(Almost all) The Graph Theory You Need

Gudfit

Contents

0 Graph Theory 3
- 0.1 Definitions and Representations 3
- 0.2 A Bestiary of Graphs 5
- 0.3 Degrees and Regularity 6
- 0.4 Matrix Representations 8
- 0.5 Graphic Sequences 8
- 0.6 Exercises 11

1 Extremal Graph Theory 13
- 1.1 Substructures 13
- 1.2 Turán's Theorem 14
- 1.3 Geometric Applications 16
- 1.4 Independent Sets 17
- 1.5 Ramsey Theory 18
- 1.6 Exercises 20

2 Connectivity 23
- 2.1 Walks, Trails, and Paths 23
- 2.2 Distance 26
- 2.3 Eulerian Tours and Trails 29
- 2.4 Exercises 31

3 Trees 34
- 3.1 Forests and Trees 34
- 3.2 Enumeration of Trees 39
- 3.3 Prüfer Sequences 43
- 3.4 Exercises 45

4 Colouring 47
- 4.1 Vertex Colouring 47
- 4.2 Bipartite Graphs 49
- 4.3 Matchings 50
- 4.4 Exercises 54

5 Planarity 56
- 5.1 Plane Drawings 56
- 5.2 Euler's Formula 57
- 5.3 Kuratowski's Theorem 60
- 5.4 Colouring Planar Graphs 60
- 5.5 Exercises 61

0

Graph Theory

Graph theory formalises the study of pairwise connections, serving as a fundamental language for discrete mathematics.

0.1 Definitions and Representations

We define a graph as a structure consisting of points and links connecting them. Unlike the continuous functions of analysis, graphs are inherently discrete.

Definition 0.1. *Graph.*
A **graph** is a pair $G = (V, E)$, where:
- V is a finite set of elements called **vertices** (or nodes).
- E is a subset of $\mathcal{P}_2(V)$, the set of all 2-element subsets of V. The elements of E are called **edges**.

The **order** of the graph is the cardinality of the vertex set, $|V|$. Two vertices $x, y \in V$ are said to be **adjacent** (or neighbours) if $\{x, y\} \in E$.
定義

Note

Unless explicitly stated otherwise, all graphs in this text are *simple* (no loops connecting a vertex to itself, no multiple edges between the same pair) and *undirected* (edges are sets $\{x, y\}$, not ordered pairs (x, y)).

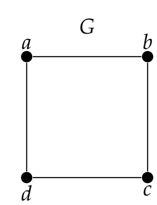


Figure 1: A representation of the graph $G = (\{a, b, c, d\}, \{\{a, b\}, \{b, c\}, \{c, d\}, \{d, a\}\})$. This is the cycle C_4.

Graphs are frequently represented geometrically by drawing vertices as points and edges as curves connecting them. It is crucial to distinguish the combinatorial object G from its visual representation; a single graph may admit multiple distinct drawings.

Example 0.1. Labelled Graphs. Consider the set of vertices $V = \{a, b, c\}$. The number of possible edges is $|\mathcal{P}_2(V)| = \binom{3}{2} = 3$. The possible edges are $\{a, b\}, \{b, c\}, \{a, c\}$. A graph on V is determined by choosing a subset of these edges. Thus, there are $2^3 = 8$ distinct graphs on these three labelled vertices.

範例

This example generalises immediately via the properties of the power set established in the previous notes

Proposition 0.1. ***Counting Labelled Graphs.***
The number of distinct graphs with a fixed vertex set V of cardinality n is

$$2^{\binom{n}{2}}.$$

命題

Proof

A graph is completely determined by its edge set $E \subseteq \mathcal{P}_2(V)$. The size of the set of all possible pairs is $|\mathcal{P}_2(V)| = \binom{n}{2}$. The number of subsets of $\mathcal{P}_2(V)$ is the cardinality of its power set, which is $2^{\binom{n}{2}}$. ■

The definition of a graph $G = (V, E)$ depends on the specific labels of the vertices. However, the structural properties of a graph (connectivity, cycles, etc.) are independent of labelling. We formalise the notion of "structural equality" via isomorphism.

Definition 0.2. ***Isomorphism.***
Two graphs $G = (V, E)$ and $G' = (V', E')$ are **isomorphic**, denoted $G \cong G'$, if there exists a bijection $\varphi : V \to V'$ such that for all $x, y \in V$:

$$\{x, y\} \in E \iff \{\varphi(x), \varphi(y)\} \in E'.$$

The map φ is called an **isomorphism**.

定義

An isomorphism renames vertices while preserving adjacency. The relation $\cong$ is an equivalence relation on the set of all graphs. When we speak of "a graph" in an abstract sense (e.g., "the triangle"), we refer to an isomorphism class.

Example 0.2. Isomorphism Classes of Order 3. While there are $2^{\binom{3}{2}} = 8$ labelled graphs on three vertices, there are only 4 up to isomorphism.

1. No edges (empty).
2. One edge.
3. Two edges (a path of length 2).
4. Three edges (a triangle).

範例

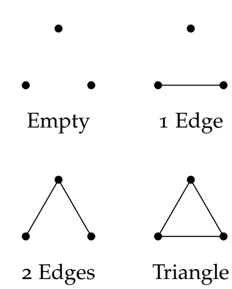


Figure 2: The four non-isomorphic graphs of order 3.

Counting graphs up to isomorphism is significantly more difficult than counting labelled graphs, as the size of isomorphism classes varies. However, we can establish useful bounds using the Pigeonhole Principle and properties of group actions.

Proposition 0.2. ***Bounds on Isomorphism Classes.***
Let $\mathcal{U}_n$ be the set of isomorphism classes of graphs of order n. Then:

$$|\mathcal{U}_n| \geq \frac{2^{\binom{n}{2}}}{n!}.$$

命题

Proof

Let $\mathcal{G}_n$ be the set of labelled graphs on the set $V = [1, n]$. We have $|\mathcal{G}_n| = 2^{\binom{n}{2}}$. Consider the mapping $\pi : \mathcal{G}_n \to \mathcal{U}_n$ that assigns each graph to its isomorphism class. For any graph G, the size of its isomorphism class is at most $n!$, since any isomorphism is determined by a permutation of the n vertices. Since the classes form a partition of $\mathcal{G}_n$, we have:

$$|\mathcal{G}_n| = \sum_{C \in \mathcal{U}_n} |C| \leq \sum_{C \in \mathcal{U}_n} n! = |\mathcal{U}_n| \cdot n!.$$

Rearranging yields the result.

■

Asymptotically, the number of unlabelled graphs behaves similarly to the number of labelled graphs, as the exponential term $2^{n^2/2}$ dominates the factorial $n!$. Using the approximation $\log_2(n!) \approx n \log_2 n$, we observe:

$$\log_2 \left(\frac{2^{\binom{n}{2}}}{n!} \right) = \binom{n}{2} - \log_2(n!) \approx \frac{n^2}{2} - n \log_2 n.$$

For large n, this growth is driven by the n^2 term.

0.2 *A Bestiary of Graphs*

We define several standard families of graphs that appear frequently in examples and counterexamples.

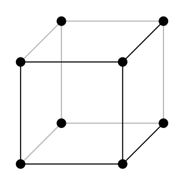


Figure 3: The 3-dimensional hypercube Q_3. Each vertex is a binary string in $\{0, 1\}^3$; edges connect strings differing in one bit.

Definition 0.3. ***Standard Graphs.***
Let $n \in \mathbb{N}^*$.

1. **Empty Graph** (S_n): $V = [1, n]$ and $E = \varnothing$. Also called the stable graph.
2. **Complete Graph** (K_n): $V = [1, n]$ and $E = \mathcal{P}_2(V)$. Every pair of distinct vertices is connected.
3. **Path Graph** (P_n): A chain of length n.

$$V = [0, n], \quad E = \{\{i, i+1\} : 0 \leq i < n\}.$$

Note that P_n has order $n+1$ and size (number of edges) n. Some authors instead index P_n by the number of vertices (order n); we follow the length convention here to match $|E| = n$.

4. **Cycle Graph** (C_n): For $n \geq 3$, a closed loop of length n.

$$V = [1, n], \quad E = \{\{i, i+1\} : 1 \leq i < n\} \cup \{\{n, 1\}\}.$$

5. **Hypercube** (Q_d): The d-dimensional cube.

$$V = \{0, 1\}^d, \quad E = \{\{x, y\} : x \text{ and } y \text{ differ in exactly one coordinate}\}.$$

定義

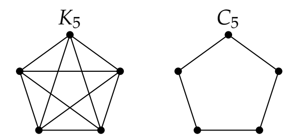


Figure 4: The complete graph K_5 and the cycle C_5.

Another crucial class of graphs arises when the vertex set can be partitioned into two disjoint sets such that edges only connect vertices from different sets.

Definition 0.4. ***Bipartite Graphs.***
The **complete bipartite graph** $K_{n,m}$ is defined by partitions of sizes n and m.

$$V = \{a_1, \dots, a_n\} \cup \{b_1, \dots, b_m\}, \quad E = \{\{a_i, b_j\} : 1 \leq i \leq n, 1 \leq j \leq m\}.$$

The order is $n + m$ and the number of edges is nm.

定義

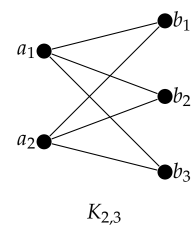


Figure 5: The complete bipartite graph $K_{2,3}$: every vertex on the left is adjacent to every vertex on the right. It has $2 \cdot 3 = 6$ edges.

Example 0.3. The Petersen Graph. The Petersen graph is a specific graph of order 10 and size 15, famous for being a counterexample to many optimistic conjectures in graph theory. It can be constructed as the Kneser graph $KG(5, 2)$:

· Vertices: The 2-element subsets of $\{1, 2, 3, 4, 5\}$.
· Edges: Two vertices (subsets) are adjacent if they are disjoint.

範例

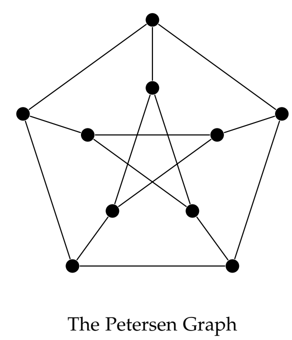


Figure 6: The Petersen graph drawn with 5-fold symmetry.

0.3 *Degrees and Regularity*

The local structure of a graph is characterised by the connectivity of individual vertices.

Definition 0.5. ***Degree.***
Let $G = (V, E)$ be a graph and $x \in V$. The **neighbourhood** of x, denoted $N(x)$, is the set of vertices adjacent to x:

$$N(x) = \{y \in V : \{x, y\} \in E\}.$$

The **degree** of x, denoted $d(x)$ (or $\deg(x)$), is the cardinality of its neighbourhood:

$$d(x) = |N(x)|.$$

A vertex of degree 0 is **isolated**; a vertex of degree 1 is a **leaf**.

定義

We characterise the global "boundedness" of the graph by its extremal degrees:

Minimum degree:

$$\delta(G) = \min_{v \in V} d(v).$$

Maximum degree:

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

If $\delta(G) = \Delta(G) = k$, every vertex has degree k, and the graph is said to be k-**regular**.

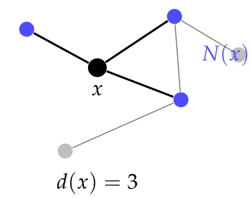


Figure 7: The neighbourhood $N(x)$ (blue vertices) consists of all vertices adjacent to x. Here $d(x) = |N(x)| = 3$.

Theorem 0.1. ***The Handshaking Lemma.***
For any graph $G = (V, E)$:

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

定理

Proof

We employ a double counting argument on the set of vertex-edge incidences. Let $S \subseteq V \times E$ be the set of pairs (v, e) such that $v \in e$:

$$S = \{(v, e) \in V \times E : v \in e\}.$$

We calculate $|S|$ in two ways:

1. Summing over vertices: For a fixed v, the number of edges containing v is $d(v)$. Thus $|S| = \sum_{v \in V} d(v)$.
2. Summing over edges: For a fixed edge $e \in E$, since G is a simple graph, e consists of exactly two distinct vertices. Thus $|S| = \sum_{e \in E} 2 = 2|E|$.

Equating the two expressions yields the result.

■

This arithmetic constraint imposes a parity restriction on the degrees.

Corollary 0.1. *Parity of Odd Degrees.* In any graph, the number of vertices with odd degree is even.

推論

Proof

Partition the vertex set V into $V_{even} = \{v : d(v) \text{ is even}\}$ and $V_{odd} = \{v : d(v) \text{ is odd}\}$. By the Handshaking Lemma:

$$2|E| = \sum_{v \in V_{even}} d(v) + \sum_{v \in V_{odd}} d(v).$$

The left-hand side is even, and the first sum on the right is a sum of even integers, hence even. Therefore, the sum $\sum_{v \in V_{odd}} d(v)$ must be even. Since this is a sum of odd integers, the number of terms $|V_{odd}|$ must be even.

■

0.4 Matrix Representations

While graphs are combinatorial objects, they can be encoded using linear algebra. Let $G = (V, E)$ with $V = \{v_1, \dots, v_n\}$ and $E = \{e_1, \dots, e_m\}$.

Definition 0.6. ***Adjacency Matrix.***
The **adjacency matrix** A_G is the $n \times n$ matrix defined by:

$$(A_G)_{ij} = \begin{cases} 1 & \text{if } \{v_i, v_j\} \in E, \\ 0 & \text{otherwise.} \end{cases}$$

For simple undirected graphs, A_G is symmetric with zeros on the diagonal. The sum of the i-th row (or column) is $d(v_i)$.

定義

Definition 0.7. ***Incidence Matrix.***
The **incidence matrix** M is the $n \times m$ matrix capturing the relationship between vertices and edges:

$$M_{ij} = \begin{cases} 1 & \text{if } v_i \in e_j, \\ 0 & \text{if } v_i \notin e_j. \end{cases}$$

定義

The incidence matrix provides an algebraic verification of the Handshaking Lemma. Summing all entries of M:

$$\sum_{i=1}^{n} \sum_{j=1}^{m} M_{ij} = \sum_{j=1}^{m} \underbrace{\left(\sum_{i=1}^{n} M_{ij} \right)}_{2} = 2m = 2|E|,$$

$$\sum_{i=1}^{n} \sum_{j=1}^{m} M_{ij} = \sum_{i=1}^{n} \underbrace{\left(\sum_{j=1}^{m} M_{ij} \right)}_{d(v_i)} = \sum_{v \in V} d(v).$$

0.5 Graphic Sequences

The **degree sequence** of a graph of order n is the tuple of vertex degrees $(d_1, \dots, d_n)$, typically sorted in non-decreasing order $d_1 \leq d_2 \leq \cdots \leq d_n$. While isomorphic graphs share the same degree

sequence, the converse is false. For example, the cycle C_6 and the disjoint union of two triangles $2K_3$ both have the degree sequence $(2, 2, 2, 2, 2, 2)$, yet C_6 is connected while $2K_3$ is not.
A sequence of integers is termed **graphic** if there exists a graph realising it as a degree sequence. The Havel-Hakimi theorem provides a recursive algorithm to test this property.

Theorem 0.2. *Havel-Hakimi.*
Let $D = (d_1, \ldots, d_n)$ be a sequence of integers satisfying $0 \leq d_1 \leq \cdots \leq d_n \leq n-1$ (the upper bound guarantees the reduction step is defined). The sequence D is graphic if and only if the reduced sequence D' of length $n-1$ is graphic, where D' is obtained by removing d_n and subtracting 1 from the d_n largest remaining terms. Formally, let $k = d_n$. The reduced sequence is formed from $(d_1, \ldots, d_{n-k-1}, d_{n-k} - 1, \ldots, d_{n-1} - 1)$ (reordered if necessary).

