct vertices. Any u-v walk in G contains a u-v path in G.

Proof. Let $W = u, v_1, \ldots, v_{k-1}, v$ be a u-v walk. If W has no repeated vertices, it is already a u-v path. Otherwise:

- Let $v_i = v_j$ be the first repeated vertices in W with $i < j$.
- Remove the subsequence $v_{i+1}, \ldots, v_j$ to obtain a shorter walk W'.
- Repeat this process until no repeated vertices remain. The result is a u-v path.

Thus, any u-v walk contains a u-v path. □

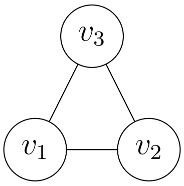
4. Connectivity, Subgraphs and Planarity

4.1 Connected Graph

Definition 4.1 (Connected Graph).
A graph $G = (V, E)$ is said to be ***connected*** if there exists a path between every pair of distinct vertices in G. Otherwise, the graph is ***disconnected***.

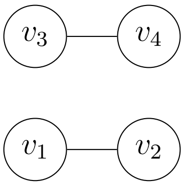
Example.

- A **connected graph**:



  Every pair of vertices is connected by a path.

- A **disconnected graph**:



  No path exists between v_1 and v_3, or between v_2 and v_4.

Remark.
Connected vertices and **adjacent vertices** are distinct concepts in graph theory:

- Two vertices are ***adjacent*** if there is an edge directly connecting them.
- Two vertices are ***connected*** if there exists a path (a sequence of edges) between them, even if they are not directly adjacent.

For example, in a path graph v_1—v_2—v_3, v_1 and v_3 are **connected** but not **adjacent**.

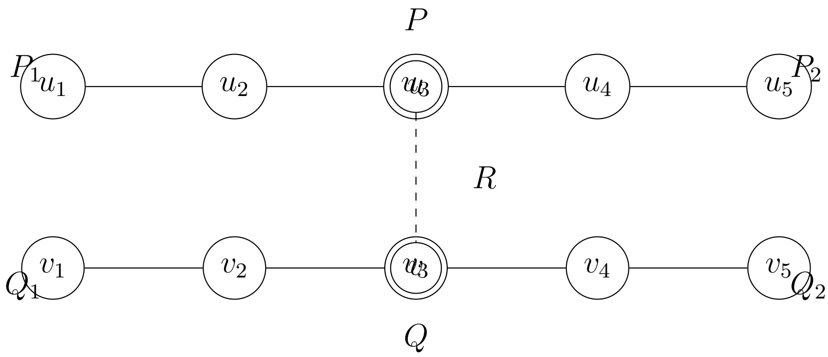
Exercise.
Prove that if P and Q are two longest paths in a connected graph, then P and Q share at least one common vertex.

Solution.
Assume, for contradiction, that P and Q are two longest paths in a connected graph G with no

common vertices. Since G is connected, there exists a path R connecting a vertex u on P to a vertex v on Q, where no interior vertex of R lies on P or Q.
Divide P into subpaths P_1 and P_2 at u, and Q into subpaths Q_1 and Q_2 at v. Without loss of generality, assume $|P_1| \geq |P_2|$ and $|Q_1| \geq |Q_2|$.
Construct a new path $P_1 \cup R \cup Q_1$. Its length is:

$$|P_1| + |R| + |Q_1| > |P|,$$

since $|P_1| \geq \frac{|P|}{2}$ and $|Q_1| \geq \frac{|Q|}{2}$. This contradicts the assumption that P and Q are the longest paths. Hence, P and Q must share at least one common vertex.



Definition 4.2 (Complement of a Graph).
Let $G = (V, E)$ be a simple graph. The ***complement*** of G, denoted $\overline{G}$, is the simple graph with the same vertex set V such that for any two distinct vertices $u, v \in V$,

$$\{u, v\} \in E(\overline{G}) \quad \text{if and only if} \quad \{u, v\} \notin E(G).$$

In other words, two vertices are adjacent in $\overline{G}$ if and only if they are not adjacent in G.
The complement of the complement of G is the original graph:

$$\overline{\overline{G}} = G.$$

Exercise.
Let G be a simple graph with n vertices, and let $\overline{G}$ be its complement. Show that:

$$|E(G)| + |E(\overline{G})| = \binom{n}{2}.$$

Solution.
In a simple graph with n vertices, the total number of possible edges is $\binom{n}{2}$. The complement $\overline{G}$ contains all edges not in G, so:

$$|E(\overline{G})| = \binom{n}{2} - |E(G)|.$$

Adding the edges of G and $\overline{G}$:

$$|E(G)| + |E(\overline{G})| = |E(G)| + \left(\binom{n}{2} - |E(G)|\right) = \binom{n}{2}.$$

4.2 Subgraphs and component of a Graph

Definition 4.3 (Subgraph).
A graph $H = (V', E')$ is a ***subgraph*** of $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq E$, and every edge in E' has both endpoints in V'.

