# Graph Theory and Its Applications

Xinmin Hou

Department of Computational and Applied Mathematics
xmhou@ustc.edu.cn

# Contents

**1 Introduction to Graph Theory** **6**

**2 Basic Concepts of Graphs** **11**

Definitions . . . . . 11

Common Terminology . . . . . 12

Ramsey Number . . . . . 16

Exercises . . . . . 19

**3 Graph Isomorphism** **21**

Introduction to the Graph Isomorphism Problem . . . . . 24

Examples of the Probabilistic Method in Graph Theory . . 25

Exercises . . . . . 26

**4 Vertex Degrees and Graph Operations** **28**

The Handshaking Lemma . . . . . 30

Graph Operations . . . . . 31

Some Definitions and Notations . . . . . 34

Turán Theorem . . . . . 36

Exercises . . . . . 43

**5 Paths and Connectivity** **45**

Turán number for special family of graphs : stars, paths and cycles . . . . . 52

Connectivity and Strong Connectivity . . . . . 57

Exercises . . . . . 61

**6 Circuits and Cycles** **63**

**7 Distance and Diameter** **68**

Bipartite Graph Characterization Theorem . . . . . 72

**8 Exercises** **79**

**9 Eulerian Graphs** **80**

Characterization of Eulerian Graphs . . . . . 81

Exercises . . . . . 86

**10 Hamiltonian Graphs** **87**

Sufficient Conditions for Hamiltonian Graphs . . . . . 92

Exercises . . . . . 101

**11 P and NP** **102**

NP-complete . . . . . 103

Well Known NP-Complete problems . . . . . 104
A proof of 3-SAT to HCP . . . . . 106

**12 Distance and Diameter** **111**
Lower Bounds . . . . . 112
$(\Delta, d)$-Moore Graphs . . . . . 115
An Upper Bound . . . . . 117
Average Distance . . . . . 120
Line Graphs . . . . . 124
Exercises . . . . . 130

**13 Matrix Representation of Graphs** **131**
Exercises . . . . . 142

# 1 Introduction to Graph Theory

- Graph Theory is a branch of discrete mathematics. It studies graphs as its objects.
- A graph in graph theory is a figure consisting of a number of given points and lines connecting pairs of points, which is usually used to describe a certain specific relationship between certain things; points represent things, and lines connecting two points represent that the corresponding two things have this relationship.
- Graph theory is a discipline with a very wide range of applications. Graph theory provides a natural structure, and the mathematical models resulting from it are suitable for almost all fields of science (natural and social sciences), as long as the subject studied in this field is the relationship between "objects" and "objects".



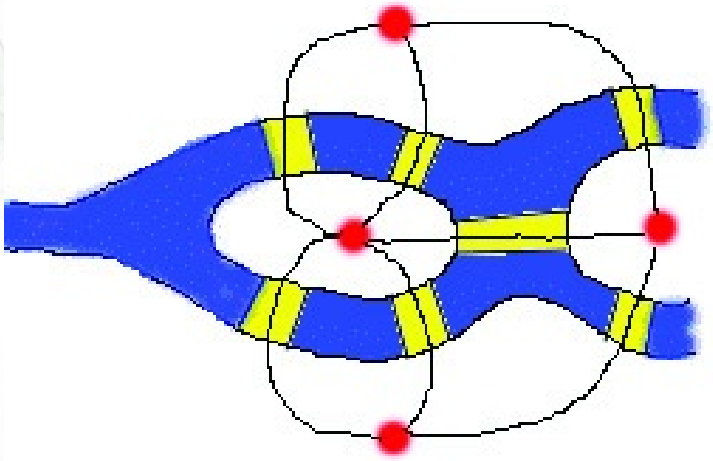


Figure 1: Leonhard Euler (1707-1783) and the Königsberg Bridge Problem Euler is recognized as one of the four greatest mathematicians in history; the other three are Archimedes (287 BC - 212 BC), Newton (1643-1727), and Gauss (1777-1855).

- Graph theory originated from the famous Königsberg Bridge Problem (1736).

- Königsberg (now Kaliningrad, Russia) was the capital of East Prussia. The Pregel River flows through the city, and there are two small islands connected to the river banks by 7 bridges.
- Legend has it that local residents wanted to design a walk, starting from somewhere, passing each bridge exactly once and returning to the starting point without repetition.This is today's "one-stroke drawing" problem, but no one could solve it at that time.
- Euler abstracted this problem into a mathematical model, drew a network-like graph with points and lines, proved that such a walk does not exist, and solved the Königsberg Bridge Problem.
- The discussion and research on such problems initiated and guided the development of two disciplines: graph theory and topology.
- According to the current status of graph theory research and development, it can be divided into the following branches: struc-

tural graph theory, algorithmic graph theory, extremal graph theory, random graph theory, topological graph theory, algebraic graph theory, hypergraph theory, and matroid theory.

References:


1 Bondy, J. A. and Murty, U. S. R., Graph Theory with Applications. The Macmillan Press Ltd, London and Basingstoke, 1976 (2nd ed., 2008, GTM244). Chinese translation by Wang Wangming et al., Science Press, 1984. This book treats undirected graphs and digraphs separately; the theory is accessible, the applications are excellent, and there are many exercises. Zhang Kemin et al., Solutions to Exercises (Tsinghua University Press, 1988), provides complete solutions.

2 Diestel, R., Graph Theory, Springer-Verlag New York, 2005. This is a graduate-level textbook, somewhat difficult, but very popular, GTM173.

3 Béla Bollobás, Modern Graph Theory, Springer-Verlag, 1998. Graduate textbook, somewhat difficult, GTM184.

# 2 Basic Concepts of Graphs

## Definitions

- An undirected simple graph is a binary structure $(V, E)$, where $V$ is a nonempty set, called the vertex set,

$$E \subseteq \binom{V}{2}$$

  where $\binom{V}{2}$ denotes the set of all 2-element subsets of $V$, $E$ is called the edge set, and elements of $V$ are called vertices (or points), and elements of $E$ are called edges.

- A directed simple graph (or digraph) is a binary structure $(V, A)$, where $V$ is a nonempty set, called the point set; $A \subseteq V \times V$ is called the arc set.

**Proposition 2.1** *(i) Let $G$ be a simple undirected graph of order $n$, then $|E(G)| \leq \frac{n(n-1)}{2}$;*
*(ii) If $D$ is a simple digraph of order $n$, then $|A(D)| \leq n(n-1)$.*

**Note:** Since the relative positions of the points representing elements of $V$ and the curvature, length of the line segments representing elements of $E$ are irrelevant, the graphical representation is not unique.

**Example 1.1.1** For example, the complete graph $K_4$ on 4 vertices.

## Common Terminology

- We use $a = (x, y) \in A(D)$ to denote a directed edge (or arc), called a directed edge from $x$ to $y$. $x$ is called the origin of $a$, and $y$ is called the terminus of $a$. The origin and terminus are collectively called the end-vertices of the edge $a$.
- We use $e = xy \in E(G)$ to denote an undirected edge, called an edge connecting $x$ and $y$.