定理

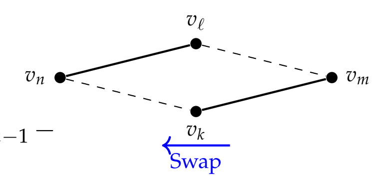


Figure 8: The edge switching argument. If v_n is connected to a low-degree vertex v_ℓ but misses a high-degree vertex v_k, and v_k has a neighbour v_m not adjacent to v_ℓ, we can swap edges $\{v_n, v_\ell\}, \{v_k, v_m\}$ for $\{v_n, v_k\}, \{v_\ell, v_m\}$ to preserve degrees while "fixing" v_n's neighbourhood.

($\Longleftarrow$)

Suppose D' is graphic. Let G' be a graph with vertex set $\{v_1, \ldots, v_{n-1}\}$ realising D'. Construct G by adding a vertex v_n and connecting it to the vertices corresponding to the d_n indices that were decremented in D'. The degrees in G are exactly D.

証明終

($\Longrightarrow$)

Suppose D is graphic. Let G be a realisation such that the vertex v_n (with degree d_n) is adjacent to the set of vertices $H = \{v_{n-d_n}, \ldots, v_{n-1}\}$, i.e., the vertices with the highest degrees among the rest. If such a G exists, removing v_n yields a graph realising D'. We show such a graph must exist. Among all graphs realising D, choose one where v_n shares the maximum number of edges with the target set H. Suppose for contradiction that $N(v_n) \neq H$. Then there exists a "missing" neighbour $v_k \in H \setminus N(v_n)$ and a "wrong" neighbour $v_\ell \in N(v_n) \setminus H$. Since $v_k \in H$ and $v_\ell \notin H$, we have $k > \ell$, implying $d(v_k) \geq d(v_\ell)$.

Case 1: *$d(v_k) = d(v_\ell)$.* We may simply swap the labels of v_k and v_ℓ. The graph structure is unchanged, but v_n is now connected to $v_k \in H$, increasing the intersection $|N(v_n) \cap H|$.

Case 2: *$d(v_k) > d(v_\ell)$.* Since v_k has strictly higher degree than v_ℓ, there must exist a vertex v_m such that $\{v_k, v_m\} \in E$ but $\{v_\ell, v_m\} \notin E$. (Note $m \neq n$). Consider the vertices $\{v_n, v_k, v_\ell, v_m\}$. We have edges $\{v_n, v_\ell\}$ and $\{v_k, v_m\}$, but $\{v_n, v_k\}$ and $\{v_\ell, v_m\}$ are absent. Perform an edge swap: remove $\{v_n, v_\ell\}$ and $\{v_k, v_m\}$; add $\{v_n, v_k\}$ and $\{v_\ell, v_m\}$. This operation preserves the degree of every vertex but connects v_n to v_k instead of v_ℓ.

In both cases, we construct a realisation with strictly greater overlap with H, contradicting the maximality assumption. Thus, a realisation connecting v_n to H exists.

証明終

Example 0.4. Determining Graphic Sequences. We determine if the following sequences are graphic.

$S_1 = (1,1,1,2,2)$. The sum of degrees is $1+1+1+2+2 = 7$, which is odd. By the Handshaking Lemma, this is impossible. Alternatively, applying Havel-Hakimi implies reducing $(1,1,1,2)$ by connecting the degree 2 vertex to the two largest: $D' = (1, 1-1, 1-1) = (1,0,0)$. A graph with one vertex of degree 1 and two of degree 0 is impossible (the degree 1 vertex needs a neighbour).

$S_2 = (1,1,1,3,4)$. Sorted: $(1,1,1,3,4)$. Remove 4, subtract 1 from four largest remaining:

$$D' = (1-1, 1-1, 1-1, 3-1) = (0,0,0,2).$$

Remove 2, subtract 1 from two largest remaining:

$$D'' = (0, 0-1, 0-1) = (0,-1,-1).$$

Negative degrees are impossible. Thus S_2 is not graphic.

$S_3 = (1,1,1,2,2,3,4,5,5)$. Remove 5: reduce $\{5,4,3,2,2\} \to \{4,3,2,1,1\}$.

$$D' = (1,1,1,1,1,2,3,4,4).$$

(Note: we retain the three 1s that were not modified). Remove 4: reduce $\{4,3,2,1\} \to \{3,2,1,0\}$.

$$D'' = (0,1,1,1,1,1,2,3).$$

Remove 3: reduce $\{2,1,1\} \to \{1,0,0\}$.

$$D''' = (0,0,0,1,1,1,1).$$

This is graphic (e.g., two disjoint edges and 3 isolated vertices). Thus S_3 is graphic.

範例

0.6 *Exercises*

1. **Degrees and Neighbours.**
   (a) Let $G_1 = (V, E)$ where $V = \{a, b, c, d, e\}$ and $E = \{\{a, c\}, \{b, d\}, \{c, d\}, \{c, e\}, \{d, e\}, \{e, e\}\}$. Compute the degree of each vertex. Is G_1 simple? List the neighbours of a and the edges incident to a.
   (b) Let $G_2 = (V, E)$ where $V = \{a, b, c\}$ and $E = \{\{a, b\}, \{a, c\}, \{a, c\}\}$. Compute degrees, check simplicity, and list neighbours/incident edges for a.
   (c) Let $G_3 = (V, E)$ where $V = \{a, b, c, d\}$ and $E = \{\{a, b\}, \{a, c\}, \{b, c\}\}$. Compute degrees, check simplicity, check for isolated vertices, and draw the graph.
2. **Basic Families.**
   (a) Draw the complete graphs K_1, K_2, K_3, K_4 and K_5.
   (b) Let G be a simple graph on vertex set V. Show that for any subset $S \subseteq V$, the induced subgraph $G[S]$ is uniquely determined by S.
   (c) Let G be a simple graph on n vertices. Prove that the edge set of the complement $\overline{G}$ satisfies $E(K_n) = E(G) \cup E(\overline{G})$ and $E(G) \cap E(\overline{G}) = \emptyset$.
3. **Handshaking and Induction.**
   (a) Prove the Handshaking Lemma ($\sum d(v) = 2|E|$) by induction on the number of edges.
   (b) Prove that the number of odd-degree vertices is even by induction on the number of edges.

   > *Remark.*
   >
   > Base case: empty graph. Inductive step: adding/removing an edge affects degrees of two vertices.

   (c) Start with K_7. Show there is a sequence of deleting one edge then one vertex such that the result is complete. Show a different sequence where the result is not complete.
4. **Counting Labelled Graphs.**
   (a) How many labelled graphs on 5 vertices have exactly 1 edge?
   (b) How many labelled graphs on 5 vertices have exactly 3 edges? Exactly 4 edges?
   (c) Prove that the total number of simple labelled graphs on n vertices is $2^{\binom{n}{2}}$.

5. **Isomorphism.**

   (a) Prove that graph isomorphism is an equivalence relation.

   (b) Let $G_1 \cong G_2$. Prove they have the same order, size, and degree sequence.

   (c) Let G_1 be the graph with $V = \{a, b, c, d\}$ and $E = \{\{a, b\}, \{a, c\}, \{a, d\}\}$. Let G_2 be the graph with $V = \{A, B, C, D\}$ and $E = \{\{B, C\}, \{C, D\}, \{B, D\}\}$. Are they isomorphic? Provide an isomorphism or prove none exists.

   (d) Draw five pairwise non-isomorphic graphs on 5 vertices. Justify your answer.

6. **The Petersen Graph.** Let G be the graph whose vertices are the 2-element subsets of $\{1, 2, 3, 4, 5\}$, with adjacency defined by disjointness.

   (a) List all vertices and determine their degrees.

   (b) Draw the graph.

   (c) Is this graph bipartite?

7. $\star$ **Hypercube Edges.** Let $e(Q_n)$ be the number of edges in the n-dimensional hypercube.

   (a) Establish the recurrence $e(Q_n) = 2e(Q_{n-1}) + 2^{n-1}$ with $e(Q_0) = 0$.

   (b) Solve this recurrence to find a closed form for $e(Q_n)$.

   (c) Verify your formula by a direct combinatorial counting argument (using the Handshaking Lemma).

1

Extremal Graph Theory

In the previous chapter, we established the vocabulary of graph theory, defining structures such as connectivity, degrees, and specific families like K_n and C_n. We now turn to extremal graph theory, a field concerned with the relationship between local constraints (such as the absence of a specific substructure) and global properties (such as the number of edges).

The central question we address is: *How dense can a graph be without containing a forbidden substructure?*

1.1 Substructures

We begin by formalising the notion of one graph being "contained" within another.

Definition 1.1. ***Subgraph.***
Let $G = (V, E)$ be a graph. A graph $H = (V', E')$ is a **subgraph** of G, denoted $H \subseteq G$, if:

$$V' \subseteq V \quad \text{and} \quad E' \subseteq E.$$

If $H \subseteq G$ and H contains *all* edges of G capable of connecting vertices in V', i.e.,

$$E' = E \cap \mathcal{P}_2(V'),$$

then H is the **induced subgraph** of G on V', denoted $G[V']$. If $V' = V$, then H is a **spanning subgraph**.

定義

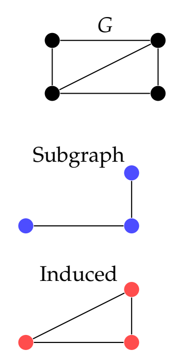


Figure 1.1: Top: Graph G. Middle: A subgraph on 3 vertices (missing edge $\{a, c\}$). Bottom: The induced subgraph $G[\{a, b, c\}]$ includes *all* edges from G.

To clarify the distinction between these types, consider the complete graph K_n. Its structure is maximal; every pair of vertices is an edge. Consequently, its subgraphs enumerates all possible simple graphs up to order n.

Example 1.1. Counting Substructures of K_n. Let K_n be the complete graph on n vertices. We determine the cardinality of the sets of its subgraphs.

Induced Subgraphs: An induced subgraph is uniquely determined by its vertex set $V' \subseteq V$. There are 2^n such subsets.

Spanning Subgraphs: The vertex set is fixed to V. A spanning subgraph is determined by choosing a subset of the edges of K_n. There are $\binom{n}{2}$ possible edges, so there are $2^{\binom{n}{2}}$ spanning subgraphs.

General Subgraphs: To form an arbitrary subgraph, we first choose k vertices (in $\binom{n}{k}$ ways), and then choose any subset of the edges available between them. Thus, the total number is:

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

範例

1.2 *Turán's Theorem*

A foundational result in extremal graph theory is Turán's Theorem. It bounds the number of edges in a graph that does not contain a complete subgraph of size r (denoted K_r).
If we wish to avoid K_3 (a triangle), the best strategy is to partition the vertices into two sets A and B and place all possible edges between A and B, but none within A or B. This yields a complete bipartite graph

$$K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil},$$

which contains no triangles. Turán's Theorem generalises this intuition.

Theorem 1.1. ***Turán's Theorem (1941).***
Let $G = (V, E)$ be a graph of order n and size m. Let $r \geq 2$. If G does not contain K_r as a subgraph, then:

$$m \leq \left(1 - \frac{1}{r-1}\right) \frac{n^2}{2}.$$

定理

Proof

We proceed by induction on n and r, ordered lexicographically: first smaller r, and for equal r smaller n. The base cases $n = 1$ (trivial) or $r = 2$ (no edges allowed, $m = 0$) hold immediately. Let $n > 1$ and $r > 2$. Assume the theorem holds for all graphs of order strictly less than n.
If G does not contain a clique of size $r - 1$ (K_{r-1}), we may apply the

induction hypothesis for $r-1$:

$$m \le \left(1-\frac{1}{r-2}\right)\frac{n^2}{2} \le \left(1-\frac{1}{r-1}\right)\frac{n^2}{2},$$

and the bound holds.

Now, assume G *does* contain a subgraph K_{r-1}. Let $A \subset V$ be the vertex set of this $(r-1)$-clique, and let $B = V \setminus A$. We partition the edge set E into three disjoint sets:

E_A: Edges with both endpoints in A. Since A induces K_{r-1}, $|E_A| = \binom{r-1}{2}$.

E_B: Edges with both endpoints in B. Since G contains no K_r, the induced subgraph $G[B]$ contains no K_r. By the inductive hypothesis on B (which has order $n-(r-1)$):

$$|E_B| \le \left(\frac{r-2}{r-1}\right)\frac{(n-r+1)^2}{2}.$$

E_{AB}: Edges connecting A and B. For any vertex $v \in B$, v can be adjacent to at most $r-2$ vertices in A. If v were adjacent to all $r-1$ vertices of A, then $A \cup \{v\}$ would form a K_r, contradicting the hypothesis. Thus:

$$|E_{AB}| \le |B|(r-2) = (n-r+1)(r-2).$$

Summing these components:

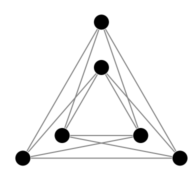
$$\begin{aligned} m &= |E_A| + |E_B| + |E_{AB}| \\ &\le \frac{(r-1)(r-2)}{2} + \frac{r-2}{2(r-1)}(n-r+1)^2 + (r-2)(n-r+1) \\ &= \frac{r-2}{2(r-1)}\left[(r-1)^2 + (n-r+1)^2 + 2(r-1)(n-r+1)\right]. \end{aligned}$$

Recognising the term in brackets as the expansion of $((r-1)+(n-r+1))^2 = n^2$, we obtain:

$$m \le \frac{r-2}{2(r-1)}n^2 = \left(1-\frac{1}{r-1}\right)\frac{n^2}{2}.$$

■

The bound is sharp. The configuration achieving equality is the **Turán Graph**, denoted $T(n, r-1)$. It is the complete multipartite graph formed by partitioning V into $r-1$ independent sets of size roughly $n/(r-1)$.



$T(6,3)$ (No K_4)

Figure 1.2: The Turán graph $T(n, r-1)$ is a complete multipartite graph formed by partitioning vertices into $r-1$ sets of equal (or nearly equal) size. It maximises edges while avoiding K_r.

Remark.

If n is a multiple of $r-1$, say $n = k(r-1)$, we partition V into $r-1$ sets $V_1, \dots, V_{r-1}$ each of size k. We connect u, v if and only if they belong to distinct sets. By the Pigeonhole Principle, any set of r vertices must contain at least two from the same partition V_i, which are non-adjacent. Thus, K_r is forbidden. The number of edges is the number of pairs of partitions times the edges between them:

$$|E| = \binom{r-1}{2} k^2 = \frac{(r-1)(r-2)}{2} \left(\frac{n}{r-1}\right)^2 = \left(1 - \frac{1}{r-1}\right) \frac{n^2}{2}.$$

1.3 Geometric Applications

Graph theoretic bounds often yield surprising results in geometry. We apply Turán's theorem to a problem involving distances in the plane.

Theorem 1.2. ***Erdős' Distance Theorem.***
Let S be a set of n points in the plane with diameter at most 1 (i.e., the maximum distance between any pair is 1). The number of pairs of points separated by a distance strictly greater than $1/\sqrt{2}$ is at most $\lfloor n^2/3 \rfloor$.