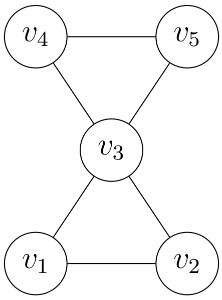
Definition 4.4 (Spanning Subgraph).
A ***spanning subgraph*** of $G = (V, E)$ is a subgraph $H = (V', E')$ where $V' = V$ and $E' \subseteq E$.

Definition 4.5 (Induced Subgraph).
An ***induced subgraph*** of $G = (V, E)$ is a subgraph $H = (V', E')$ where $V' \subseteq V$ and E' contains all edges in E that have both endpoints in V'.

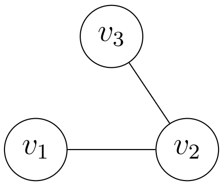
Definition 4.6 (Edge-Induced Subgraph).
An ***edge-induced subgraph*** of $G = (V, E)$ is a subgraph $H = (V', E')$ where $E' \subseteq E$ and V' is the set of vertices incident to edges in E'.

—

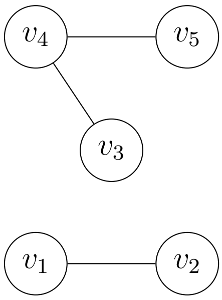
Examples.
Consider the graph G:



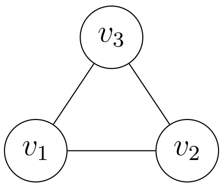
- **Subgraph**: A subgraph H of G could be:



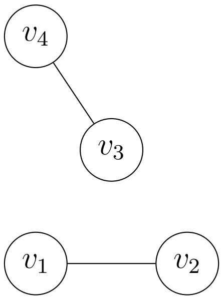
  Here, $V' = \{v_1, v_2, v_3\}$ and $E' = \{\{v_1, v_2\}, \{v_2, v_3\}\}$.
- **Spanning Subgraph**: A spanning subgraph H of G could be:



  Here, $V' = V$ and $E' = \{\{v_1, v_2\}, \{v_3, v_4\}, \{v_4, v_5\}\}$.
- **Induced Subgraph**: An induced subgraph H of G with $V' = \{v_1, v_2, v_3\}$ is:



  Here, $E' = \{\{v_1, v_2\}, \{v_2, v_3\}, \{v_3, v_1\}\}$, which includes all edges of G with both endpoints in V'.
- **Edge-Induced Subgraph**: An edge-induced subgraph H of G with $E' = \{\{v_1, v_2\}, \{v_3, v_4\}\}$ is:



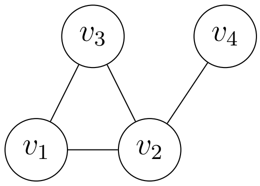
Here, $V' = \{v_1, v_2, v_3, v_4\}$, which are the vertices incident to edges in E'.

Definition 4.7 (Maximal Subgraph).
A subgraph $H = (V', E')$ of $G = (V, E)$ is called a ***maximal subgraph*** with respect to a property P if:

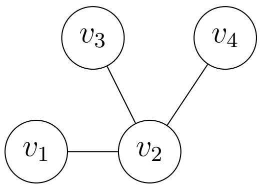
- H satisfies P, and
- No larger subgraph of G containing H satisfies P.

In other words, H is maximal if adding any additional vertex or edge from G to H would cause it to no longer satisfy P.

Example.
Consider the graph G:



- A **maximal acyclic subgraph** of G is:



This subgraph is acyclic, and adding any additional edge from G (e.g., $\{v_1, v_3\}$) would create a cycle, violating the acyclic property.

Exercise.
Let G be a graph with n vertices and m edges. How many induced subgraphs and how many spanning subgraphs does G have?

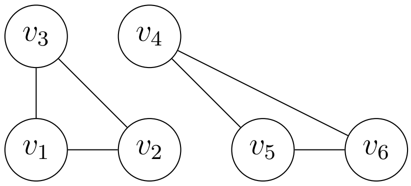
Solution.

- **Induced Subgraphs**: There are 2^n induced subgraphs, as each subset of vertices defines a unique induced subgraph.
- **Spanning Subgraphs**: There are 2^m spanning subgraphs, as each subset of edges defines a unique spanning subgraph.

Definition 4.8 (Component).
A ***component*** of a graph $G = (V, E)$ is a maximal connected subgraph of G.

Remark.
A graph is ***connected*** if it has exactly one component. Otherwise, it is ***disconnected***.

Example.
Consider the graph G:



This graph has two components:

- One component contains v_1, v_2, v_3, forming a triangle.
- The other component contains v_4, v_5, v_6, forming another triangle.