- An edge and its two endpoints are called incident; Two endpoints incident with the same edge, or two edges incident with the same vertex, are called adjacent.
- An edge with identical endpoints is called a loop. Two edges with a common origin and a common terminus are called parallel edges or called multiple edges.
- Two directed edges with identical endpoints but opposite directions are called symmetric edges.
- For a digraph $D$, the undirected graph $G$ obtained by removing the directions of its edges is called the underlying graph.
- Given an undirected graph $G$, assigning a direction to each edge of $G$ to obtain a digraph $D$, such a digraph $D$ is called an oriented graph.
- Given an undirected graph $G$, replacing each edge of $G$ with a pair of symmetric edges to obtain a digraph $D$, is called the

symmetric digraph.

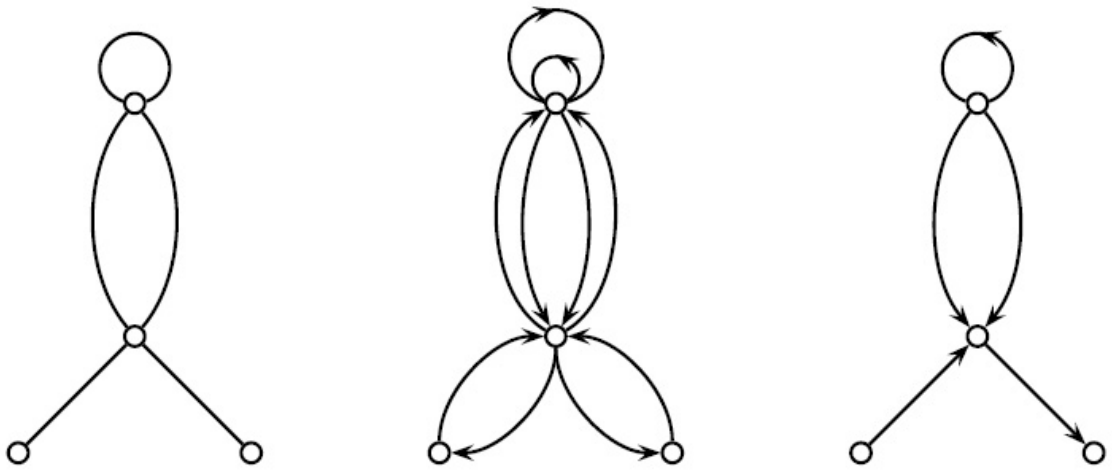


Figure 2: (a) Undirected graph $G$ (b) The symmetric digraph $D$ of $G$ (c) The orientation $H$ of $G$

- The number of vertices in a graph is called the **order**. A simple graph of order 1 is called a **trivial graph**. A graph with zero

edges is called an **empty graph**. A graph with both a finite number of vertices and edges is called a **finite graph**.

- $K_n$ denotes the complete graph on $n$ vertices, and $K_3$ is called a triangle.

Example 1.1.4 Prove: among any six people, there must exist three people who either all know each other, or none of them know each other.

# Ramsey Number

In general, the Ramsey number $R(r,s)$, is the least $N$ such that any blue-red edge-coloring of a complete graph $K_N$, has a monochromatic subgraph $K_r$ or $K_s$. We know little about the exact values of $R(r,s)$.

**Theorem 2.2 (Ramsey, 1930, Erdős, 1947)**

$$2^{r/2} < R(r,r) \leq 2^{2r-3}.$$

Table 1: Major asymptotic Ramsey results

| **Result** | **Paper and authors** | **Journal publication** |
|---|---|---|
| Diagonal upper bound $R(k,k) \le (4-\delta)^k$ | Campos, Griffiths, Morris, and Sahasrabudhe. *An exponential improvement for diagonal Ramsey* | **2026**, *Ann. of Math.* **203**(3), 869–932. Preprint: 2023. |
| $R(3,t) = \Theta(t^2/\log t)$ | Jeong Han Kim. *The Ramsey number $R(3,t)$ has order of magnitude $t^2/\log t$* | **1995**, *Random Struct. Algorithms* **7**(3), 173–207. |
| $R(3,t) \ge \left(\frac{1}{2} + o(1)\right) t^2/\log t$ | Hefty, Horn, King, and Pfender. *Improving $R(3,k)$ in just two bites* | **2025** preprint; revised 2026. Journal publication not verified. |
| $R(4,t) = \Omega(t^3/\log^4 t)$ | Mattheus and Verstraete. *The asymptotics of $r(4,t)$* | **2024**, *Ann. of Math.* **199**(2), 919–941. |
| Fixed-ratio lower bound $R(k, \lceil Ck \rceil) \ge \left(p_C^{-1/2} + \varepsilon\right)^k$ | Jie Ma, Wujie Shen, and Shengjie Xie. *An exponential improvement for Ramsey lower bounds* | **2026**, *Invent. Math.* **245**, 1059–1115. Preprint: 2025. |

For fixed $C > 1$, $\varepsilon = \varepsilon(C) > 0$ and $C = \log p_C / \log(1 - p_C)$ with $p_C \in (0, 1/2)$; the final bound holds for sufficiently large $k$. Publication details checked September 8, 2026.



Figure 3: Erdős and Terence Tao (age 10) discussing a problem

To date, only 9 exact values of $R(r,s)$ $(r,s \geq 3)$ are known. Paul Erdős once used a metaphor to describe the difficulty of finding Ramsey numbers: "Imagine an alien army landing on Earth, demanding the value of $R(5,5)$, or they will destroy the Earth. In this situation, we should concentrate all our computers and mathematicians to find this value. If they asked for the value of R(6,6), we should try to destroy the alien army."

**Small Ramsey number $R(k,l)$: best known exact values/bounds.** Source: Radziszowski, Small Ramsey Numbers, Electron. J. Combin. DS1.18 (rev. #18, 2026-04-24), Tables Ia/Ib.

| $k\backslash l$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **3** | **6** | **9** | **14** | **18** | **23** | **28** | **36** | 40–41 | 47–50 | 53–59 | 61–68 | 67–77 | 74–87 |
| **4** | 9 | **18** | **25** | 36–40 | 49–58 | 59–79 | 73–105 | 92–135 | 102–170 | 128–210 | 139–256 | 148–307 | 159–364 |
| **5** | 14 | 25 | **43–46** | 59–85 | 80–133 | 101–193 | 133–282 | 149–381 | 183–511 | 203–672 | 233–860 | 267–1081 | 275–1341 |
| **6** | 18 | 36–40 | 59–85 | **102–160** | 115–270 | 134–423 | 183–651 | 204–944 | 262–1346 | 294–1855 | 347–2499 | 401–3301 | $\leq 4305$ |
| **7** | 23 | 49–58 | 80–133 | 115–270 | **205–492** | 219–832 | 252–1368 | 292–2119 | 405–3197 | 417–4665 | 511–6653 | $\leq 9260$ | $\leq 12635$ |
| **8** | 28 | 59–79 | 101–193 | 134–423 | 219–832 | **282–1518** | 329–2662 | 343–4402 | 457–7040 | 817–10836 | 873–27485 | $\leq 41525$ | $\leq 63609$ |
| **9** | 36 | 73–105 | 133–282 | 183–651 | 252–1368 | 329–2662 | **565–4956** | 581–8675 | $\leq 14631$ | $\leq 38832$ | $\leq 64864$ | | |
| **10** | 40–41 | 92–135 | 149–381 | 204–944 | 292–2119 | 343–4402 | 581–8675 | **798–16064** | 1313–45881 | $\leq 81123$ | | | |

**exact** | **diagonal $R(k,k)$ (open)** | interval lower–upper | only upper bound $\leq$ | symmetric $= R(l,k)$

# Exercises

1 Prove that $R(3,4) = 9$.

2 Nine mathematicians meet at an international conference. Among any three of them, at least two speak the same language. Prove: if each mathematician speaks at most three lan-

guages, then at least three mathematicians can converse in the same language (7th US Mathematical Olympiad problem, May 2, 1978).