定理

Proof

Construct a graph $G = (S, E)$ where an edge $\{x, y\}$ exists if and only if $d(x, y) > 1/\sqrt{2}$. We claim that this graph cannot contain a complete subgraph of order 4 (K_4).
Suppose for contradiction that $\{x, y, z, t\} \subseteq S$ forms a K_4. This implies the distance between any two of these points is greater than $1/\sqrt{2}$. Consider the geometric configuration of these four points. In any set of four planar points, either one lies in the convex hull of the others (so some angle at that point is at least $120°$) or they form a convex quadrilateral (so some interior angle is at least $90°$). In either case, there exists a triplet, say x, y, z, forming an angle $\angle xyz \geq 90°$. By the Law of Cosines (or simply Pythagoras if $\angle xyz = 90°$):

$$d(x, z)^2 \geq d(x, y)^2 + d(y, z)^2.$$

Since $d(x, y) > 1/\sqrt{2}$ and $d(y, z) > 1/\sqrt{2}$, we have:

$$d(x, z)^2 > \frac{1}{2} + \frac{1}{2} = 1 \implies d(x, z) > 1.$$

This contradicts the assumption that the diameter of S is at most 1.

Thus, G is K_4-free. Applying Turán's Theorem with $r = 4$:

$$|E| \leq \left(1 - \frac{1}{3}\right) \frac{n^2}{2} = \frac{n^2}{3}.$$

■

1.4 *Independent Sets*

We defined an independent set (or stable set) implicitly when discussing Turán graphs. We now treat them formally.

Definition 1.2. ***Independent Set.***
Let $G = (V, E)$. A subset $S \subseteq V$ is an **independent set** if no two vertices in S are adjacent. The **independence number** $\alpha(G)$ is the cardinality of the largest independent set in G.

定義

This concept is dual to that of a clique via the complement graph.

Definition 1.3. ***Complement Graph.***
The **complement** of $G = (V, E)$ is the graph $\overline{G} = (V, \overline{E})$, where $\{u, v\} \in \overline{E}$ if and only if $\{u, v\} \notin E$ (for $u \neq v$).

定義

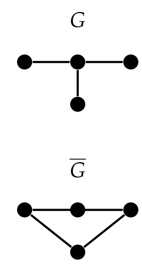


Figure 1.3: A graph G and its complement $\overline{G}$. The clique $\{a, c, d\}$ in $\overline{G}$ is an independent set in G.

Note
A set of vertices forms a clique in G if and only if it forms an independent set in $\overline{G}$.

Using this duality, we can translate Turán's theorem into a lower bound for the size of independent sets.

Corollary 1.1. *Lower Bound for* $\alpha(G)$. Let G be a graph of order n and size m. Then:

$$\alpha(G) \geq \frac{n^2}{2m + n}.$$

推論

Proof
Let $\alpha = \alpha(G)$. Since an independent set in G is a clique in $\overline{G}$, the graph $\overline{G}$ contains a clique of size α, but no clique of size $\alpha + 1$. Applying Turán's Theorem to $\overline{G}$, which has size $\binom{n}{2} - m$:

$$|E(\overline{G})| \leq \left(1 - \frac{1}{\alpha + 1 - 1}\right) \frac{n^2}{2} = \left(\frac{\alpha - 1}{\alpha}\right) \frac{n^2}{2}.$$

Substituting $|E(\overline{G})| = \frac{n(n-1)}{2} - m$:

$$\frac{n^2 - n}{2} - m \leq \frac{\alpha - 1}{2\alpha} n^2.$$

Multiplying by 2α:

$$\alpha(n^2 - n - 2m) \leq (\alpha - 1)n^2 = \alpha n^2 - n^2.$$

Cancelling αn^2 from both sides:

$$-\alpha(n + 2m) \leq -n^2 \iff \alpha(2m + n) \geq n^2.$$

Rearranging yields the result.

■

1.5 *Ramsey Theory*

While Turán's theorem determines the maximum density of edges before a specific substructure *must* appear, Ramsey theory poses a more fundamental question: is total disorder possible? The central result, Ramsey's Theorem, asserts that in any sufficiently large graph, one can find either a highly connected substructure (a clique) or a completely disconnected one (an independent set).

Definition 1.4. ***Ramsey Numbers.***
Let $s, t \in \mathbb{N}^*$. The **Ramsey number** $R(s, t)$ is the minimum integer n such that for any graph G of order n, either:

- G contains K_s as a subgraph, or
- G contains S_t (an independent set of size t) as an induced subgraph.

Equivalently, referencing the complement graph, any graph of order $R(s, t)$ satisfies $K_s \subseteq G$ or $K_t \subseteq \overline{G}$.

定義

The existence of such numbers is not immediately obvious; a priori, one might construct arbitrarily large graphs avoiding both structures. Ramsey's theorem guarantees their finiteness. We begin with the elementary boundary values.

Proposition 1.1. ***Basic Properties.***
For all $s, t \geq 2$:

1. **Symmetry:** $R(s, t) = R(t, s)$.
2. **Triviality:** $R(1, t) = 1$.
3. **Pairs:** $R(2, t) = t$.

命題

Proof

1. Let $n = R(s, t)$. For any graph G of order n, either $K_s \subseteq G$ or $S_t \subseteq G$. Taking complements, for any graph $H = \overline{G}$, either $S_s \subseteq H$ or $K_t \subseteq H$. Thus $R(t, s) \leq R(s, t)$. The reverse inequality holds similarly.

2. A graph of order 1 contains K_1 (a single vertex).
3. We show $R(2,t) = t$.
   - Upper bound: Let G be a graph of order t. If G contains at least one edge, then $K_2 \subseteq G$. If G contains no edges, then $G \cong S_t$. In either case, the condition is satisfied.
   - Lower bound: Consider the graph S_{t-1}. It has order $t-1$, contains no edges (no K_2), and has size strictly less than t (no S_t). Thus $R(2,t) > t-1$.

■

The finiteness of Ramsey numbers for general s,t is established via a recurrence relation, essentially a graph-theoretic application of the Pigeonhole Principle.

Theorem 1.3. *Ramsey Recurrence.*
For integers $s,t \geq 2$, the Ramsey numbers satisfy:

$$R(s,t) \leq R(s-1,t) + R(s,t-1).$$

定理

Proof

Let $n = R(s-1,t) + R(s,t-1)$. Consider an arbitrary graph $G = (V,E)$ of order n. Pick any vertex $v \in V$. We partition the remaining $n-1$ vertices into two sets:

$$A = N(v) = \{u \in V : \{u,v\} \in E\},$$

$$B = V \setminus (A \cup \{v\}) = \{u \in V : \{u,v\} \notin E\}.$$

Since $|A| + |B| = n-1 = R(s-1,t) + R(s,t-1) - 1$, the Pigeonhole Principle implies that either $|A| \geq R(s-1,t)$ or $|B| \geq R(s,t-1)$.

Case 1: $|A| \geq R(s-1,t)$. Consider the subgraph induced by A, $G[A]$. By definition of the Ramsey number, $G[A]$ must contain either K_{s-1} or S_t.

- If $S_t \subseteq G[A]$, then $S_t \subseteq G$, and we are done.
- If $K_{s-1} \subseteq G[A]$, let K be the set of vertices forming this clique. Since v is adjacent to every vertex in A (and thus every vertex in K), the set $K \cup \{v\}$ forms a clique of size s (K_s) in G.

Case 2: $|B| \geq R(s,t-1)$. Consider the induced subgraph $G[B]$. It must contain either K_s or S_{t-1}.

- If $K_s \subseteq G[B]$, then $K_s \subseteq G$, and we are done.
- If $S_{t-1} \subseteq G[B]$, let S be this independent set. Since v is adjacent to no vertex in B, $S \cup \{v\}$ forms an independent set of size t (S_t) in G.

In all scenarios, G satisfies the Ramsey condition. Thus $R(s,t) \leq n$. ■

Example 1.2. The Party Problem ($R(3,3)$). We determine the value of $R(3,3)$. From the recurrence relation and the base case $R(2,3) = 3$:

$$R(3,3) \leq R(2,3) + R(3,2) = 3 + 3 = 6.$$

To show $R(3,3) > 5$, we must exhibit a graph of order 5 containing neither a triangle (K_3) nor an independent set of size 3 (S_3). Consider the cycle C_5 (see *Figure 6* or the Bestiary section).

Cliques: The maximum clique size in C_5 is 2 (an edge), so it is K_3-free.

Independent Sets: The maximum independent set size is 2 (any pair of non-adjacent vertices forces the remaining three to be connected enough to prevent a third independent choice). Thus it is S_3-free.

Since C_5 serves as a counterexample for $n = 5$, we conclude $R(3,3) = 6$.

Interpretation: In any group of 6 people, there are either 3 mutual friends or 3 mutual strangers.

範例

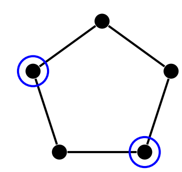


Figure 1.4: The cycle C_5 witnesses $R(3,3) > 5$. Blue circles mark a maximum independent set of size 2; adding any third vertex creates an edge.

1.6 *Exercises*

1. **Substructures of K_n.**
   (a) Prove that the number of induced subgraphs of K_n is 2^n.
   (b) Prove that the number of spanning subgraphs of K_n is $2^{\binom{n}{2}}$.
   (c) Show that the total number of subgraphs of K_n is $\sum_{k=0}^{n} \binom{n}{k} 2^{\binom{k}{2}}$.
2. **Defining Subgraphs.** Let $G = (V, E)$ be a simple graph.
   (a) Show that for any subset $S \subseteq V$, the induced subgraph $G[S]$ is uniquely determined by S.
   (b) Let H be an arbitrary subgraph of G. Prove that there exists a unique vertex subset $S \subseteq V$ and a unique edge subset $F \subseteq E(G[S])$ such that $H = (S, F)$.
3. **Triangle-Free Graphs.** Let G be a graph of order n containing no triangles (K_3-free).
   (a) Use Turán's Theorem ($r = 3$) to show that $|E(G)| \leq \lfloor n^2/4 \rfloor$.

(b) Describe the graphs that achieve equality in this bound.

4. **Turán Bound Practice.** Let $n_1, \ldots, n_k$ be positive integers summing to n. Consider K_n partitioned into disjoint sets V_i of size n_i.

   (a) Explain why $\sum_{i=1}^{k} \binom{n_i}{2}$ counts the edges *inside* the parts.

   (b) Show that $\sum_{i=1}^{k} \binom{n_i}{2} \leq \binom{n}{2}$. When does equality hold?

5. ⋆ **Sharpness of Turán's Theorem.** Let $T(n, r-1)$ be the Turán graph (complete $(r-1)$-partite graph with roughly equal parts).

   (a) Show that the number of edges is $|E(T(n, r-1))| = \frac{1}{2}(n^2 - \sum_{i=1}^{r-1} n_i^2)$.

   (b) Deduce that $|E(T(n, r-1))| \geq \left(1 - \frac{1}{r-1}\right) \frac{n^2}{2}$, with equality when all parts are as equal as possible.

   (c) Conclude that the bound in Turán's theorem cannot be improved.

6. ⋆ **Uniqueness of Turán Graphs.** Let G be a K_r-free graph on n vertices. Prove that if G has the maximum possible number of edges among all such graphs, then G must be isomorphic to the Turán graph $T(n, r-1)$.

7. **Independence and Cliques.** Let $G = (V, E)$ and $\overline{G}$ be its complement.

   (a) Prove that $S \subseteq V$ is a clique in G if and only if S is an independent set in $\overline{G}$.

   (b) Show that $\alpha(G) = \omega(\overline{G})$, where $\omega(G)$ is the clique number (size of the largest clique).

8. **Independence Bounds.** Verify the inequality $\alpha(G) \geq \frac{n^2}{2m+n}$ for the following families:

   (a) The complete graph K_n.

   (b) The cycle graph C_n.

   (c) The path graph P_n (length n, order $n+1$).

   Comment on the tightness of the bound in each case.

9. **Independence and Average Degree.** Let $\bar{d}$ be the average degree of a graph G on n vertices. Show that:

$$\alpha(G) \geq \frac{n}{\bar{d}+1}.$$

> *Remark.*
>
> Hint: Relate $\bar{d}$ to the number of edges m and use the bound from the previous exercise.

10. **Geometry Lemma.** Show that for any 4 points in the plane, there

exist three of them forming an angle of at least 90°.

> *Remark.*
>
> Consider the convex hull.

11. **Erdős Distance Proof Step.** Complete the missing step in the proof of Erdős' theorem: Prove that if $\angle xyz \geq 90^\circ$ and $d(x,y) > 1/\sqrt{2}$ and $d(y,z) > 1/\sqrt{2}$, then $d(x,z) > 1$.

12. ⋆ **Long Distances Construction.** Construct a set of n points in the plane with diameter 1 such that the number of pairs with distance $> 1/\sqrt{2}$ is of the order $n^2/3$.

> *Remark.*
>
> Place points in tight clusters near the vertices of an equilateral triangle.

13. **Ramsey Basics.**
    (a) Prove $R(1,t) = 1$ and $R(2,t) = t$.
    (b) Prove the symmetry $R(s,t) = R(t,s)$.

14. **Ramsey's Theorem.** Use the recurrence $R(s,t) \leq R(s-1,t) + R(s,t-1)$ and the base cases to prove by induction on $s+t$ that $R(s,t)$ is finite for all positive integers s,t.

15. **The Party Problem.** Re-prove $R(3,3) = 6$ by showing:
    (a) $R(3,3) \leq 6$ using the recurrence.
    (b) C_5 contains no K_3 and no S_3.
    (c) Deduce $R(3,3) > 5$.

16. ⋆ **Bounding $R(3,4)$.** Use the recurrence relation and known values to derive an explicit upper bound for $R(3,4)$.

2 Connectivity

Having established the structural definitions of graphs and their subgraphs, we now turn to the fundamental topological notion of connectivity. Intuitively, a graph is connected if it is possible to travel between any two vertices along the edges of the graph. To formalise this, we must define the precise nature of "travel" within a discrete structure.

2.1 *Walks, Trails, and Paths*

We distinguish between sequences of adjacent vertices based on whether they repeat vertices or edges.

Definition 2.1. ***Walks and Paths.***
Let $G = (V, E)$ be a graph.

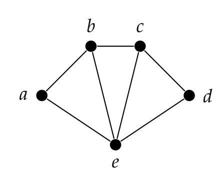
- A **walk** of length k is a sequence of vertices $(v_0, v_1, \ldots, v_k)$ such that $\{v_i, v_{i+1}\} \in E$ for all $0 \leq i < k$. The vertex v_0 is the **start** and v_k is the **end** (or terminus).
- A **trail** is a walk in which all edges $\{v_i, v_{i+1}\}$ are distinct.
- A **path** is a walk in which all vertices v_i are distinct.

If $v_0 = v_k$, the walk is **closed**. A closed path (where only $v_0 = v_k$ are repeated) is a **cycle**.

定義

Note

In the previous chapter’s, we defined the Path Graph P_n and Cycle Graph C_n. A path of length k in G is a subgraph isomorphic to P_{k+1}, and a cycle of length k is isomorphic to C_k.



Walk: (a, b, e, b, c, d)
Trail: (a, e, b, c, e, d)
Path: (a, b, c, d)

Figure 2.1: Walk, trail, and path from a to d. The walk repeats edge $\{b, e\}$; the trail repeats vertex e but no edge; the path has no repetitions.

It is immediate that every path is a walk. Conversely, while a walk may wander and loop back on itself, the existence of a walk implies the existence of a path.