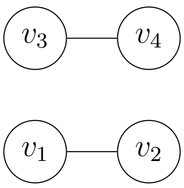
Each component is a maximal connected subgraph because:

- No additional vertices or edges can be added to either component without disconnecting it.
- The graph G is ***disconnected*** because it has more than one component.

4.3 Cliques and Stable Sets

Definition 4.9 (Stable Set).
A **stable set** (or **independent set**) in a graph $G = (V, E)$ is a subset of vertices $S \subseteq V$ such that no two vertices in S are adjacent. A **maximal stable set** is a stable set that cannot be extended by adding another vertex. A **maximum stable set** is a stable set of the largest possible size. The size of a maximum stable set is denoted by $\alpha(G)$, the **stability number** of G.

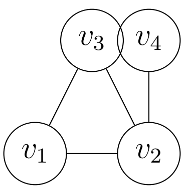
Example.
Consider the graph G:



- The sets $\{v_1, v_3\}$, $\{v_1, v_4\}$, $\{v_2, v_3\}$, and $\{v_2, v_4\}$ are maximal stable sets.
- The maximum stable set has size $\alpha(G) = 2$.

Definition 4.10 (Clique).
A **clique** in a graph $G = (V, E)$ is a subset of vertices $C \subseteq V$ such that every two distinct vertices in C are adjacent. A **maximal clique** is a clique that cannot be extended by adding another vertex. A **maximum clique** is a clique of the largest possible size. The size of a maximum clique is denoted by $\omega(G)$, the **clique number** of G.

Example.
Consider the graph G:



- The sets $\{v_1, v_2, v_3\}$ and $\{v_2, v_4\}$ are maximal cliques.
- The maximum clique has size $\omega(G) = 3$.

4.4 Trees

Definition 4.11 (Tree).
A **tree** is a connected acyclic graph.

Definition 4.12 (Leaf).
In a tree, a *leaf* is a vertex of degree 1.

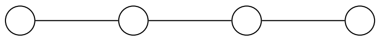
Definition 4.13.
A *spanning tree* of a connected graph G is a spanning subgraph of G that is a tree.

Definition 4.14.
A *forest* is a graph where each connected component is a tree. Equivalently, a forest is a graph with no cycles.

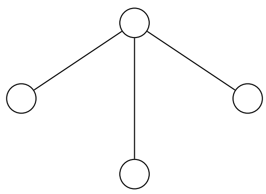
Example.
Here are some examples of trees with drawings:

- A single vertex (trivial tree):

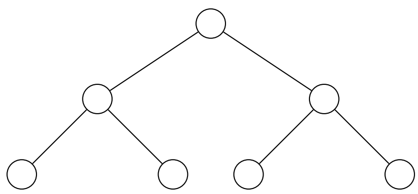
- A path graph P_4 with 4 vertices:



- A star graph S_4 with one central vertex connected to 3 leaves:



- A binary tree of height 2:



Theorem 4.15.
A graph G is a tree if and only if every pair of vertices is connected by a unique path.

Proof.

- ($\Longrightarrow$) If G is a tree, it is connected and acyclic. By definition, there is exactly one path between any two vertices.
- ($\Longleftarrow$) If every pair of vertices in G has a unique path, then G is connected. If G had a cycle, there would be two distinct paths between some pair of vertices, which contradicts the assumption. Thus, G is acyclic and hence a tree.

□

Theorem 4.16.
A graph G is a tree if and only if it has $n-1$ edges.

Proof.

- ($\implies$) If G is a tree, it has $n-1$ edges. This can be proved by induction on n.
- ($\impliedby$) If G has $n-1$ edges and is connected, it is a tree. If G had a cycle, removing an edge from the cycle would leave G connected with $n-2$ edges, which contradicts the fact that a connected graph must have at least $n-1$ edges. Thus, G is acyclic and hence a tree.

□

Theorem 4.17.
Let T be a tree with n_i vertices of degree i, where $i = 1, 2, \ldots, k$ and $k = \Delta(T)$. Then:

$$n_1 = 2 + n_3 + 2n_4 + 3n_5 + \cdots + (k-2)n_k.$$

Proof.

- By the Handshaking Lemma, $\sum_{i=1}^{k} i \cdot n_i = 2(n-1)$.
- Since T is a tree, it has $n-1$ edges.
- Rearrange the equation to solve for n_1:

$$n_1 = 2 + n_3 + 2n_4 + 3n_5 + \cdots + (k-2)n_k.$$

□

Theorem 4.18.
A forest with n vertices and k trees has $n-k$ edges.

Proof. Each of the k trees in the forest is a connected component. A tree with n_i vertices has $n_i - 1$ edges. Summing over all k trees, the total number of edges is:

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

since $\sum_{i=1}^{k} n_i = n$. □