**Problem 2.3** *Determine the exact value of small Ramsey numbers:*

- $R(5,5) = 43 \sim 46$*: lower bound 43(Exoo 1989); upper bound* $48 \to 46$*(Angeltveit-McKay 2023, JGT 2026).*
- $R(3,t)$*:* $R(3,10) = 40 \sim 41,\ R(3,11) = 47 \sim 50.$
- $R(4,t)$*:* $R(4,6) = 36 \sim 40,\ R(4,7) = 49 \sim 58.$

# 3 Graph Isomorphism

- $D = (V(D), E(D))$ and $H = (V(H), E(H))$ are called isomorphic, denoted by $D \cong H$, if there exists a bijection

$$\theta : V(D) \to V(H),$$

such that for any

$$(x, y) \in E(D) \Leftrightarrow (\theta(x), \theta(y)) \in E(H).$$

The mapping $\theta$ is called an isomorphic mapping between $D$ and $H$.

- To prove that two graphs are isomorphic, one must point out an isomorphic mapping between them.
- It is easy to see that the isomorphism relation on graphs is an equivalence relation. This equivalence relation partitions graphs

of the same order into several equivalence classes; two isomorphic graphs belong to the same class. Graphs in the same class have the same structure, differing only in the names of vertices and edges.

complete graph : a simple undirected graph in which every pair of distinct vertices is connected by an edge. The complete graph of order $n$ is denoted by $K_n$.

Complete digraph: the symmetric digraph of a complete graph. The complete digraph of order $n$ is denoted by $K_n^*$.

tournament: an orientation of a complete graph.

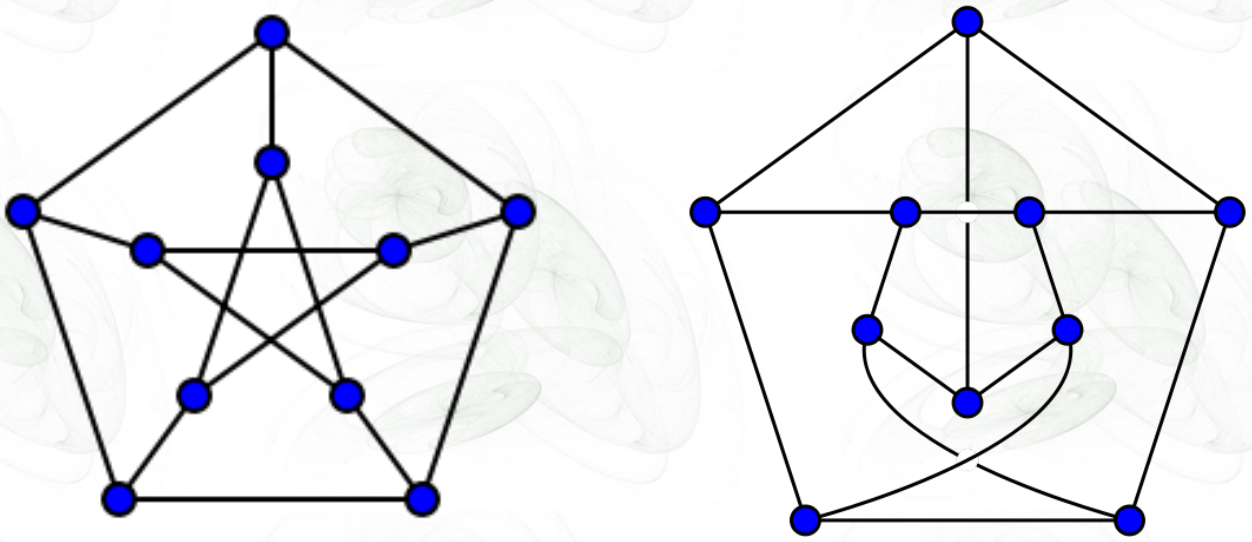


Figure 4: Two drawings of the Petersen graph

(The reason for this name is that this graph vividly describes the results of a round-robin competition with $n$ players. If player $x$ beats player $y$, then edge $xy$ is directed from $x$ to $y$.)

Bipartite graph: if the vertex set of a loopless graph can be partitioned into two nonempty subsets $X$ and $Y$ such that no two

vertices in $X$ are connected by an edge and no two vertices in $Y$ are connected by an edge, then the graph is called a bipartite graph, and $\{X, Y\}$ is called a **bipartition**. A bipartite graph with bipartition $\{X, Y\}$ is denoted by $(X \cup Y, E)$.

Similarly, one can define general $k$-partite graphs ($k \geq 2$), complete $k$-partite graphs, etc.

## Introduction to the Graph Isomorphism Problem

- The set of all automorphisms of a graph $G$, with the operation of composition of permutations, is a permutation group on $V(G)$ (a subgroup of the symmetric group on $V(G)$). This is the automorphism group of $G$, denoted $Aut(G)$.
- Two algorithmic questions arising from the above definitions are graph isomorphism and finding the automorphism group.

- The graph isomorphism problem belongs to NP, but it is not known whether it is in P or is NP-complete. Babai showed that it can be solved in quasipolynomial time ($\exp((\log n)^{O(1)})$, STOC '16).
- $Aut(K_n) = Aut(K_n^c) = S_n$ (symmetric group); $Aut(C_n) = D_n$ (Dihedral group), etc.

## Examples of the Probabilistic Method in Graph Theory

Example 1.1.4 Prove: if $\binom{n}{k}(1-2^{-k})^{n-k} < 1$, then there exists a tournament on $n$ vertices such that for any chosen $k$ players, there exists a player who beats all these $k$ players.

**Proof:**

- Consider a random tournament on $V = \{1, 2, \cdots, n\}$. For any fixed $k$-element subset $K$ of $V$, let $A_K$ denote the event that no player beats all players in $K$.
- Because for any $v \in V - K$, the probability that $v$ cannot beat all players in $K$ is $1 - 2^{-k}$, and these $n - k$ events are mutually independent, $Pr(A_K) = (1 - 2^{-k})^{n-k}$.
- Hence

$$Pr(\cup_{\substack{K \subset V \\ |K|=k}} A_K) \leq \sum_{\substack{K \subset V \\ |K|=k}} Pr(A_K) = \binom{n}{k}(1 - 2^{-k})^{n-k} < 1.$$

  This shows that the probability that $A_K$ does not occur is positive, that is, the conclusion holds.