Lemma 2.1. ***Walk Reduction.***
If there exists a walk connecting u to v in G, then there exists a path connecting u to v.

引理

Proof

Let $W = (u = v_0, v_1, \ldots, v_k = v)$ be a walk from u to v of minimal length. Suppose W is not a path. Then the vertices are not distinct, so there exist indices $i < j$ such that $v_i = v_j$. We can form a new sequence W' by excising the segment between i and j:

$$W' = (v_0, \ldots, v_i, v_{j+1}, \ldots, v_k).$$

Since $v_i = v_j$, the pair $\{v_i, v_{j+1}\}$ is the edge $\{v_j, v_{j+1}\}$, which exists in G. Thus W' is a valid walk from u to v with length strictly less than W. This contradicts the minimality of W. Therefore, the minimal walk is a path.

■

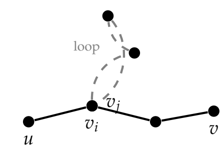


Figure 2.2: Reducing a walk to a path. If a walk intersects itself at $v_i = v_j$, the loop (dashed) can be removed to form a shorter walk.

Connected Components

We define a binary relation on the vertex set V to capture global cohesion.

Definition 2.2. ***Connectivity Relation.***
We say two vertices $x, y \in V$ are **connected**, denoted $x \sim y$, if there exists a walk (and hence, by *lemma* 2.1, a path) starting at x and ending at y.

定義

Theorem 2.1. ***Connectivity is an Equivalence Relation.***
The relation $\sim$ is an equivalence relation on V.

定理

Proof

1. **Reflexivity:** For any $x \in V$, the trivial sequence (x) is a walk of length 0. Thus $x \sim x$.
2. **Symmetry:** If $x \sim y$, there is a walk $(x, v_1, \ldots, v_{k-1}, y)$. Reversing this sequence yields $(y, v_{k-1}, \ldots, v_1, x)$, which is a valid walk since edges are undirected sets $\{u, v\}$. Thus $y \sim x$.
3. **Transitivity:** If $x \sim y$ and $y \sim z$, there exist walks $W_1 = (x, \ldots, y)$ and $W_2 = (y, \ldots, z)$. The concatenation $W_1 \cdot W_2 = (x, \ldots, y, \ldots, z)$ is a walk from x to z. Thus $x \sim z$.

■

The equivalence classes under $\sim$ are called the **connected components** of G. A graph is **connected** if it has exactly one connected component; otherwise, it is **disconnected**.

Example 2.1. Components of $2K_3$. Recall the graph $2K_3$ consisting of two disjoint triangles with vertices $\{1,2,3\}$ and $\{4,5,6\}$. No edge connects the first set to the second. The relation $\sim$ partitions V into $C_1 = \{1,2,3\}$ and $C_2 = \{4,5,6\}$. These are the connected components.

範例

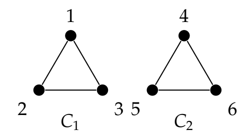


Figure 2.3: The graph $2K_3$: two disjoint triangles forming two connected components.

Cuts and Partitions

Connectivity can be equivalently characterised by the absence of a "cut" — a partition of the vertices into two sets with no edges crossing between them. This dual perspective is often more useful for proofs involving contradiction.

Theorem 2.2. ***The Cut Condition.***
A graph $G = (V, E)$ is connected if and only if for every partition of V into two non-empty sets A and B, there exists an edge $\{u, v\} \in E$ such that $u \in A$ and $v \in B$.

定理

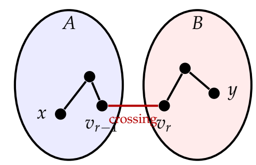


Figure 2.4: If x and y are connected, any path between them must cross the boundary between A and B at least once.

$(\Longrightarrow)$

Assume G is connected. Let $V = A \cup B$ be a partition with $A, B \neq \varnothing$. Pick arbitrary vertices $x \in A$ and $y \in B$. Since G is connected, there exists a path $P = (v_0, \dots, v_k)$ with $v_0 = x$ and $v_k = y$. We traverse the path from x. Since $v_0 \in A$ and $v_k \in B$, there must be a first index r where the path leaves A. Let

$$r = \min\{i : v_i \in B\}.$$

Since $v_0 \in A$, we have $r > 0$. By definition of r, $v_{r-1} \notin B$, so $v_{r-1} \in A$. The edge $\{v_{r-1}, v_r\}$ connects a vertex in A to a vertex in B.

証明終

$(\Longleftarrow)$

Assume that for every non-trivial partition, a crossing edge exists. Suppose for contradiction that G is disconnected. Let C be a connected component of G. Since G is disconnected, $C \subsetneq V$, so let $A = C$ and $B = V \setminus C$. Both sets are non-empty. By the hypothesis, there exists an edge $\{u, v\}$ with $u \in A$ and $v \in B$. Since $u \in C$ and $\{u, v\} \in E$, there is a walk from any vertex in C to v. By transitivity, v must belong to the connected component C. This implies $v \in A$, contradicting $v \in B$. Thus, G must be connected.

証明終

This theorem is particularly useful when proving global connectivity properties from local degree conditions.

Proposition 2.1. ***Minimum Degree and Connectivity.***
If G is a graph of order n such that $\delta(G) \geq \lfloor n/2 \rfloor$, then G is connected.

命題

Proof

Suppose G is disconnected. Then the vertex set can be partitioned into sets A and B with no edges between them. Let $u \in A$. All neighbours of u must lie within A, so $d(u) \leq |A| - 1$. Thus $|A| \geq d(u) + 1 \geq \lfloor n/2 \rfloor + 1 > n/2$. Similarly, for any $v \in B$, $|B| \geq d(v) + 1 > n/2$. Summing the sizes:

$$|V| = |A| + |B| > n/2 + n/2 = n.$$

This is a contradiction. Therefore, no such partition exists, and G is connected by the Cut Condition (*theorem* 2.2).

■

2.2 *Distance*

When a graph is connected, we can measure the separation between vertices.

Definition 2.3. ***Geodesic Distance.***
Let $G = (V, E)$ be a connected graph. The **distance** between two vertices $s, t \in V$, denoted $d_G(s, t)$, is the length of the shortest path connecting s and t. If G is not connected and s, t lie in different components, we define $d_G(s, t) = \infty$.

定義

While the definition relies on graph structure, the function d_G satisfies the axioms of a metric space on the set of vertices, provided the graph is connected.

Proposition 2.2. ***Metric Properties.***
The function $d_G : V \times V \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ satisfies:

1. **Separation:** $d_G(s, t) = 0 \iff s = t$.
2. **Symmetry:** $d_G(s, t) = d_G(t, s)$.
3. **Triangle Inequality:** For all $s, x, t \in V$,

$$d_G(s, t) \leq d_G(s, x) + d_G(x, t).$$

命題

Proof

Separation and symmetry follow immediately from the definition of a path (length 0 implies a single vertex; edges are undirected).

For the triangle inequality, observe that the concatenation of a shortest path from s to x and a shortest path from x to t forms a walk from s to t of length $d_G(s,x) + d_G(x,t)$. By *lemma* 2.1, this walk contains a path from s to t of equal or lesser length. Thus, the shortest path from s to t cannot exceed this sum.

■

Algebraic Counting of Walks

The adjacency matrix A_G encodes the existence of edges (paths of length 1). Its powers generalise this to paths of arbitrary length.

Note

Although the provided source text refers to "trails", standard algebraic graph theory confirms that matrix powers count **walks** (sequences where vertices and edges may repeat). We present the mathematically correct statement here.

Theorem 2.3. ***Counting Walks.***
Let G be a graph with vertex set $V = \{v_1, \ldots, v_n\}$ and adjacency matrix A_G. For any $k \in \mathbb{N}$, the number of walks of length k connecting v_i to v_j is given by the entry $(A_G^k)_{ij}$.

定理

We proceed by induction on k.

Base Case ($k = 0$)

$A_G^0 = I_n$. A walk of length 0 from v_i to v_j exists if and only if $v_i = v_j$, which corresponds to the identity matrix entries.

証明終

Inductive Step

Assume the property holds for k. A walk of length $k+1$ from v_i to v_j consists of a walk of length k from v_i to some neighbour v_ℓ, followed by the edge $\{v_\ell, v_j\}$. Summing over all possible penultimate vertices v_ℓ:

$$N_{k+1}(v_i, v_j) = \sum_{v_\ell \in V} N_k(v_i, v_\ell) \cdot (A_G)_{\ell j}.$$

By the inductive hypothesis, $N_k(v_i, v_\ell) = (A_G^k)_{i\ell}$. Thus:

$$N_{k+1}(v_i, v_j) = \sum_{\ell=1}^{n} (A_G^k)_{i\ell}(A_G)_{\ell j} = (A_G^k \cdot A_G)_{ij} = (A_G^{k+1})_{ij}.$$

証明終

Corollary 2.1. *Distance via Matrices.* The distance between distinct vertices v_i, v_j is the smallest power for which the matrix entry is non-zero:

$$d_G(v_i, v_j) = \min\{k \geq 1 : (A_G^k)_{ij} \neq 0\}.$$

推論

Weighted Graphs

In many applications, edges are not uniform; they carry costs such as length, time, or resistance.

Definition 2.4. ***Weighted Graph.***
A **weighted graph** is a triple (V, E, μ), where $\mu : E \to \mathbb{R}_{>0}$ is a **valuation function**. The valuation (or weight) of a walk $\gamma = (v_0, \dots, v_k)$ is the sum of its edge weights:

$$\mu(\gamma) = \sum_{i=1}^{k} \mu(\{v_{i-1}, v_i\}).$$

The weighted distance $d_\mu(s, t)$ is the minimum valuation of any walk connecting s and t.

定義

Note
If $\mu(e) = 1$ for all $e \in E$, the weighted distance d_μ coincides with the geodesic distance d_G.

Dijkstra's Algorithm

To compute $d_\mu(s, t)$ efficiently, we employ Dijkstra's Algorithm. This greedy method maintains the shortest known distance from a source s to all other vertices, iteratively "settling" the closest vertex.
Algorithm State: Let s be the source. We maintain two arrays:
- $C[v]$: The current minimal cost found from s to v. Initially $C[s] = 0$ and $C[v] = \infty$ for $v \neq s$.
- $T[v]$: The predecessor of v on the optimal path.

We partition vertices into two sets: $V_{settled}$ (distance known) and W_{active} (distance tentative). Initially $V_{settled} = \varnothing$ and $W_{active} = V$.
Procedure: While W_{active} contains a vertex with finite C value: 1. Select $u \in W_{active}$ with minimal $C[u]$. 2. Move u from W_{active} to $V_{settled}$. 3. For each neighbour x of u in W_{active}: If the path through u is shorter ($C[u] + \mu(\{u, x\}) < C[x]$):
- Update $C[x] \leftarrow C[u] + \mu(\{u, x\})$.
- Update $T[x] \leftarrow u$.

The complexity of this algorithm is $\mathcal{O}((n + m) \log n)$, making it highly efficient for sparse graphs. The correctness relies on the non-negativity of weights: once a vertex u is settled, no shorter path to u

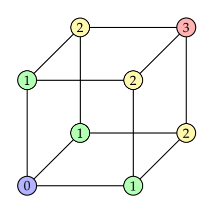


Figure 2.5: The Hypercube Q_3 with distances from a source vertex. Dijkstra's algorithm settles vertices in order of increasing distance.

can be found through more distant vertices.

2.3 Eulerian Tours and Trails

We examine trails and cycles that exhaustively visit the structural elements of a graph.

Eulerian Graphs

Definition 2.5. ***Eulerian Definitions.***
Let $G = (V, E)$ be a graph.

- An **Eulerian tour** is a closed trail that traverses every edge of G exactly once (and therefore visits every non-isolated vertex).
- An **Eulerian graph** is a graph that admits an Eulerian tour.
- An **Eulerian trail** is a trail that traverses every edge of G exactly once (not necessarily closed).
- A **semi-Eulerian graph** is a graph that admits an Eulerian trail.

定義

The existence of such tours is determined by the connectivity and degree parity of the graph.

Theorem 2.4. ***Euler's Theorem.***
Let G be a graph of order at least 2. G is Eulerian if and only if it is connected and all its vertices have even degree.

定理

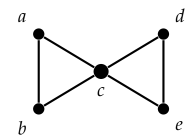


Figure 2.6: The bowtie graph is Eulerian: all degrees even, admitting the tour shown.

($\Longrightarrow$)

Let $G = (V, E)$ be Eulerian and let $\gamma = (v_0, v_1, \ldots, v_k = v_0)$ be an Eulerian tour. G is necessarily connected, as the tour visits every vertex and the sequence of edges connects any pair of vertices on the tour. For any vertex $v \in V$, let E_v be the set of edges incident to v. We classify an edge $e \in E_v$ as *incoming* if it appears in the sequence as $\{v_{i-1}, v_i\}$ with $v_i = v$, and *outgoing* if it appears as $\{v_i, v_{i+1}\}$ with $v_i = v$. The map $\{v_{i-1}, v_i\} \mapsto \{v_i, v_{i+1}\}$ (with the convention that $\{v_{k-1}, v_k\} \mapsto \{v_0, v_1\}$) constitutes a bijection between the incoming and outgoing edges incident to v. Consequently, $|E_v| = d(v)$ must be even.

証明終

($\Longleftarrow$)

Assume G is connected and every vertex has even degree. Let $\gamma = (v_0, \ldots, v_k)$ be a trail in G with no repeated edges and of maximal length.

Claim 1: ***γ is closed ($v_k = v_0$).*** Suppose $v_k \neq v_0$. The number of edges in γ incident to v_k would be odd (one entry, plus pairs of entry/exit for any previous visits). However, $d(v_k)$ is even in G. Thus, there exists an edge incident to v_k not used in γ. Extending γ by this edge yields a longer trail, contradicting maximality.

Claim 2:* γ *visits every edge. Suppose there exists an edge $a = \{v_i, w\}$ not in γ with v_i on γ (possible since G is connected). By Claim 1, γ is closed, so we may cyclically permute it to start and end at v_i:

$$\gamma = (v_i, v_{i+1}, \ldots, v_k = v_0, v_1, \ldots, v_{i-1}, v_i).$$

Appending the unused edge yields the trail

$$\gamma' = (v_i, v_{i+1}, \ldots, v_{i-1}, v_i, w),$$

which contains every edge of γ plus a, hence is strictly longer—contradicting the maximality of γ.

Claim 3:* γ *visits every vertex. Since G is connected and γ visits every edge, any vertex with non-zero degree is visited. If an isolated vertex existed, G would not be connected (unless the order is 1, which is excluded).

証明終

We extend this characterisation to semi-Eulerian graphs.

Corollary 2.2. *Semi-Eulerian Characterisation.* A graph G is semi-Eulerian if and only if it is connected and the number of vertices of odd degree is 0 or 2.

推論

($\Longrightarrow$)

Let γ be an Eulerian trail. G is connected. As shown in the previous proof, any vertex strictly internal to the trail (not the start or end) must have even degree. The start and end vertices have odd degree unless the trail is closed. Thus the count of odd-degree vertices is 0 (if closed) or 2 (if open).

証明終

($\Longleftarrow$)

If G has 0 vertices of odd degree, it is Eulerian and thus semi-Eulerian. Suppose G has exactly two vertices, u and v, of odd degree. Construct a new graph G' by adding a vertex s and edges $\{u, s\}$ and $\{v, s\}$. In G', u and v now have even degree (degree increased by 1), and s has degree 2. Thus G' is Eulerian. Let γ' be an Eulerian tour of G'. The edges $\{u, s\}$ and $\{v, s\}$ must appear consecutively in γ' (as s has degree 2). Removing s and these two edges breaks the cycle into a trail connecting u and v that covers all edges of G.