## Exercises

**1** Prove that $Aut(C_n) = D_n$, where $D_n$ denotes the dihedral group.

# 4 Vertex Degrees and Graph Operations

- Let $G$ be an undirected graph. The vertex degree of $x \in V(G)$ is defined as the number of edges incident with $x$ in $G$ (a loop is counted twice), denoted by $d_G(x)$.
- A vertex of degree zero is called an **isolated vertex**. We use $\Delta(G)$ and $\delta(G)$ to denote the **maximum** and **minimum vertex degree** of $G$, respectively. That is,

$$\Delta(G) = \max\{d_G(x) : x \in V(G)\},$$
$$\delta(G) = \min\{d_G(x) : x \in V(G)\}.$$

- An undirected graph $G$ is called $k$**-regular** if for every $x \in V(G)$, $d_G(x) = k$.
- Let $D$ be a digraph. The vertex out-degree of a vertex $y$ in $D$ is defined as the number of directed edges in $D$ starting at $y$,

denoted by $d_D^+(y)$. The vertex in-degree of $y \in V(D)$ is defined as the number of directed edges in $D$ ending at $y$, denoted by $d_D^-(y)$. The vertex degree of $y \in V(D)$ is defined as $d_D^+(y) + d_D^-(y)$, denoted by $d_D(y)$.

- If $d_D^+(y) = d_D^-(y)$, then $y$ is called a balanced vertex. A digraph in which every vertex is balanced is called a balanced digraph.
- The symmetric digraph of an undirected graph is balanced.
-

$$\begin{aligned}\Delta^+(D) &= \max\{d_D^+(y) : y \in V(D)\};\\ \Delta^-(D) &= \max\{d_D^-(y) : y \in V(D)\};\\ \delta^+(D) &= \min\{d_D^+(y) : y \in V(D)\};\\ \delta^-(D) &= \min\{d_D^-(y) : y \in V(D)\}.\end{aligned}$$

denote the **maximum** and **minimum vertex out-degree**, **in-degree** of $D$, respectively. A digraph in which all four parameters equal $k$ is called a $k$**-regular digraph**.

# The Handshaking Lemma

**Theorem 4.1 (L. Euler, 1736)** *For any undirected graph $G$,*

$$\sum_{x \in V(G)} d_G(x) = 2|E(G)|.$$

**Proof:** (**Double-counting**) Count the number of elements in the set

$$A = \{(x, e) : x \in V, e \in E \text{ and } x \in e\}$$

.

**Corollary 1** *In any undirected graph, the number of vertices of odd degree must be even.*

**Proof** Let $V_o$ and $V_e$ be the sets of vertices of odd degree and even degree in $G$, respectively. By Theorem 4.1,

$$\sum_{x \in V_o} d_G(x) + \sum_{x \in V_e} d_G(x) = 2|E(G)|.$$

Both the right-hand side and the second term on the left-hand side of the above equation are even, so the first term on the left-hand side is also even. Since $d_G(x)$ is odd when $x \in V_o$, $|V_o|$ is even. ∎

**Corollary 2** *For any digraph $D$,*

$$\sum_{x \in V(D)} d_D^+(x) = \sum_{x \in V(D)} d_D^-(x) = |A(D)|.$$

**Proof** Omitted

## Graph Operations

- Let $G = (V(G), E(G))$ and $H = (V(H), E(H))$ be two graphs. If $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$, then $H$ is called a subgraph, denoted by $H \subseteq G$. $G$ is called a supergraph.

- If $H \subseteq G$ and $V(H) = V(G)$, then $H$ is called a spanning subgraph of $G$. If $H \subseteq G$ and $H \neq G$, then $H$ is called a proper subgraph, denoted by $H \subset G$.

- Let $V'$ be a nonempty proper subset of $V(G)$, the subgraph with vertex set $V'$ and edge set consisting of edges of $G$ both of whose endpoints are in $V'$ is called the subgraph of $G$ induced by $V'$, abbreviated as induced subgraph, denoted by $G[V']$. The induced subgraph $G[V \setminus V']$ is denoted by $G - V'$. If $V' = \{x\}$, then $G - \{x\}$ is abbreviated as $G - x$.

- Let $E'$ be a nonempty subset of $E(G)$. The subgraph with edge set $E'$ and vertex set consisting of the endpoints of edges in $E'$ is called the subgraph of $G$ induced by $E'$, denoted by $G[E']$. $G - E'$ denotes the subgraph obtained from $G$ by deleting $E'$ (but not deleting the endpoints).

- Let $G_1 \subseteq G, G_2 \subseteq G$. If $V(G_1) \cap V(G_2) = \emptyset$, then $G_1$ and $G_2$ are **vertex-disjoint**; if $E(G_1) \cap E(G_2) = \emptyset$, then $G_1$ and $G_2$ are **edge-disjoint**.

- The **union** $G_1 \cup G_2$ of $G_1$ and $G_2$ is the subgraph of $G$ with vertex set $V(H) = V(G_1) \cup V(G_2)$ and edge set $E(H) = E(G_1) \cup E(G_2)$ as its edge set.

- If $G_1$ and $G_2$ are vertex-disjoint, then denote $G_1 \cup G_2 = G_1 + G_2$. If for every $i$ $(1 \leq i \leq n)$, $G_i \cong H$, then denote $G_1 + G_2 + \cdots + G_n = nH$. If $G_1$ and $G_2$ are edge-disjoint, then denote $G_1 \cup G_2 = G_1 \oplus G_2$.

- If $V(G_1) \cap V(G_2) \neq \emptyset$, then one can similarly define the **intersection** $G_1 \cap G_2$.

# Some Definitions and Notations

- Let graph $G = (V, E)$, $S \subseteq V$, and use $N_G(S)$ to denote the set of vertices in $G$ adjacent to vertices in $S$, called the neighborhood of $S$.
- If $G = (V, E)$ is a simple graph, then $d_G(v) = |N_G(v)|$ for $v \in V$.
- Let $S, T \subseteq V(G)$, and use $E_G(S)$ and $E_G(S, T)$ to denote the set of edges both of whose endpoints are in $S$ and the set of edges with one endpoint in $S$ and the other in $T$, respectively. Denote $e_G(S) = |E_G(S)|$ and $e_G(S, T) = |E_G(S, T)|$.
- Similarly, we can define $N_D^+(S)$, called the **out-neighborhood** of $S$, and $N_D^-(S)$, called the **in-neighborhood** of $S$.

**Example 1**: $2n$ people attend a party. Prove that at least two people have an even number of common friends.

**Proof:**

- Define a graph $G = (V, E)$ on $2n$ vertices, where $xy \in E$ if and only if $x$ and $y$ are friends.
- The statement holds if and only if there exist $x, y \in V$ such that $|N(x) \cap N(y)|$ is even.
- (By contradiction) Assume that for any $x, y \in V$, $|N(x) \cap N(y)|$ is odd.
- Claim: every vertex of $G$ has even degree.

  Indeed, by the Handshaking Lemma applied to $G[N(x)]$: each $y \in N(x)$ has, inside this subgraph, degree $|N(y) \cap N(x)|$, odd by assumption; hence $|N(x)|$ must be even.
- Take a vertex $x \in V$; then $d(x) \equiv 0 \pmod 2$.

- For any $y \in V - (\{x\} \cup N(x))$, we have $|N(y) \cap N(x)| \equiv 1 \pmod 2$.
- Since $d(y) \equiv 0 \pmod 2$, we have $|N(y) \cap (V - (\{x\} \cup N(x))| \equiv 1 \pmod 2$.
- Let $H = G[V - (\{x\} \cup N(x))]$; all vertices of $H$ are odd-degree vertices, hence $|V - (\{x\} \cup N(x))| \equiv 0 \pmod 2$.
- Thus $|V| = |N(x)| + |V - (\{x\} \cup N(x))| + 1 \equiv 1 \pmod 2$, a contradiction.

## Turán Theorem

**Theorem 4.2 (Mantel, 1907)** *Let $G$ be a simple undirected graph of order $n$ with no triangle. Then $e(G) \leqslant \lfloor \frac{n^2}{4} \rfloor$, with equality if and only if $G = K_{\lceil n/2 \rceil, \lfloor n/2 \rfloor}$.*

**Proof**

- If $G$ is an empty graph, then $e(G) = 0$, and the conclusion holds.
- Now assume $G$ is nonempty. Take any $xy \in E(G)$. Since $G$ is a simple graph with no triangle,

$$(d_G(x) - 1) + (d_G(y) - 1) \leqslant n - 2,$$

that is

$$n \geqslant d_G(x) + d_G(y).$$

- Summing both sides over all edges of $G$, and by the Cauchy

inequality and Theorem 4.1, we obtain

$$\begin{aligned} n \cdot e(G) &\geq \sum_{e=xy \in E(G)} (d_G(x) + d_G(y)) \\ &= \sum_{x \in V} d_G^2(x) \text{ (double counting)} \\ &\geq \frac{1}{n} \left( \sum_{x \in V} d_G(x) \right)^2 \text{ (Cauchy inequality or Jensen inequa} \\ &= \frac{4}{n} e(G)^2. \end{aligned}$$

Consequently, $e(G) \leqslant \frac{n^2}{4}$.

•

$$\begin{aligned}\text{Equality holds} &\Leftrightarrow d_G(x)+d_G(y)=n, \quad xy\in E(G)\\ &\Leftrightarrow d_G(x)=d_G(y)=\frac{n}{2}, \quad xy\in E(G)\\ &\Leftrightarrow G=K_{\frac{n}{2},\frac{n}{2}}\end{aligned}$$

∎

Note: Cauchy inequality: $(\sum_{i=1}^{n} a_i b_i)^2 \le (\sum_{i=1}^{n} a_i^2)(\sum_{i=1}^{n} b_i^2)$.

- In general, for a given graph $H$, define

$$\mathrm{ex}(n,H)=\max\{e(G) \mid G \text{ is an } n\text{-vertex } H\text{-free graph}\},$$

called the Turán number. A graph with $\mathrm{ex}(n,H)$ edges containing no $H$ as a subgraph is called an extremal graph for $H$. Here an $H$-free graph is a graph containing no $H$ as a subgraph

- Mantel's theorem tells us $\mathrm{ex}(n, K_3) = \lfloor \frac{n^2}{4} \rfloor$.
- Let $T_r(n)$ denote the balanced complete $r$-partite graph, where each part has $\lceil \frac{n}{r} \rceil$ or $\lfloor \frac{n}{r} \rfloor$ vertices.
- Clearly, $\delta(T_r(n)) = n - \lceil \frac{n}{r} \rceil$ and $\Delta(T_r(n)) = n - \lfloor \frac{n}{r} \rfloor$; moreover, $T_r(n)$ is a maximal graph with no $K_{r+1}$ (adding any edge to $T_r(n)$ induces a $K_{r+1}$ subgraph), hence

$$\mathrm{ex}(n, K_{r+1}) \geq e(T_r(n)).$$

**Theorem 4.3 (Turán, 1941)** *For $r \geq 2$*

$$\mathrm{ex}(n, K_{r+1}) = e(T_r(n))$$

*and $T_r(n)$ is the unique extremal graph for $K_{r+1}$.*

**Proof**:

- By induction on $n$.
- For $n \leq r+1$, the result is trivial. Suppose $n > r+1$ and the result holds for smaller $n$.
- Let $G$ be a graph on $n$ vertices, $e(T_r(n))$ edges and it has no $K_{r+1}$.
- Then $\delta(G) \leq \delta(T_r(n)) \leq \Delta(T_r(n))$ (WHY?).
- Choose $x \in V(G)$ with $d(x) = \delta(G)$. Then

$$e(G-x) = e(G) - d(x) \geq e(T_r(n-1)).$$

- Note that $G-x$ has no $K_{r+1}$ as a subgraph. By induction hypothesis $G-x \cong T_r(n-1)$.
- Hence $d(x) = \delta(T_r(n)) = n - \lceil \frac{n}{r} \rceil$. $x$ is joined to all but $n-1-(n-\lceil \frac{n}{r} \rceil) = \lfloor \frac{n-1}{r} \rfloor$ vertices of $G-x$.

- As $x$ cannot be joined to a vertex in each class of $G-x$ (WHY?), it has to be joined to all but a small class of $G-x$. This shows that $G \cong T_r(n)$.

Example 1.3.4 A loopless graph $G$ has a bipartite spanning subgraph $H$ such that for every vertex $x$, $d_H(x) \geq d_G(x)/2$. and hence $e(H) \geq e(G)/2$.

**Proof**

- Choose $H = (V = X \dot{\cup} Y, E) \subseteq G$ such that $e(H)$ is maximum.
- Then for each $x \in X$ or $Y$, $d_{G[X]}(x) \leq d_H(x)$.
- Since $d_G(x) = d_H(x) + d_{G[X]}(x)$, $d_H(x) \geq \frac{1}{2} d_G(x)$.
- $e(H) \geqslant e(G)/2$.

Open problems related to Example 1.3.4:

**Problem 4.4** *For any given cubic (3-regular) graph $G = (V, E)$, if there is a partition $X \cup Y$ of $V$ such that $\Delta(X) \leq 1$ and $\Delta(Y) \leq 1$ and $||X| - |Y|| \leq 2$.*

**Problem 4.5** *For any $d$-regular graph $G = (V, E)$, if there is a partition $X \cup Y$ of $V$ such that $\delta(X) \geq d/2$ and $\delta(Y) \geq d/2$.*

## Exercises

**1.3.12** Prove: Let $G$ be a simple undirected graph with vertex set $V = \{x_1, x_2, \cdots, x_n\}$ containing no subgraph $K_{r+1}$, then there exists a complete $r$-partite graph $H$ with vertex set $V$ such that for every $i$ $(1 \leq i \leq n)$, $d_G(x_i) \leq d_H(x_i)$, with equality if and only if $G \cong H$. (P. Erdős, 1970)

# 5 Paths and Connectivity

- A walk (or chain) $W$ connecting vertices $x$ and $y$ in a graph is called an $xy$-walk, which is an alternating sequence of vertices $x_i$ and edges $a_j$:

$$W = (x =)x_{i_0}a_{i_1}x_{i_1}a_{i_2}\cdots a_{i_k}x_{i_k}(= y),$$

  where the two vertices $x_{i_{j-1}}$ and $x_{i_j}$ adjacent to edge $a_{i_j}$ are exactly the two endpoints of $a_{i_j}$ (possibly $x_{i_{j-1}} = x_{i_j}$). $x$ and $y$ are called the end-vertices of $W$, and the other vertices are called internal vertices. The number $k$ of edges in $W$ is called the length of $W$.

- A walk with distinct edges is called a trail.

- A trail with distinct vertices is called a path.

- A walk (trail) with identical end-vertices is called a closed walk (closed trail).