証明終

Hamiltonian Graphs

We define the corresponding concept for vertices.

Definition 2.6. ***Hamiltonian Definitions.***

- A **Hamiltonian cycle** is a cycle that visits every vertex of the graph exactly once. Its length is equal to the order n of the graph.
- A **Hamiltonian graph** is a graph that admits a Hamiltonian cycle.
- A **Hamiltonian chain** is a chain that visits every vertex of the graph exactly once.
- A **semi-Hamiltonian graph** is a graph that admits a Hamiltonian chain.

定義

Unlike the Eulerian case, there is no known simple characterisation (like "all degrees even") for Hamiltonian graphs. Determining if a graph is Hamiltonian is an NP-complete problem.

Example 2.2. Contrasting Eulerian and Hamiltonian. Consider the graph formed by two triangles sharing a vertex (the "butterfly" or "bowtie" graph).

- **Eulerian?** Yes. The central vertex has degree 4, and the four outer vertices have degree 2. Since all degrees are even and the graph is connected, it is Eulerian.
- **Hamiltonian?** No. Any cycle must pass through the central vertex. Once it enters one triangle and returns to the center, it cannot enter the second triangle without visiting the center a second time, which is forbidden.

Conversely, the complete graph K_n ($n \geq 3$) is always Hamiltonian, but is Eulerian only if n is odd (so degrees $n-1$ are even).

範例

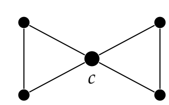


Figure 2.7: The butterfly (bowtie) graph: two triangles sharing vertex c. Eulerian (all degrees even: c has degree 4, others have degree 2), but not Hamiltonian.

2.4 *Exercises*

1. **Tree Properties.**

   (a) Let T be a tree with n_k vertices of degree k. Prove that the number of leaves is:

   $$n_1 = 2 + \sum_{k=3}^{\infty} (k-2) n_k.$$

   (b) Draw all non-isomorphic trees on 6 vertices.

(c) Prove that the centre of a tree consists of either a single vertex or two adjacent vertices. (The centre is the set of vertices minimising the maximum distance to any other vertex).

2. **Refined Cayley Practice.**

(a) Calculate the number of trees on the vertex set $\{1, 2, 3, 4, 5, 6\}$ where degrees are $d(1) = 3, d(2) = 3, d(3) = 1, d(4) = 1, d(5) = 1, d(6) = 1$.

(b) Verify your answer by listing the possible structures (up to relabelling vertices with the same degree).

3. **Prüfer Encoding.**

(a) Find the Prüfer sequence of the path graph P_{n-1} on vertices $1, 2, \ldots, n$ in natural order (edges $\{i, i+1\}$).

(b) Find the Prüfer sequence of the star graph $K_{1,n-1}$ with centre at vertex n.

(c) Construct the tree corresponding to the sequence $(1, 3, 5, 5, 3)$ on vertices $\{1, \ldots, 7\}$.

4. **Counting Forests.** Let F be a forest on n labelled vertices with k connected components. Prove that the number of such forests where vertices $1, 2, \ldots, k$ belong to distinct components is kn^{n-k-1}.

> *Remark.*
>
> Generalise the Prüfer argument or use Cayley's formula on a slightly modified graph.

5. **Spanning Trees.**

(a) Calculate the number of spanning trees of the complete bipartite graph $K_{2,m}$.

(b) Let G be the cycle C_n. How many spanning trees does it have?

(c) $\star$ Let G be the "ladder graph" $P_2 \times P_n$. Find a recurrence for the number of spanning trees.

6. **Matrix Tree Theorem Application.** The number of spanning trees of a graph G is any cofactor of its Laplacian matrix $L = D - A$.

(a) Write down the Laplacian matrix for K_4.

(b) Compute a cofactor to verify Cayley's formula for $n = 4$ ($4^{4-2} = 16$).

7. **Tree Diameter.** Let T be a tree.

(a) Prove that for every $k \geq 1$, the intersection of all paths of length at least k is either empty or a path.

(b) Prove that if the diameter of T is d, then T has at least d leaves? (False. Find a counter-example). Correct statement: If $\Delta(T) \geq k$, T has at least k leaves.

8. $\star$ **Double Counting.** Let $N(n,k)$ be the number of forests on n labelled vertices with k edges. Show that:

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

> *Remark.*
>
> This generalises Cayley's formula (case $k = n - 1$).

3
Trees

In the previous chapter, we explored connectivity as a topological property: can one navigate between any two vertices? We now refine this question to ask: what is the *minimal* structure required to maintain connectivity? Conversely, what is the *maximal* structure one can build without creating redundant loops?

The answer to both questions lies in the concept of a **tree**. Trees form the skeleton of graph theory; they are the simplest connected graphs, yet they admit a rich set of equivalent characterisations.

3.1 *Forests and Trees*

We begin by formally excluding the existence of cycles. Recall that a cycle is a closed walk with no repeated vertices (other than the start/end).

> **Definition 3.1.** ***Trees and Forests.***
> Let $G = (V, E)$ be a graph.
> · G is a **forest** if it is **acyclic** (contains no cycles).
> · G is a **tree** if it is acyclic and **connected**.
>
> 定義

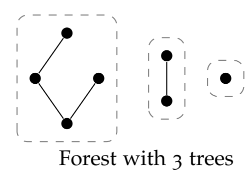


Figure 3.1: A forest: an acyclic graph whose connected components are trees.

The connected components of a forest are trees. The absence of cycles imposes strict constraints on the "ends" of the graph.

> **Definition 3.2.** ***Leaves.***
> A vertex v is a **leaf** (or pendant vertex) if $d(v) = 1$.
>
> 定義

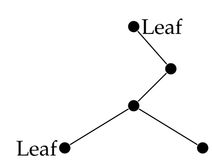


Figure 3.2: A tree of order 5. The vertices of degree 1 are leaves.

Every finite tree (with at least one edge) must have an "end". This topological intuition is formalised via the maximal path argument.

> **Lemma 3.1.** ***Existence of Leaves.***
> Let $T = (V, E)$ be a forest with $E \neq \varnothing$. Then T contains at least two leaves.
>
> 引理

Proof

Consider a path $P = (v_0, v_1, \ldots, v_k)$ in T of **maximal length**. Since $E \neq \emptyset$, such a path exists with $k \geq 1$. We claim v_0 and v_k are leaves. Suppose for contradiction that $d(v_0) > 1$. Then v_0 has a neighbour $u \neq v_1$.

- If u lies on the path (i.e., $u = v_i$ for some $i > 1$), then the edge $\{u, v_0\}$ completes a cycle $(v_0, v_1, \ldots, v_i, v_0)$, contradicting the acyclicity of T.
- If u does not lie on the path, we can extend P to $(u, v_0, v_1, \ldots, v_k)$, creating a path of length $k+1$. This contradicts the maximality of P.

Thus, $d(v_0) = 1$. By symmetry, $d(v_k) = 1$.

■

This lemma provides the engine for inductive proofs on trees: one can "prune" a leaf to reduce the order of the graph while preserving the tree structure.

Theorem 3.1. ***Tree Edges.***
Let $T = (V, E)$ be a tree of order n. Then $|E| = n - 1$.

定理

We proceed by induction on n.

Base Case ($n = 1$)

A graph with 1 vertex and no cycles must have 0 edges. $0 = 1 - 1$.

証明終

Inductive Step

Assume the statement holds for all trees of order $n - 1$. Let T be a tree of order $n \geq 2$. Since T is connected and $n \geq 2$, E is non-empty. By *lemma 3.1*, T contains a leaf v. Let $\{u, v\}$ be the unique edge incident to v. Consider the graph $T' = T - v$ (removing v and its incident edge).

Acyclicity: Removing vertices/edges cannot create cycles. T' is acyclic.

Connectivity: Since v was a leaf, it was not an internal node of any path between two other vertices $x, y \in V \setminus \{v\}$. Thus T' remains connected.

Therefore, T' is a tree of order $n - 1$. By the inductive hypothesis, $|E(T')| = (n-1) - 1 = n - 2$. The edge set of T is $E(T') \cup \{\{u, v\}\}$, so $|E(T)| = (n-2) + 1 = n - 1$.

証明終

Corollary 3.1. *Degrees in Trees.* In any tree of order $n \geq 2$, $\sum_{v \in V} d(v) = 2n - 2$.

推論

Proof

Combine the previous theorem with the *theorem* 0.1.

■

Characterisations

The definition of a tree (acyclic and connected) is merely one of many equivalent ways to specify this structure. The following theorem asserts that any two of the following properties (connectedness, acyclicity, and size $n-1$), imply the third (mostly). Furthermore, trees are precisely the graphs that are "minimally connected" or "maximally acyclic".

Theorem 3.2. ***The Big Theorem on Trees.***
Let $G = (V, E)$ be a graph of order n. The following statements are equivalent:

1. G is a tree (connected and acyclic).
2. For every pair $u, v \in V$, there exists a **unique** path connecting them.
3. G is connected and $|E| = n - 1$.
4. G is acyclic and $|E| = n - 1$.
5. G is connected, but removing any edge renders it disconnected (minimally connected).
6. G is acyclic, but adding any edge between non-adjacent vertices creates a cycle (maximally acyclic).

定理

Proof

We establish the cycle of implications.

(1) $\implies$ ***(2):*** Connectivity implies existence. For uniqueness, suppose two distinct paths P_1, P_2 connect u and v. The symmetric difference of their edge sets must contain a cycle (see *lemma* 2.1 intuition: diverge at some point x and reconverge at y), contradicting acyclicity.

(2) $\implies$ ***(1):*** Existence implies connectivity. If G contained a cycle, any two vertices on that cycle would be connected by at least two paths (the two arcs of the cycle), contradicting uniqueness. Thus G is acyclic.

(1) $\implies$ ***(3):*** Proven in the previous section.

(3) $\implies$ ***(4):*** Suppose G is connected with $n - 1$ edges. If G contained a cycle, we could remove an edge from that cycle without destroying connectivity (the "detour" remains). Repeating this until G is acyclic yields a tree T with $V(T) = V(G)$ and $E(T) \subsetneq E(G)$. But a tree on n vertices must have $n-1$ edges. Since G already has $n-1$ edges, no edges could be removed. Thus G was already acyclic.

(4) $\implies$ (1): Suppose G is acyclic with $n-1$ edges. Let k be the number of connected components $G_1, \dots, G_k$. Each G_i is a tree of order n_i. By the theorem on tree edges, $|E(G_i)| = n_i - 1$. Summing over components:

$$|E| = \sum_{i=1}^{k}(n_i - 1) = \left(\sum n_i\right) - k = n - k.$$

Given $|E| = n-1$, we have $n - k = n - 1 \implies k = 1$. Thus G is connected.

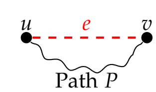
(1) $\implies$ (5): G is connected. Let $e = \{u, v\}$. Since G is acyclic, there is no path between u and v other than the edge e itself (otherwise e plus that path would form a cycle). Removing e eliminates the only path between u and v, disconnecting them.

(5) $\implies$ (1): G is connected. If G had a cycle, removing an edge from that cycle would preserve connectivity. Since removing *any* edge disconnects G, no cycles can exist.

(1) $\implies$ (6): G is acyclic. Let u, v be non-adjacent. Since G is connected, there is a path P between them. Adding $e = \{u, v\}$ closes this path into a cycle.

(6) $\implies$ (1): G is acyclic. Suppose G is disconnected. Let u, v be in different components. Adding $e = \{u, v\}$ connects two components but cannot create a cycle (a cycle requires entering and leaving a component, implying two edges crossed the gap). This contradicts (6). Thus G is connected.

■



Adding e creates cycle $P \cup \{e\}$

Figure 3.3: The Maximally Acyclic property (6). If a unique path exists between u and v, adding an edge closes the loop.

Isomorphism with Linear Algebra

The behaviour of trees is strikingly similar to the behaviour of bases in vector spaces. This is not a coincidence; it is the foundation of algebraic graph theory and matroid theory.

Recall that for a vector space W of dimension n, a set of vectors B is a basis if and only if it generates W and is linearly independent. The parallels are exact:

Vector Space	Graph Theory
Vector Space W	Complete Graph K_n
Vector v	Edge e
Linear Independence	Acyclic (Forest)
Generating / Spanning	Connected
Basis	Spanning Tree
Dimension n	Spanning tree has $n-1$ edges

Table 3.1: Structural analogy between Linear Algebra and Graph Theory.

Remark.

Compare *theorem 3.2* with the standard characterisation of a basis B in a vector space of dimension n:

- B is a basis (Independent + Generating).
- Every vector has a **unique** representation as a linear combination of B.
- B is generating and $|B| = n$.
- B is independent and $|B| = n$.
- B is minimally generating (removing any vector destroys the span).
- B is maximally independent (adding any vector creates a dependency).

The graph theoretic "dimension" of a connected graph on n vertices is $n-1$.

Spanning Trees

The correspondence above suggests that every connected graph contains a tree that "spans" the vertices.

Definition 3.3. ***Spanning Tree.***
Let $G = (V, E)$ be a connected graph. A subgraph $T = (V, E')$ is a **spanning tree** of G if T is a tree and $E' \subseteq E$.

定義

Proposition 3.1. ***Existence of Spanning Trees.***
Every connected graph G contains a spanning tree.

命題

Proof

If G is acyclic, G itself is a tree. If G contains a cycle, remove an edge from that cycle. The graph remains connected. Repeat this process until no cycles remain. The resulting subgraph is connected, acyclic, and contains all vertices of G, hence it is a spanning tree.

■

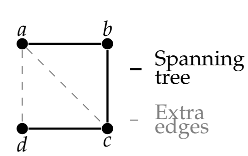


Figure 3.4: A graph G with 5 edges. The spanning tree (solid) has $n-1 = 3$ edges; the 2 dashed edges form the fundamental cycles.

This leads to a numerical invariant for graphs, similar to the dimension of the null space.

> **Definition 3.4.** ***Cyclomatic Number.***
> Let $G = (V, E)$ be a connected graph with n vertices and m edges. The **cyclomatic number** (or cycle rank) is:
>
> $$\nu(G) = m - n + 1.$$
>
> This integer counts the number of "fundamental cycles" in G. Relative to any spanning tree T, adding one of the $\nu(G)$ edges in $E \setminus E(T)$ creates exactly one unique cycle.
>
> 定義

3.2 *Enumeration of Trees*

Having characterised the structure of trees, we turn to the problem of enumeration. Specifically, we seek to determine the number of distinct trees that can be formed on a fixed set of n labelled vertices, say $V = [1, n]$. We denote this quantity by $t(n)$.
For small values of n, direct enumeration yields:
- $n = 1$: 1 tree (single vertex).
- $n = 2$: 1 tree (edge $\{1, 2\}$).
- $n = 3$: 3 trees (path $1 - 2 - 3$, $2 - 1 - 3$, or $1 - 3 - 2$; central vertex determines the tree).
- $n = 4$: 16 trees.

The sequence suggests the closed form n^{n-2}, a result famously attributed to Cayley.