- A closed trail is called a circuit. A closed trail with distinct internal vertices is called a cycle.
- For a simple graph, a walk $W$ of length $k$ is uniquely determined by its vertex sequence $x_{i_0}x_{i_1}\cdots x_{i_k}$, so we abbreviate it as $W = (x_{i_0}, x_{i_1}, \cdots, x_{i_k})$ or $W = x_{i_0}x_{i_1}\cdots x_{i_k}$.
- Orient an $xy$-walk (trail, path) $W$ in a digraph $D$ from $x$ to $y$. If all edges in $W$ are oriented consistently, then $W$ is called a **directed walk** (**trail**, **path**) from $x$ to $y$ in $D$, denoted as an $(x, y)$-walk (trail, path).
- A path of maximum length in a graph is called a longest path.
- A path (cycle) containing every vertex of a graph is called a Hamilton path (cycle).
- A graph containing a Hamilton cycle is called Hamiltonian.

**Theorem 5.1** *Let $G$ be a simple undirected graph, $\delta = \delta(G)$, then $G$ contains a path of length at least $\delta$ and a cycle of length at least $\delta + 1$ (if $\delta \geq 2$) (exercise).*

**Proof**

- Let $P = (x_0, x_1, \cdots, x_k)$ be a longest path in $G$,
- By the maximality of $P$, $N_G(x_0) \subseteq \{x_1, x_2, \cdots, x_k\}$.
- Since $G$ is a simple graph, $k \geq |N_G(x_0)| = d_G(x_0) \geqslant \delta$.

∎

**Theorem 5.2** *Let $G$ be a connected graph with order $n \geq 3$ such that for any two non-adjacent vertices $x, y$ we have*

$$d(x) + d(y) \geq k.$$

*If $k = n$ then $G$ is Hamiltonian, and if $k < n$ then $G$ contains a path of length at least $k$ and a cycle of length at least $\frac{k+2}{2}$.*

**Proof**.

- Suppose $G$ is not Hamiltonian.
- Let

$$P = x_0 x_1 \cdots x_\ell$$

be a longest path in $G$. Then $x_0 x_\ell \notin E(G)$ (WHY?).

- Hence $N(x_0), N(x_\ell) \subseteq \{x_1, \cdots, x_{\ell-1}\}$.
- Let $N(x_0) = \{x_i \mid x_0 x_i \in E(G)\}$ and $N^+(x_\ell) = \{x_{i+1} \mid x_\ell x_i \in E(G)\}$. Then $N(x_0) \cap N^+(x_\ell) = \emptyset$ (WHY?).
- Note that $N^+(x_\ell) \subset \{x_2, x_3, \ldots, x_\ell\}$. Hence

$$k \leq d(x_0) + d(x_\ell) = |N(x_0) \cup N^+(x_\ell)| \leq \ell \leq n - 1.$$

- If $k = n$ then we have a CONTRADICTION. So $G$ is Hamiltonian.
- If $k < n$ then $\ell \geq k$. We are DONE.
- Without loss of generality, assume $d(x_0) \geq d(x_\ell)$. Then $d(x_0) \geq \frac{k}{2}$.
- Let $N(x_0) = \{x_{i_1}, \cdots, x_{i_s}\}$. Then $i_s \geq \frac{k}{2}$.
- Hence $C = x_0 x_1 \cdots x_{i_s} x_0$ is a cycle of length at least $\frac{k}{2} + 1$. ∎

**Remark:** The following theorems are corollaries of Theorem 5.2.

**Theorem 5.3 (Dirac's Theorem, 1952)** *Let $G$ be a connected graph on $n$ vertices with $\delta(G) \geq k$. If $k \geq n/2$ (Dirac's condition) then $G$ is Hamiltonian, and if $n > 2k$ then $G$ contains a path of length at least $2k$.*

**Theorem 5.4 (O. Ore, 1960, Ore's condition)** *Let $G$ be a simple graph of order $n$ ($\geq 3$). If for any two non-adjacent vertices $x$ and $y$ in $G$, $d_G(x)+d_G(y) \geq n$ (Ore's condition), then $G$ is Hamiltonian.*

Theorem 1.4.1 (L. Rédei, 1934) Every tournament contains a Hamilton directed path.

**Proof**

- Let $D$ be a tournament of order $n$, and $P = (x_1, x_2, \cdots, x_k)$ be a longest directed path in $D$.
- If $k = n$, the conclusion holds.
- Assume $k < n$. Then there exists $x \in V(D) \setminus V(P)$ with $(x, x_k) \in A(D)$ and $(x_1, x) \in A(D)$.

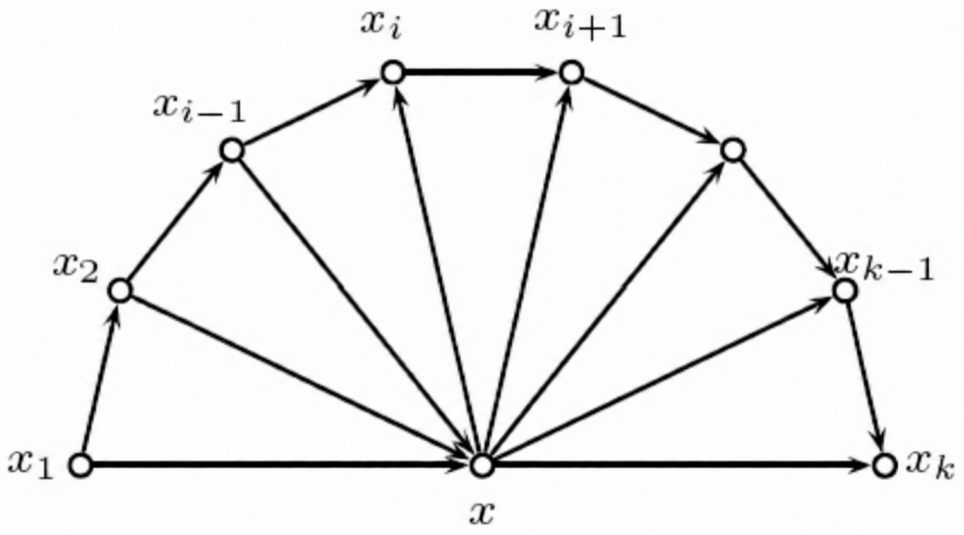



Figure 5: Proof of Theorem 1.4.1

- Hence there exists $i$ $(1 < i \leq k)$ such that $(x_{i-1}, x), (x, x_i) \in A(D)$ (see Figure 5).
- Then $(x_1, x_2, \cdots, x_{i-1}, x, x_i, x_{i+1}, \cdots, x_k)$ is a directed path in $D$ longer than $P$ by 1, contradicting the maximality of $P$.

- Therefore, $P$ is a Hamilton directed path.

∎

## Turán number for special family of graphs : stars, paths and cycles

**Theorem 5.5**

$$\mathrm{ex}(n, K_{1,k+1}) \leq \frac{kn}{2},$$

*with equality holds if and only if $G$ is a $k$-regular graph.*

**Theorem 5.6 (Erdős-Gallai Theorem, 1959)**

$$\mathrm{ex}(n, P_{k+1}) \leq \frac{(k-1)n}{2},$$

*with equality holds if and only if $G$ is the union of disjoint complete graphs $K_k$, where $P_{k+1}$ is a path of length $k$.*

**Proof**.

- By induction on $n$.
- If $n \leq k$ then the result is clearly true.
- Suppose $n > k$ and the result is true for smaller values of $n$.
- If $G$ is disconnected then the induction hypothesis implies the result.
- If $G$ is connected, then $G$ contains no $K_k$ (WHY?) and, by Theorem 5.2, $G$ has a vertex $x$ with $d(x) \leq \frac{k-1}{2}$.
- As $G-x$ is not an extremal graph (WHY?), by the induction hypothesis, $e(G-x) < \frac{(k-1)(n-1)}{2}$.