> **Theorem 3.3.** ***Cayley's Formula.***
> For any integer $n \geq 1$, the number of distinct trees on the vertex set $[1, n]$ is:
> $$t(n) = n^{n-2}.$$
>
> 定理

A direct inductive proof of this formula is difficult because removing a vertex from a tree usually results in a forest, complicating the recurrence. Instead, we prove a stronger result that accounts for the specific degrees of the vertices. This allows us to control the structure during the inductive step by "pruning" leaves.

Trees with Fixed Degrees

Let $D = (d_1, \ldots, d_n)$ be a sequence of positive integers. We denote by $t(d_1, \ldots, d_n)$ the number of trees on the vertex set $[1, n]$ such that the degree of vertex i is exactly d_i.

A necessary condition for such trees to exist is given by the Handshaking Lemma (*theorem 0.1*) and the edge count of trees (*theorem 3.2*):

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

Theorem 3.4. ***Refined Cayley Formula.***
Let $n \geq 2$ and let $d_1, \ldots, d_n$ be positive integers such that $\sum_{i=1}^{n} d_i = 2n-2$. The number of trees on $[1, n]$ with degree sequence $(d_1, \ldots, d_n)$ is given by the multinomial coefficient:

$$t(d_1, \ldots, d_n) = \frac{(n-2)!}{(d_1-1)!(d_2-1)! \ldots (d_n-1)!}.$$

定理

We proceed by induction on n.

Base Case ($n = 2$)

The condition $\sum d_i = 2(2) - 2 = 2$ with $d_i \geq 1$ implies $d_1 = 1$ and $d_2 = 1$. There is exactly one tree on two vertices (the single edge). The formula yields $\frac{0!}{0!0!} = 1$. The base case holds.

証明終

Inductive Step

Assume the formula holds for all sequences of length $n - 1$ summing to $2(n - 1) - 2$. Consider a sequence $(d_1, \ldots, d_n)$ summing to $2n - 2$. By *corollary 3.1*, any tree with this degree sequence must have at least two leaves. Thus, at least one d_i must be equal to 1. Without loss of generality (by reordering if necessary), assume $d_n = 1$.

In any tree realising this sequence, vertex n is a leaf. Let j be the unique neighbour of n. Since n is connected to j, removing n yields a tree on the vertex set $[1, n - 1]$. In this smaller tree T', the degrees are:

$$d'_k = \begin{cases} d_k & \text{if } k \neq j, \\ d_j - 1 & \text{if } k = j. \end{cases}$$

Note that $d_j \geq 1$, and if $d_j = 1$, vertex j would be a leaf in the original tree connected only to n, implying $n = 2$ (which is covered by the base case). For $n > 2$, if $d_j = 1$, j becomes an isolated vertex in T', which is impossible for a tree. Thus $d_j \geq 2$ implies $d'_j \geq 1$. The sum of degrees in T' is:

$$\sum_{k=1}^{n-1} d'_k = \left(\sum_{k=1}^{n} d_k \right) - d_n - 1 = (2n-2) - 1 - 1 = 2(n-1) - 2.$$

Thus, the inductive hypothesis applies to T'.

The set of trees with degrees $(d_1, \ldots, d_n)$ can be partitioned based on the neighbour j of the leaf n. The possible values for j are indices $k \in [1, n-1]$ such that $d_k \geq 2$ (since $d_k - 1$ must be valid). However, if $d_k = 1$, the term $(d_k - 2)!$ in the denominator would be undefined (or effectively zero contribution), so we may formally sum over all $j \in [1, n-1]$. Using the addition principle:

$$t(d_1, \ldots, d_n) = \sum_{j=1}^{n-1} t(d_1, \ldots, d_j - 1, \ldots, d_{n-1}).$$

Substituting the inductive formula:

$$\begin{aligned} t(d_1, \ldots, d_n) &= \sum_{j=1}^{n-1} \frac{(n-3)!}{(d_1-1)! \ldots (d_j-2)! \ldots (d_{n-1}-1)!} \\ &= \frac{(n-3)!}{\prod_{k=1}^{n-1}(d_k-1)!} \sum_{j=1}^{n-1}(d_j - 1). \end{aligned}$$

We compute the sum in the final term:

$$\sum_{j=1}^{n-1}(d_j - 1) = \left(\sum_{j=1}^{n-1} d_j\right) - (n-1).$$

Since $\sum_{i=1}^{n} d_i = 2n - 2$ and $d_n = 1$, we have $\sum_{j=1}^{n-1} d_j = 2n - 3$.

$$\sum_{j=1}^{n-1}(d_j - 1) = (2n-3) - (n-1) = n - 2.$$

Substituting back:

$$t(d_1, \ldots, d_n) = \frac{(n-3)!}{\prod_{k=1}^{n-1}(d_k-1)!} \cdot (n-2) = \frac{(n-2)!}{\prod_{k=1}^{n-1}(d_k-1)!}.$$

Since $d_n = 1$, $(d_n - 1)! = 0! = 1$. We may include it in the denominator to recover the symmetric form:

$$t(d_1, \ldots, d_n) = \frac{(n-2)!}{(d_1-1)! \ldots (d_n-1)!}.$$

証明終

Recovering Cayley's Formula

To find the total number of trees $t(n)$, we sum $t(d_1, \ldots, d_n)$ over all valid degree sequences. This summation is handled elegantly by the Multinomial Theorem.

Lemma 3.2. ***Multinomial Theorem.***
For any integer $m \geq 0$ and variables $x_1, \ldots, x_k$:

$$(x_1 + \cdots + x_k)^m = \sum_{\substack{a_1+\cdots+a_k=m \\ a_i \geq 0}} \frac{m!}{a_1! \ldots a_k!} x_1^{a_1} \ldots x_k^{a_k}.$$

引理

Proof

Consider the expansion of the product:

$$\underbrace{(x_1 + \cdots + x_k) \cdot (x_1 + \cdots + x_k) \ldots (x_1 + \cdots + x_k)}_{m \text{ factors}}.$$

To form a term in the product, we must select exactly one variable x_j from each of the m factors. Suppose we select x_1 exactly a_1 times, x_2 exactly a_2 times, ..., and x_k exactly a_k times. Since we make exactly m selections in total, we must have $\sum_{i=1}^{k} a_i = m$, where each $a_i \geq 0$. The resulting term is the product $x_1^{a_1} x_2^{a_2} \ldots x_k^{a_k}$.
The coefficient of this term corresponds to the number of ways to assign these specific counts to the m distinct positions (factors).
This is equivalent to partitioning the set of m positions into k disjoint subsets of sizes $a_1, a_2, \ldots, a_k$. The number of such partitions is given by the multinomial coefficient:

$$\binom{m}{a_1, a_2, \ldots, a_k} = \frac{m!}{a_1! a_2! \ldots a_k!}.$$

Summing over all possible non-negative integer solutions to $a_1 + \cdots + a_k = m$ yields the full expansion.

■

Proof of Cayley's Formula

The total number of trees is the sum of $t(d_1, \ldots, d_n)$ over all tuples $(d_1, \ldots, d_n)$ satisfying $d_i \geq 1$ and $\sum d_i = 2n - 2$. Let $k_i = d_i - 1$. The conditions transform to $k_i \geq 0$ and:

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

Using *theorem* 3.4:

$$t(n) = \sum_{\substack{k_1+\cdots+k_n=n-2 \\ k_i \geq 0}} \frac{(n-2)!}{k_1! \ldots k_n!}.$$

This is precisely the coefficient expansion of the Multinomial Theorem (*lemma* 3.2) with $m = n - 2$ and $x_1 = x_2 = \cdots = x_n = 1$:

$$t(n) = (1 + 1 + \cdots + 1)^{n-2} = n^{n-2}.$$

■

3.3 *Prüfer Sequences*

There exists a more direct method to encode any tree on $[1, n]$ as a sequence of length $n-2$ with elements in $[1, n]$. This encoding, known as the **Prüfer sequence**, provides a constructive proof of *Cayley's Formula*.

Definition 3.5. ***Prüfer Sequence Construction.***
Let T be a tree on vertices $[1, n]$ with $n \geq 2$. We generate a sequence $P = (p_1, \ldots, p_{n-2})$ iteratively:

1. Find the leaf with the smallest label. Let this be u.
2. Let v be the unique neighbour of u.
3. Record v as the next element in the sequence.
4. Remove u from the tree.
5. Repeat this process $n-2$ times.

The resulting sequence of length $n-2$ is the Prüfer sequence of T. If $n = 2$, the sequence is empty.

定義

Example 3.1. Calculating a Prüfer Sequence. Consider a tree on $\{1, 2, 3, 4, 5, 6\}$ with edges:

$$E = \{\{1,4\}, \{2,4\}, \{3,4\}, \{4,5\}, \{5,6\}\}.$$

1. Leaves are $\{1, 2, 3, 6\}$. Smallest is 1. Neighbour is 4. Sequence: (4). Remove 1.
2. Leaves are $\{2, 3, 6\}$. Smallest is 2. Neighbour is 4. Sequence: $(4, 4)$. Remove 2.
3. Leaves are $\{3, 6\}$. Smallest is 3. Neighbour is 4. Sequence: $(4, 4, 4)$. Remove 3.
4. Leaves are $\{4, 6\}$. Smallest is 4. Neighbour is 5. Sequence: $(4, 4, 4, 5)$. Remove 4.
5. Remaining edge is $\{5, 6\}$. Stop.

The sequence is $(4, 4, 4, 5)$.

範例

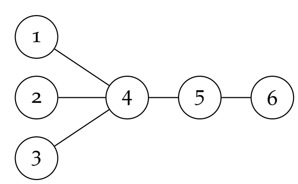


Figure 3.5: The tree from the Prüfer example. Vertex 4 appears thrice in the sequence $(4, 4, 4, 5)$, indicating $d(4) = 1 + 3 = 4$.

The fundamental utility of this encoding lies in its relationship with vertex degrees.

Claim 3.1. Degree Property. Let T be a tree on $[1, n]$ with Prüfer sequence P.
For any vertex v, the degree $d(v)$ in T is equal to $1 + (\text{number of occurrences of } v \text{ in } P)$.

主張

Proof

We proceed by induction on n. For $n = 2$, the sequence is empty. Both vertices have degree 1, and $1 + 0 = 1$. For $n \geq 3$, let u be the smallest leaf removed in the first step, and let v be its neighbour. The sequence P begins with v, followed by the Prüfer sequence P' of the tree $T' = T - u$.

For u: Since u is a leaf, $d(u) = 1$. Since u was removed, it never appears as a neighbour of a subsequently removed leaf. Thus u appears 0 times in P. $1 + 0 = 1$.

For $x \neq u, v$: The degree of x is unchanged in T'. By the inductive hypothesis, $d_{T'}(x) = 1 + (\text{count in } P')$. Thus $d_T(x) = 1 + (\text{count in } P)$.

For v: $d_T(v) = d_{T'}(v) + 1$. By induction, $d_{T'}(v) = 1 + (\text{count in } P')$. Thus $d_T(v) = 1 + ((\text{count in } P') + 1) = 1 + (\text{count in } P)$.

■

This property allows us to establish the bijection.

Theorem 3.5. ***Prüfer Bijection.***
For $n \geq 2$, the mapping $T \mapsto P(T)$ is a bijection between the set of trees on $[1, n]$ and the set of sequences of length $n-2$ with elements in $[1, n]$.

定理

Proof

Injectivity: Let T_1, T_2 be distinct trees. If their degree sequences differ, their Prüfer sequences differ (by *claim* 3.1). If their degree sequences are identical, they share the same set of leaves. Let u be the smallest leaf. If u has different neighbours in T_1 and T_2, the first term of the sequences differs. If u has the same neighbour v, we remove u to get T_1', T_2'. Since $T_1 \neq T_2$, we must have $T_1' \neq T_2'$ (otherwise adding edge $\{u, v\}$ would yield identical trees). By induction, $P(T_1') \neq P(T_2')$, so $P(T_1) \neq P(T_2)$.

Surjectivity: Let $S = (s_1, \ldots, s_{n-2})$ be a sequence. We reconstruct T. Let L be the set of labels $[1, n]$. At step i (from 1 to $n - 2$), let u be the smallest element in the current set of available labels that *does not* appear in the remaining sequence suffix $(s_i, \ldots, s_{n-2})$. Add edge $\{u, s_i\}$ and remove u from the available labels. After $n - 2$ steps, exactly two labels remain. Join them with an edge. This process constructs a tree whose Prüfer sequence is S.

■

Since there are n^{n-2} such sequences, this confirms *Cayley's Formula*. Moreover, counting sequences where specific numbers appear specific times recovers *Refined Cayley Formula*.

Otter's Formula

Throughout this chapter, we have counted *labelled* trees, where the identity of the vertices matters. If we consider trees up to isomorphism (ignoring labels), the problem becomes significantly harder. For $n = 4$, there are 16 labelled trees but only 2 isomorphism classes: the path P_3 and the star $K_{1,3}$. Let $\tilde{t}(n)$ denote the number of unlabelled trees on n vertices. The sequence begins $1, 1, 1, 2, 3, 6, 11, 23, \ldots$ and has no simple closed form. However, its asymptotic behaviour is known.

> **Theorem 3.6.** ***Otter's Formula (1948).***
> The number of unlabelled trees on n vertices satisfies:
>
> $$\tilde{t}(n) \sim \beta \cdot \alpha^n \cdot n^{-5/2},$$
>
> where $\alpha \approx 2.95576$ and $\beta \approx 0.53494$.

定理

3.4 Exercises

1. **Small Trees.**
   (a) Draw two non-isomorphic trees on 4 vertices. For each, verify that $|E| = 4 - 1 = 3$.
   (b) Draw a tree on 6 vertices and identify all its leaves.
2. **Forests.** Let F be a forest with k connected components and n vertices. Prove that $|E(F)| = n - k$.

   > *Remark.*
   >
   > Sum the edge counts of each tree component.
3. **Leaves.**
   (a) Prove directly (using a longest path argument) that every tree with at least one edge has at least two leaves.
   (b) Construct a tree on $n \geq 3$ vertices with exactly 2 leaves. Construct one with $n - 1$ leaves.
4. **Pruning.** Let T be a tree and v a leaf. Prove that $T - v$ is a tree.
5. **Inductive Proofs.**
   (a) Prove by induction on n that $|E(T)| = n - 1$ for any tree of

order n.

(b) Prove that $\sum_{v \in V} d(v) = 2n - 2$ for any tree of order $n \geq 2$.

6. **Characterisations.** Let G be a connected graph with no cycles.
   (a) Prove that between any two vertices, there is a unique path.
   (b) Show that adding any edge between non-adjacent vertices creates exactly one cycle.
7. **Spanning Trees.**
   (a) Prove that every connected graph contains a spanning tree by repeatedly deleting cyclic edges.
   (b) Let G be connected with n vertices and m edges. Let T be a spanning tree. Show that $|E(G) \setminus E(T)| = m - n + 1$ (the cyclomatic number).
8. **Refined Cayley Formula.**
   (a) Use the formula to find the number of trees on $\{1, \ldots, 5\}$ with degrees $(3, 2, 1, 1, 1)$.
   (b) How many trees on $\{1, \ldots, n\}$ are stars (one vertex of degree $n - 1$, others degree 1)? Verify using the formula.
9. **Prüfer Sequences.**
   (a) Compute the Prüfer sequence for the path $1 - 2 - 3 - 4 - 5 - 6$.
   (b) Compute the Prüfer sequence for the star $K_{1,5}$ with centre 1.
   (c) Decode the sequence $(1, 1, 1, 1)$ on vertices $\{1, \ldots, 6\}$.
   (d) Decode the sequence $(2, 3, 2, 3)$ on vertices $\{1, \ldots, 6\}$.
10. **Leaves via Prüfer.** Prove that the leaves of a tree T are exactly the labels from $\{1, \ldots, n\}$ that do *not* appear in its Prüfer sequence. Deduce that every tree ($n \geq 2$) has at least two leaves.
11. $\star$ **Direct Counting.** Use the property of Prüfer sequences (label i appears $d_i - 1$ times) to derive the Refined Cayley Formula directly from the number of permutations of a multiset.
12. **Unlabelled Trees.** List all non-isomorphic trees on $n = 5$ vertices. Compare the count (3) with the number of labelled trees ($5^3 = 125$).

4
Colouring

In the chapter on Extremal Graph Theory, we introduced the concept of an **??**. We now generalise this notion by asking whether the vertex set of a graph can be partitioned entirely into independent sets. This process, known as **colouring**, assigns labels to vertices such that adjacent vertices receive distinct labels.
This framework models problems of resource allocation and conflict resolution, but mathematically, it provides a rigorous way to classify graphs based on their local structural constraints.

4.1 Vertex Colouring

Definition 4.1. ***Colouring and Chromatic Number.***
Let $G = (V, E)$ be a graph and C be a set of labels, called **colours**. A function $f : V \to C$ is a **proper colouring** if for every edge $\{u, v\} \in E$, we have $f(u) \neq f(v)$. The **chromatic number** of G, denoted $\chi(G)$, is the minimum cardinality of C such that a proper colouring exists. If $\chi(G) \leq k$, we say G is k-**colourable**.

定義

Note

If f is a proper colouring with k colours, the preimages $f^{-1}(c)$ for each $c \in C$ form a partition of V into k independent sets, often called **colour classes**.

We immediately observe the values for standard families defined in previous chapters.

Proposition 4.1. ***Elementary Chromatic Numbers.***

1. $\chi(G) = 1$ if and only if $E = \varnothing$ (i.e., G is an independent set).
2. $\chi(K_n) = n$. Since every vertex is adjacent to every other, no two can share a colour.
3. For the path graph P_n ($n \geq 2$), $\chi(P_n) = 2$.

4. For the cycle graph C_n:

$$\chi(C_n) = \begin{cases} 2 & \text{if } n \text{ is even,} \\ 3 & \text{if } n \text{ is odd.} \end{cases}$$

5. For the hypercube Q_d, $\chi(Q_d) = 2$. Vertices can be coloured by the parity of their Hamming weight.

命题

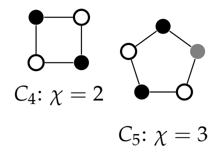


Figure 4.1: Cycle colouring. C_4 alternates two colours. C_5 requires a third colour (gray) to close the cycle without conflict.

The Greedy Algorithm

Determining $\chi(G)$ is generally NP-hard. However, we can obtain upper bounds via a constructive approach known as the **Greedy Algorithm**. This method orders the vertices and assigns the smallest available colour index to each.

Algorithm: Let $V = \{v_1, \ldots, v_n\}$ be an ordering of the vertices. We map $f : V \to \mathbb{Z}^+$. For $i = 1$ to n:

1. Identify the set of colours used by the neighbours of v_i that precede it in the ordering:

$$C_i = \{f(v_j) : \{v_i, v_j\} \in E \text{ and } j < i\}.$$

2. Assign $f(v_i) = \min\{c \in \mathbb{Z}^+ : c \notin C_i\}$.

The efficiency of this algorithm depends heavily on the chosen vertex ordering. While there exists an ordering that produces exactly $\chi(G)$ colours, finding it is difficult. However, the worst-case performance provides a fundamental bound based on the maximum degree $\Delta(G)$.

Theorem 4.1. ***The Greedy Bound.***
For any graph G,

$$\chi(G) \leq \Delta(G) + 1.$$

定理

Proof

Let vertices be ordered arbitrarily as $v_1, \ldots, v_n$. When the algorithm considers vertex v_i, it examines its neighbours. The number of neighbours preceding v_i is at most the total degree $d(v_i)$, which is bounded by $\Delta(G)$. Thus, the set of forbidden colours C_i has size $|C_i| \leq \Delta(G)$. The set $\{1, \ldots, \Delta(G) + 1\}$ contains more elements than C_i, guaranteeing that at least one colour in this range is available for v_i.

■

4.2 *Bipartite Graphs*

The class of graphs with $\chi(G) \leq 2$ merits special attention. These are the **bipartite graphs**.

Definition 4.2. ***Bipartite Graph.***
A graph $G = (V, E)$ is **bipartite** if it is 2-colourable. Equivalently, V admits a partition $V = A \sqcup B$ such that every edge has one endpoint in A and one in B. The sets A and B are the **parts** of the bipartition.
定義

Examples include the complete bipartite graph $K_{n,m}$ (defined in Extremal Graph Theory), the hypercubes Q_d, and all trees.

Characterisation by Cycles

A triangle (K_3 or C_3) requires 3 colours. Intuitively, any odd cycle creates a parity conflict that prevents 2-colouring. It turns out this is the *only* obstruction.

Theorem 4.2. ***Bipartite Characterisation.***
A graph G is bipartite if and only if it contains no cycle of odd length.

定理

($\implies$)

Let G be bipartite with partition $V = A \sqcup B$. Consider a cycle $C = (v_0, v_1, \ldots, v_k = v_0)$. Without loss of generality, let $v_0 \in A$. Since edges only connect distinct parts, $v_1 \in B$, $v_2 \in A$, and by induction, $v_i \in A$ if i is even, and $v_i \in B$ if i is odd. For the cycle to close at $v_k = v_0 \in A$, the index k must be even. Thus, the length of the cycle is even.
証明終

($\impliedby$)

Assume G contains no odd cycles. It suffices to consider a single connected component (as G is bipartite if and only if all components are). Fix a base vertex $s \in V$. We partition V based on the parity of path lengths from s. Define:

$$X = \{v \in V : \exists \text{ a path between } s \text{ and } v \text{ of even length}\},$$
$$Y = \{v \in V : \exists \text{ a path between } s \text{ and } v \text{ of odd length}\}.$$

Since G is connected, $V = X \cup Y$. We must show that $X \cap Y = \varnothing$ and that no edges exist within X or within Y.
Suppose $v \in X \cap Y$. Then there exists an even path P_{even} from s to v and an odd path P_{odd} from s to v. The concatenation of P_{even}

and the reversal of P_{odd} forms a closed walk starting and ending at s with length equal to the sum of an even and an odd integer, which is odd. From the chapter on Connectivity, we know a closed walk reduces to a set of cycles. If a closed walk has odd length, it must contain at least one cycle of odd length (since a sum of even integers is even). This contradicts the hypothesis. Thus $X \cap Y = \varnothing$, and $V = X \sqcup Y$ is a valid partition.
Finally, consider an edge $\{u, v\} \in E$. If $u \in X$, there is an even path from s to u. Extending this path to v yields a walk of odd length, so $v \in Y$ (and specifically $v \notin X$ by the disjointness proved above). Similarly, if $u \in Y$, then $v \in X$. Thus, edges only connect X and Y, making G bipartite.

証明終

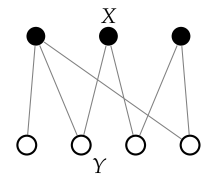


Figure 4.2: A bipartite graph with 2-colouring. Filled vertices form one colour class X, hollow vertices form Y. All edges connect distinct classes.

This theorem provides an efficient algorithm for checking bipartiteness: perform a Breadth-First Search (BFS). If we encounter an edge between two vertices at the same layer (distance from root), an odd cycle exists, and the graph is not bipartite. Otherwise, the layers form the sets X and Y.

4.3 *Matchings*

We now consider the dual problem: selecting a subset of edges such that no two share a vertex. This concept, known as a **matching**, models pairings in a population, such as job assignments or chemical bonding.

Definition 4.3. ***Matching.***
Let $G = (V, E)$ be a graph. A subset $M \subseteq E$ is a **matching** if no two edges in M share a common vertex. A vertex $v \in V$ is **saturated** by M if it is an endpoint of some edge in M; otherwise, it is **unsaturated**. A matching is **perfect** if it saturates every vertex in V.

定義

Example 4.1. Matchings in Complete Structures.

1. **Complete Bipartite Graph $K_{n,n}$:** Let the partition be $X \sqcup Y$ with $|X| = |Y| = n$. A perfect matching corresponds to a bijection $\sigma : X \to Y$. Thus, $K_{n,n}$ admits $n!$ distinct perfect matchings.
2. **Complete Graph K_{2n+1}:** Since the order is odd, no matching can saturate all vertices. A perfect matching is impossible.
3. **Complete Graph K_{2n}:** We construct a perfect matching by choosing a partner for the first vertex ($2n - 1$ choices), then a partner for the next available vertex ($2n - 3$ choices), and so on. The

number of perfect matchings is the double factorial:

$$(2n-1)!! = (2n-1)(2n-3)\cdots 3 \cdot 1 = \frac{(2n)!}{2^n n!}.$$

範例

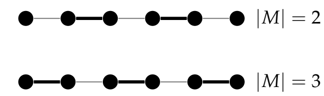


Figure 4.3: Two matchings on P_6. Top: maximal (no edge can be added) but not maximum. Bottom: maximum (perfect) matching.

Maximality and Augmentation

We distinguish between two notions of "largest" matchings.

Definition 4.4. ***Maximal vs Maximum.***
- A matching M is **maximal** if no edge can be added to it without violating the matching property (i.e., it is not a proper subset of another matching).
- A matching M is **maximum** if it has the largest possible cardinality among all matchings in G.

定義

Clearly, every maximum matching is maximal, but the converse is false. To systematically improve a matching, we look for paths that alternate between being "in" and "out" of the matching.

Definition 4.5. ***Alternating and Augmenting Paths.***
Let M be a matching in G. An **alternating path** is a path in G whose edges alternate between $E \setminus M$ and M. An **augmenting path** is an alternating path that starts and ends at distinct unsaturated vertices.

定義

If an augmenting path P exists, we can swap the edges along P: those in M leave the matching, and those not in M enter it. Since P starts and ends with edges not in M, the new set $M' = M \oplus E(P)$ is a valid matching with $|M'| = |M| + 1$. This observation is the "easy" direction of Berge's Theorem.

Theorem 4.3. ***Berge's Theorem (1957).***
A matching M in G is maximum if and only if G contains no augmenting path with respect to M.

定理

$(\Longrightarrow)$

Suppose G contains an augmenting path P. As described above, the symmetric difference $M' = M \oplus E(P)$ is a matching with cardinality $|M| + 1$. Thus M was not maximum.

証明終

$(\Longleftarrow)$

We prove the contrapositive. Suppose M is not maximum. Let M^*

be a maximum matching, so $|M^*| > |M|$. Consider the graph $H = (V, M \oplus M^*)$ induced by the symmetric difference of the edge sets. The maximum degree in H is 2, since every vertex is incident to at most one edge from M and one edge from M^*. Consequently, the connected components of H are either isolated vertices, paths, or cycles.

Cycles: Must be of even length, alternating between M and M^*. They contain an equal number of edges from both sets.

Paths: Must alternate between M and M^*.

Since $|M^*| > |M|$, there must be at least one component in H with strictly more edges from M^* than from M. Cycles have equal counts, so this component must be a path P. For P to have more M^* edges than M edges, it must start and end with an edge from M^*. This implies the endpoints of P are saturated by M^* but not by M (in the context of H, and thus in G relative to M). Therefore, P is an augmenting path for M.

証明終

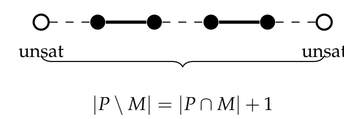


Figure 4.4: An augmenting path. Solid edges are in M, dashed edges are not. Hollow vertices are unsaturated. Swapping edge membership yields $|M| + 1$.

Hall's Marriage Theorem

For bipartite graphs, the existence of specific matchings is governed by neighbour sets. Let $G = (A \sqcup B, E)$ be bipartite. For a subset $U \subseteq A$, let $N(U)$ denote the set of neighbours of vertices in U. If a matching saturates A, then every vertex in A is mapped to a distinct vertex in B. A necessary condition is therefore $|N(U)| \geq |U|$ for all subsets $U \subseteq A$. Hall proved this is also sufficient.

Theorem 4.4. ***Hall's Marriage Theorem (1935).***
Let $G = (A \sqcup B, E)$ be a bipartite graph. There exists a matching that saturates A if and only if:

$$\forall U \subseteq A, \quad |N(U)| \geq |U|.$$

定理

Proof

The condition is clearly necessary. We prove sufficiency by contradiction. Assume the condition holds ($|N(U)| \geq |U|$ for all U), but G admits no matching saturating A. Let M be a **maximum** matching. By assumption, M does not saturate A, so there exists an unsaturated vertex $s \in A$.
We construct the set of vertices reachable from s via alternating

paths. Let:

$$Z = \{v \in V : \exists \text{ an alternating path from } s \text{ to } v\}.$$

Let $U = Z \cap A$ and $V = Z \cap B$. Note that $s \in U$.

Claim 1: $N(U) \subseteq V$**.** Let $u \in U$ and $v \in N(u)$. If $\{u, v\} \in M$, then v lies on the alternating path to u, so $v \in Z \cap B = V$. If $\{u, v\} \notin M$, then extending the alternating path ending at u (which must end with an edge in M or be s) by the edge $\{u, v\}$ creates a valid alternating path to v. Thus $v \in V$.

Claim 2: V ***is matched into*** $U \setminus \{s\}$**.** Let $v \in V$. Since v is reachable from s by an alternating path starting with a non-matching edge, the path enters v via a non-matching edge. If v were unsaturated, the path from s to v would be augmenting, contradicting the maximality of M (by Berge's Theorem). Thus, v must be saturated by some edge $e \in M$. Let $e = \{v, u'\}$. The path can be extended through e to u', so $u' \in Z \cap A = U$. Moreover $u' \neq s$ since s is unsaturated. This defines a bijection between V and a subset of $U \setminus \{s\}$ formed by the edges of M.

From Claim 2, we have $|V| = |U \setminus \{s\}| = |U| - 1$. From Claim 1, we have $N(U) \subseteq V$, so $|N(U)| \leq |V|$. Combining these:

$$|N(U)| \leq |U| - 1 < |U|.$$

This contradicts the Hall condition for the set U. Thus, a matching saturating A must exist.

■

Corollary 4.1. *Regular Bipartite Graphs.* For $k > 0$, every k-regular bipartite graph admits a perfect matching.

推論

Proof

Let $G = (A \sqcup B, E)$ be k-regular. First, we show $|A| = |B|$. Counting edges via A: $|E| = \sum_{v \in A} d(v) = k|A|$. Counting edges via B: $|E| = \sum_{v \in B} d(v) = k|B|$. Thus $k|A| = k|B| \implies |A| = |B|$. A matching saturating A will therefore be perfect.

We check Hall's condition. Let $U \subseteq A$. Let E_U be the set of edges incident to U. Since degrees are k, $|E_U| = k|U|$. These edges must be incident to $N(U)$. The sum of degrees in $N(U)$ is $k|N(U)|$. Since all edges in E_U land in $N(U)$, we must have:

$$|E_U| \leq \sum_{v \in N(U)} d(v) = k|N(U)|.$$

Substituting $|E_U| = k|U|$:

$$k|U| \leq k|N(U)| \implies |U| \leq |N(U)|.$$

By Hall's Theorem, a perfect matching exists. ■

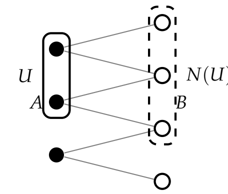


Figure 4.5: Hall's condition: for $U = \{a_1, a_2\} \subseteq A$, the neighbourhood $N(U)$ has $|N(U)| \geq |U|$.