- Therefore,

$$\begin{aligned} e(G) &= d(x) + e(G-x) \\ &< \frac{k-1}{2} + \frac{(k-1)(n-1)}{2} \\ &= \frac{(k-1)n}{2}. \end{aligned}$$

In fact, we have

**Theorem 5.7**

$$\mathrm{ex}(n, P_{k+1}) = \frac{(k-1)n}{2} - \frac{r(k-r)}{2},$$

*where* $n \equiv r \pmod{k}$.

**Theorem 5.8 (Erdős-Sós Conjecture)**

$$\mathrm{ex}(n, T_{k+1}) \leq \frac{(k-1)n}{2},$$

*where $T_{k+1}$ is a tree (simple connected graph containing no cycle) on $k+1$ vertices.*

**Remark:** Theorem 5.5 and Theorem 5.6 imply that Erdős-Sós Conjecture is true for $K_{1,k}$ and $P_{k+1}$. The Erdős–Sós conjecture has inspired an extensive literature, including a proof announced by Ajtai, Komlós, Simonovits, and Szemerédi (at early 1990s) for sufficiently large trees that remained unpublished. Remarkably, GPT-6 recently produced a proof that fits on a single page.

**Theorem 5.9 (Erdős-Gallai Theorem)**

$$\mathrm{ex}(n, C_{\geq k+1}) \leq \frac{k(n-1)}{2},$$

*with equality holds if and only if $G$ is connected and all its blocks are complete graphs of order $k$.*

# Connectivity and Strong Connectivity

- Let $x, y \in V$. If there exists a path (not necessarily directed) connecting $x$ and $y$ in $D$ (or $G$), then $x$ and $y$ are said to be connected.
- The connectivity relation on elements of $V$ is an equivalence relation on $V$. This equivalence relation partitions elements of $V$ into equivalence classes $V_1, V_2, \cdots, V_\omega$. $x$ and $y$ belong to the same class $V_i \Leftrightarrow x$ and $y$ are connected.
- The subgraph $D[V_i]$ (or $G[V_i]$) $(1 \leq i \leq \omega)$ is called a connected component (connected component) of $D$ (or $G$); $\omega$ is called the number of components (number of components), denoted by $\omega = \omega(D)$ (or $\omega(G)$).
- If $\omega(D) = 1$ (or $\omega(G) = 1$), then $D$ (or $G$) is called a connected digraph (or graph). Otherwise, it is called a disconnected digraph (or graph).

- Let $D$ be a digraph, $x, y \in V(D)$. If $D$ contains both an $(x,y)$-path and a $(y,x)$-path, then $x$ and $y$ are strongly connected.
- The strong connectivity relation on elements of $V$ is an equivalence relation on $V$. The equivalence classes $V_i$ of $V$ induce subgraphs $D[V_i]$ in $D$, called the **strongly connected components** of $D$. The **number of strongly connected components** of $D$ is denoted by $\overrightarrow{\omega}(D)$.
- If $\overrightarrow{\omega}(D) = 1$, then $D$ is called a strongly connected graph. Otherwise, $D$ is called a non-strongly-connected graph.
- Let $x \in V(G), e \in E(G)$. If $\omega(G-x) > \omega(G)$, then $x$ is called a cut vertex; otherwise $x$ is called a non-cut vertex (connected vertex).
- If $\omega(G-e) > \omega(G)$, then $e$ is called a cut edge, also called a bridge (bridge); otherwise $e$ is called a non-bridge edge (connected edge).

- A connected graph with no cut vertex is called a block. A maximal connected subgraph of $G$ with no cut vertex is called a block of $G$.

**Property 1.4.2** (a) $D$ **(or** $G$**) is connected** $\Leftrightarrow$ **for every nonempty proper subset** $S$ **of** $V$**,** $E(S, \overline{S}) \neq \emptyset$.

(b) $D$ **is strongly connected** $\Leftrightarrow$ **for every nonempty proper subset** $S$ **of** $V$**,** $E^+(S, \overline{S}) \neq \emptyset$ **and** $E^-(\overline{S}, S) \neq \emptyset$. ∎

**Proposition 5.10** *Let* $G$ *be a simple undirected graph,* $V(G) = \{x_1, x_2, \cdots, x_n\}$ *with* $d_G(x_1) \leq d_G(x_2) \leq \cdots \leq d_G(x_n)$. *If for every natural number* $k (1 \leqslant k \leqslant n - d_G(x_n) - 1)$, $d_G(x_k) \geq k$, *then* $G$ *is connected.*

**Proof** (By contradiction)

- If $G$ is disconnected, by Theorem 1.4.2(a), there exists a nonempty proper subset $S \subset V(G)$ such that $E(S, \overline{S}) = \emptyset$.

- Without loss of generality, let $x_n \in \overline{S}$, and set $k = |S|$. Since $G$ is a simple graph, $|\overline{S}| \geq d_G(x_n) + 1$, and $k = |S| = n - |\overline{S}| \leqslant n - d_G(x_n) - 1$. By assumption, $d_G(x_k) \geq k$.
- Since $d_G(x) \leq |S| - 1 = k - 1$ for every $x \in S$, we have $x_k \in \overline{S}$. Hence $S$ contains at most $(k-1)$ vertices, contradicting $|S| = k$ (space obstruction). Therefore, $G$ is connected. ∎

**Proposition 5.11** *Let $D$ be a simple digraph of order $n$ with $|E(D)| > (n-1)^2$. Then $D$ is strongly connected.*

**Proof** (By contradiction) If $D$ is not strongly connected, by Theorem 1.4.2(b), there exists a nonempty proper subset $S \subset V(D)$ such that $E^+(S, \overline{S}) = \emptyset$. Let $k = |S|$, then

$$\begin{aligned} |E(D)| \leqslant & \quad k(k-1) + (n-k)(n-k-1) + k(n-k) \\ = & \quad (n-1)^2 - (k-1)(n-k-1) \leqslant (n-1)^2, \end{aligned}$$

This contradicts the hypothesis. Therefore, $D$ is strongly connected. ∎

# Exercises

1.4.3 Prove: let $D$ be a simple graph, and $k = \max\{\delta^+, \delta^-\}$, then $D$ contains a directed path of length at least $k$.

1.4.7 Let $G$ be a simple graph of order $n$, and $\omega = \omega(G)$. Prove: (a) $e(G) \leqslant \frac{1}{2}(n-\omega)(n-\omega+1)$;

(b) If $e(G) > \frac{1}{2}(n-1)(n-2)$, then $G$ is connected.