4.4 Exercises

1. **Subgraph Monotonicity.** Let G be a graph and H be a subgraph of G. Prove that $\chi(H) \leq \chi(G)$.

   > *Remark.*
   >
   > Any valid colouring of G induces a valid colouring of H.

2. **Clique Lower Bound.** Prove that if a graph G contains a clique of size r (a subgraph isomorphic to K_r), then $\chi(G) \geq r$.
3. **Critical Graphs.** A graph G is called k**-critical** if $\chi(G) = k$, but for every proper subgraph $H \subset G$, $\chi(H) < k$. Using the logic of the Greedy Bound, prove that if G is k-critical, then the minimum degree of G satisfies $\delta(G) \geq k - 1$.

   > *Remark.*
   >
   > Suppose there is a vertex v with degree $< k - 1$. Consider the graph $G - v$.

4. **Greedy Algorithm Trace.** Draw a graph with 6 vertices (e.g., a cycle with a chord). Apply the Greedy Algorithm using two different vertex orderings. Do they result in the same number of colours?
5. **Ordering Sensitivity.** Construct a specific example of a graph and two different vertex orderings such that the Greedy Algorithm uses a different number of colours for each ordering.

   > *Remark.*
   >
   > Consider a path of length 3 (P_3) or a bipartite graph like P_4 with orderings $1, 2, 3, 4$ vs $1, 3, 2, 4$.

6. **Parity Constraint.** Prove that if a graph G has a perfect matching, then the number of vertices $|V(G)|$ must be even.
7. **Failing Hall's Condition.** Construct a bipartite graph $G = (A \sqcup B, E)$ with $|A| = |B| = 4$ that has no perfect matching. Explicitly identify a subset $U \subseteq A$ that violates Hall's condition (i.e., show a set U where $|N(U)| < |U|$).

8. **Necessity of Hall's Condition.** Write a formal argument explaining why a matching that saturates A cannot exist if there is a subset $U \subseteq A$ with $|N(U)| < |U|$.

> *Remark.*
>
> This is the pigeonhole principle applied to the edges of the matching incident to U.

5 Planarity

Until now, we have treated graphs as abstract structures defined solely by vertex adjacency. However, many practical applications (such as printed circuit board design or cartography), require realising these structures in physical space. We now investigate graphs that can be drawn on a 2-dimensional surface without edges crossing. This chapter bridges combinatorics and topology. While a rigorous treatment requires the Jordan Curve Theorem (a deep result in topology), we will proceed with a level of formality appropriate for algebraic graph theory, accepting the topological foundations as intuitive axioms.

5.1 Plane Drawings

Definition 5.1. ***Planar Graphs.***
A **plane drawing** (or embedding) of a graph G is a representation in $\mathbb{R}^2$ where:

- Vertices are distinct points.
- Edges are simple continuous curves (arcs) connecting their endpoints.
- The intersection of any two distinct edges is empty, except possibly at their endpoints.

A graph is **planar** if it admits a plane drawing.

定義

To formalise the notion of a "curve", one technically relies on continuous injective functions $\gamma : [0,1] \to \mathbb{R}^2$. A **Jordan curve** is such a curve that is closed ($\gamma(0) = \gamma(1)$). The fundamental property governing plane graphs is:

Theorem 5.1. ***Jordan Curve Theorem.***
Every Jordan curve $\mathcal{C}$ in the plane partitions $\mathbb{R}^2 \setminus \mathcal{C}$ into two disjoint connected open sets: the **interior** (which is bounded) and the **exterior** (which is unbounded). The curve $\mathcal{C}$ is the boundary of both.

定理

Consequently, any edge connecting a vertex in the interior to a vertex in the exterior must intersect the boundary curve.

Faces

A plane drawing partitions the plane into disjoint connected regions called **faces**.

> **Definition 5.2.** ***Faces.***
> Let G be a plane graph. The connected components of $\mathbb{R}^2 \setminus G$ are called **faces**. We denote the set of faces by F. There is exactly one unbounded face, called the **external face**.
>
> 定義

The boundary of a face consists of a sequence of edges and vertices. If the graph is connected, this boundary corresponds to a closed walk in G.

> **Lemma 5.1.** ***Edge-Face Incidence.***
> Let e be an edge in a plane graph.
> · If e belongs to a cycle, it separates two distinct faces.
> · If e does not belong to a cycle (i.e., it is a bridge), it lies within a single face (the "slit" in the region).
>
> 引理

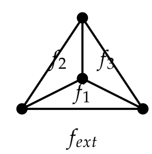


Figure 5.1: A plane embedding of K_4. It has 4 faces (3 internal bounded, 1 external unbounded). This structure corresponds to the projection of a tetrahedron.

5.2 *Euler's Formula*

The fundamental invariant of planar graphs connects the number of vertices (n), edges (m), and faces (f). This relation was first observed for polyhedra.

> **Lemma 5.2.** ***Trees in the Plane.***
> Every tree is a planar graph. Any plane drawing of a tree has exactly one face (the external face).
>
> 引理

> **Theorem 5.2.** ***Euler's Formula.***
> Let G be a connected planar graph with n vertices, m edges, and f faces. Then:
> $$n - m + f = 2.$$
>
> 定理

We proceed by induction on the number of edges m.

> ***Base Case ($m = 0$):***
> The graph consists of a single vertex ($n = 1$). There is one face (the

entire plane).

$$1 - 0 + 1 = 2.$$

The formula holds.

証明終

Inductive Step:

Assume the formula holds for all connected planar graphs with fewer than m edges. Let G have m edges.

Case 1:* G *is a tree. Then $m = n - 1$. Since there are no cycles, there is only 1 face ($f = 1$).

$$n - (n - 1) + 1 = 1 + 1 = 2.$$

Case 2:* G *contains a cycle. Let e be an edge belonging to a cycle. By the incidence lemma, e separates two distinct faces, say F_1 and F_2. Consider the graph $G' = G - e$.

- G' remains connected (since e was on a cycle).
- The number of vertices $n' = n$.
- The number of edges $m' = m - 1$.
- Removing e merges F_1 and F_2 into a single face, so $f' = f - 1$.

By the induction hypothesis:

$$n' - m' + f' = 2 \implies n - (m - 1) + (f - 1) = 2 \implies n - m + f = 2.$$

証明終

Bounds on Edge Density

Euler's formula imposes a strict limit on the number of edges a planar graph can support. To derive this, we count the boundary walks of the faces.

Corollary 5.1. *Planar Edge Density*. Let G be a planar graph with $n \geq 3$ vertices. Then:

$$m \leq 3n - 6.$$

推論

Proof

It suffices to consider G connected, since for fixed n adding components cannot increase m. Let F be the set of faces. For each face $\phi \in F$, let $\ell(\phi)$ be the number of edges bounding ϕ. Since G has no multiple edges and $n \geq 3$, every face (including the external one) must be bounded by at least 3 edges. Thus $\ell(\phi) \geq 3$. Summing over all faces:

$$\sum_{\phi \in F} \ell(\phi) \geq 3f.$$

Each edge bounds at most two faces. Thus, the sum counts every edge at most twice:

$$\sum_{\phi \in F} \ell(\phi) \leq 2m.$$

Combining these inequalities yields $3f \leq 2m$, or $f \leq \frac{2}{3}m$. Substituting this into Euler's formula ($n - m + f = 2$):

$$n - m + \frac{2}{3}m \geq 2 \implies n - \frac{1}{3}m \geq 2 \implies 3n - 6 \geq m.$$

■

This necessary condition allows us to prove non-planarity for dense graphs.

Example 5.1. Non-planarity of K_5. The complete graph K_5 has $n = 5$ and $m = \binom{5}{2} = 10$. Testing the bound:

$$3(5) - 6 = 15 - 6 = 9.$$

Since $10 \not\leq 9$, K_5 is not planar.

範例

For bipartite graphs, the absence of odd cycles strengthens the bound, as the smallest face must be a quadrilateral ($\ell(\phi) \geq 4$).

Corollary 5.2. *Bipartite Planar Bound.* Let G be a planar bipartite graph with $n \geq 3$. Then:

$$m \leq 2n - 4.$$

推論

Proof

Again, we may assume G is connected; extra components only reduce m for fixed n. Similar to the previous proof, we have $4f \leq 2m \implies f \leq \frac{1}{2}m$. Euler's formula gives $n - m + \frac{1}{2}m \geq 2 \implies 2n - 4 \geq m$.

■

Example 5.2. Non-planarity of $K_{3,3}$. The complete bipartite graph $K_{3,3}$ has $n = 6$ and $m = 3 \times 3 = 9$. Testing the bipartite bound:

$$2(6) - 4 = 8.$$

Since $9 \not\leq 8$, $K_{3,3}$ is not planar.

範例

5.3 *Kuratowski's Theorem*

We have identified two fundamental non-planar graphs: K_5 and $K_{3,3}$. It turns out that *any* non-planar graph contains the structure of one of these two.

Definition 5.3. ***Subdivision.***
A **subdivision** of a graph G is a graph obtained by replacing edges of G with paths. Formally, we repeatedly apply the operation of replacing an edge $\{u, v\}$ with $\{u, x\}$ and $\{x, v\}$, where x is a new vertex.
定義

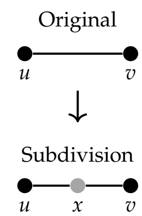


Figure 5.2: Subdivision: an edge $\{u, v\}$ is replaced by a path through a new degree-2 vertex x.

Clearly, if G is non-planar, any subdivision of G is non-planar (adding vertices of degree 2 does not help resolve crossings).

Theorem 5.3. ***Kuratowski's Theorem (1930).***
A graph is planar if and only if it does not contain a subgraph that is a subdivision of K_5 or $K_{3,3}$.
定理

The rigorous proof of this theorem is lengthy, but the intuition is that K_5 and $K_{3,3}$ represent the minimal "obstructions" to planarity. Informally, consider a Hamiltonian cycle in $K_{3,3}$ or K_5. Any chords connecting vertices on the cycle must be drawn inside or outside. In these specific graphs, the chords form an "incompatible" system where no valid assignment of inside/outside prevents all crossings.

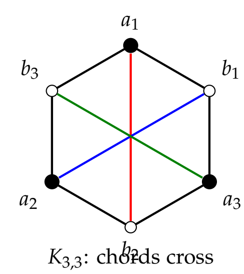


Figure 5.3: $K_{3,3}$ with Hamiltonian cycle (black). The 3 chord edges must cross: no planar embedding exists.

5.4 *Colouring Planar Graphs*

A famous problem in cartography asks: how many colours are required to colour a map such that no two adjacent regions share a colour? Graph theoretically, this is the chromatic number of the planar graph dual to the map. Euler's formula provides a crucial structural lemma.

Corollary 5.3. *Structural Lemma.* Every planar graph contains a vertex of degree at most 5.
推論

Proof

Let G be planar with n vertices and m edges. Assume for contradiction that $\delta(G) \geq 6$. Then every vertex has degree at least 6. By the Handshaking Lemma:

$$2m = \sum_{v \in V} d(v) \geq 6n.$$

Thus $m \geq 3n$. However, we know $m \leq 3n - 6$.

$$3n \leq 3n - 6 \implies 0 \leq -6,$$

a contradiction.

■

This allows us to bound the chromatic number by induction.

Theorem 5.4. ***The 6-Colour Theorem.***
Every planar graph is 6-colourable.

定理

Proof

We proceed by induction on n. The base cases $n \leq 6$ are trivial. Assume all planar graphs of order $n - 1$ are 6-colourable. Let G be a planar graph of order n. By the Structural Lemma, G contains a vertex v with $d(v) \leq 5$. Consider the graph $H = G - v$. H is a subgraph of a planar graph, hence planar. By the inductive hypothesis, H admits a proper colouring using 6 colours. We now reinsert v. The vertex v has at most 5 neighbours in G. These neighbours use at most 5 distinct colours. Since 6 colours are available, there is at least one colour not used by the neighbours of v. Assign this colour to v. This yields a valid 6-colouring of G.

■

Remark.

Refining this argument to prove 5-colourability is possible but requires considering chains of alternating colours to rearrange the colouring of the neighbours. The **4-Colour Theorem**, proven by Appel and Haken in 1976 using computer assistance, asserts that $\chi(G) \leq 4$ for all planar graphs.

5.5 *Exercises*

1. **Subgraph Monotonicity.**
   (a) Prove that if a graph G has a subgraph H that is not planar, then G is not planar.
   (b) Deduce that for every $n \geq 6$, the complete graph K_n is not planar.
2. **Almost Non-Planar.** The graphs K_5 and $K_{3,3}$ are the fundamental obstructions to planarity. Show that removing just one edge makes them planar by drawing a plane embedding of:
   (a) K_5 minus one edge ($K_5 - e$).
   (b) $K_{3,3}$ minus one edge ($K_{3,3} - e$).
3. **The Petersen Graph.** Prove that the Petersen graph is not planar.

> *Remark.*
>
> Hint: Use Kuratowski's Theorem by finding a subdivision of $K_{3,3}$ or K_5, or use the corollary $m \leq 3n - 6$ adapted for graphs with girth $g = 5$, which states $m \leq \frac{g}{g-2}(n-2)$.

4. **Disconnected Graphs.** Euler's formula $n - m + f = 2$ requires the graph to be connected.

   (a) Use induction to prove a formula for planar graphs that have exactly two connected components.

   (b) Generalise this to a graph with k connected components. Prove that $n - m + f = 1 + k$.

5. **Topological Surfaces.** For graphs embedded on a torus (a doughnut shape) such that all faces resemble discs, the Euler characteristic differs from the plane. Given that K_5 and $K_{3,3}$ can be embedded on a torus, and assuming the standard Euler logic applies, determine the value of the constant C in the formula $n - m + f = C$ for a torus.

   > *Remark.*
   >
   > Hint: K_5 has $n = 5, m = 10$. How many faces would it need? K_5 triangulation on torus implies $3f = 2m$.

6. **Feasibility Check.** For each of the following sets of conditions, either draw a connected, simple planar graph that satisfies them, or explain why one cannot exist:

   (a) $n = 15, m = 12$.

   (b) $n = 10, m = 33$.

   (c) $n = 5, m = 8$.

   (d) $n = 6, m = 9$, and the embedding has $f = 6$. (Note: Check Euler's formula consistency).

7. **Planarity of Complements.**

   (a) Show that if G is a simple planar graph with $n \geq 11$ vertices, then the complement graph $\overline{G}$ is not planar.

   > *Remark.*
   >
   > Hint: Consider the total number of edges in K_n ($m(G) + m(\overline{G})$) versus the maximum allowed in two planar graphs ($2(3n - 6)$).

   (b) Find a planar graph with $n = 8$ vertices whose complement is also planar.

8. **Self-Dual Graphs.** We define the **dual graph** G^* of a plane graph G by placing a vertex in every face of G and connecting two such vertices if their corresponding faces share an edge. A planar embedding is **self-dual** if G is isomorphic to G^*.

   (a) Prove that if a connected planar graph G is self-dual, then $2n - 2 = m$.

   (b) Find a self-dual planar embedding for the complete graph K_4.