# 6 Circuits and Cycles

- A circuit (cycle) of odd (even) length is called an **odd** or **even circuit** (**cycle**). Let $C_n$ denote a cycle of length $n$.
- A cycle of maximum (minimum) length in a graph is called a **longest** (or **shortest**) **cycle**.
- The length of the shortest cycle (or directed cycle) in $G$ (or $D$) is called the girth, denoted by $g(G)$ (or $g(D)$).
- The length of the longest cycle (or directed cycle) in $G$ (or $D$ ) is called the circumference, denoted by $c(G)$ (or $c(D)$).
- A simple graph $G$ (or $D$) is called **pancyclic** , if for every integer $k$ $(3 \leq k \leq |V|)$, $G$ (or $D$) contains a cycle (directed cycle) of length $k$.
- $G$ (or $D$) is called **vertex-pancyclic** , if for every vertex $u$ and every integer $k$ $(3 \leq k \leq |V|)$, $G$ (or $D$) contains a cycle (directed

cycle) of length $k$ containing $u$.

**Proposition 6.1** *Let $G$ be an undirected graph. If $\delta = \delta(G) \geq 2$, then $G$ contains a cycle. If $G$ is simple, then $G$ contains a cycle of length at least $\delta + 1$.*

**Proof** Omitted; see the preceding exercises.

**Theorem 1.5.1** (J. W. Moon, 1966) **Every strongly connected tournament of order at least 3 is vertex-pancyclic.**

**Proof** Let $D$ be a strongly connected tournament of order $n$ ($\geq 3$), $u$ any vertex of $D$, and $k$ ($3 \leq k \leq n$) any integer. Induct on $k \geq 3$.

- For $k = 3$, let $S = N_D^+(u), T = N_D^-(u)$. Since $D$ is strongly connected, $S \neq \emptyset, T \neq \emptyset$.

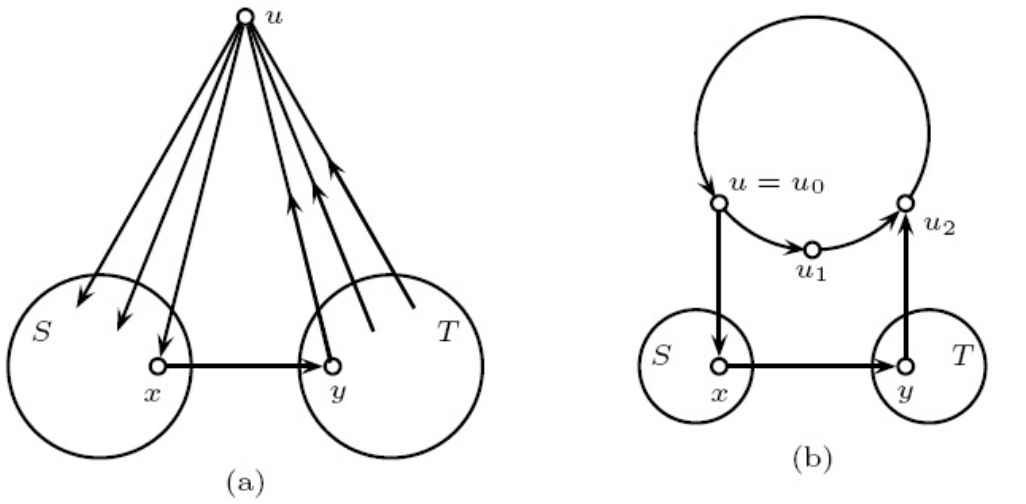



Figure 7: Illustration for the proof of Theorem 1.5.1

- Since $D$ is a tournament, $T \cup \{u\} = \overline{S}$. By Theorem 1.4.2(b), $E^+(S,T) = E^+(S,\overline{S}) \neq \emptyset$.
- Hence there exist $x \in S, y \in T$ with $(x,y) \in E(D)$; then $(u,x,y,u)$ is a directed cycle of length 3 in $D$ containing $u$ (see Figure 7(a)).

The base case holds.

- Assume $u$ lies on a directed cycle of length $k$ $(3 \leq k < n)$, we show that $u$ also lies on a directed cycle of length $k+1$. Let $C = (u_0, u_1, \cdots, u_k)$ be a directed cycle of length $k$ in $D$, where $u_0 = u_k = u$.

- If there exists $x \in V \setminus V(C)$ such that $N_D^+(x) \cap V(C) \neq \emptyset$ and $N_D^-(x) \cap V(C) \neq \emptyset$, then since $D$ is a tournament, there exist adjacent vertices $u_i$ and $u_{i+1}$ in $C$ such that $(u_i, x), (x, u_{i+1}) \in E(D)$. Then $u$ lies on the directed cycle $(u_0, u_1, \cdots, u_i, x, u_{i+1}, \cdots, u_k)$ of length $k+1$.

- Now suppose that for every $x \in V \setminus V(C)$, either $N_D^+(x) \cap V(C) = \emptyset$, or $N_D^-(x) \cap V(C) = \emptyset$. Let

$$S = \{x \in V \setminus V(C) : \ N_D^+(x) \cap V(C) = \emptyset\},$$
$$T = \{y \in V \setminus V(C) : \ N_D^-(y) \cap V(C) = \emptyset\}.$$

Clearly, $S \neq \emptyset, T \neq \emptyset$ and $E^+(S, T) \neq \emptyset$.

- Choose $x \in S, y \in T$ with $(x, y) \in E(D)$ (see Figure 7(b)). Then $u$ lies on the directed cycle $(u_0, x, y, u_2, \cdots, u_k)$ of length $k+1$. ∎

# 7 Distance and Diameter

- Let $G$ be a graph, $x, y \in V(G)$. The minimum length of all $x$-$y$ paths in $G$ is called the distance from $x$ to $y$, denoted by $d_G(x, y)$. A path whose length equals the distance is called a shortest path.

- If no $x$-$y$ path exists in $G$, we set $d_G(x, y) = \infty$.

$$\begin{aligned} d(G) = diam(G) &= \max\{d_G(x, y) : \forall\ x, y \in V(G)\} \\ &= \max_{x \in V} \max_{y \in V} d_G(x, y) \end{aligned}$$

  is called the diameter of $G$.

- Define the radius of $G$

$$r(G) = rad(G) = \min_{x \in V} \max_{y \in V} d_G(x, y).$$

A vertex $x \in V(G)$ is called a central vertex if $\max_{y \in V} d_G(x,y) = rad(G)$.

**Proposition 7.1 (Triangle inequality)** *For $x, y, z \in V(G)$,*

$$d_G(x,y) \leq d_G(x,z) + d_G(z,y).$$

**Proposition 7.2**

$$rad(G) \leq d(G) \leq 2rad(G).$$

**Proposition 7.3** *Let $G$ be a graph with $\Delta(G) \leq \Delta$ $(\Delta \geq 3)$ and $rad(G) = k$. Then*

$$|V(G)| < \frac{\Delta}{\Delta - 2}(\Delta - 1)^k.$$

**Proof:**

- Let $z$ be a central vertex of $G$ and let

$$N_i(z) = \{x \in V(G) \mid d_G(x,z) = i\}$$

for $i \in [0,k]$. Then $V(G) = \cup_{i=0}^{k} N_i(z)$.

- Moreover, $|N_0(z)| = 1, |N_1(z)| \leq \Delta, |N_2(z)| \leq \Delta(\Delta-1)$ and generally, $|N_i(z)| \leq \Delta(\Delta-1)^{i-1}$ for $1 \leq i \leq k$.

- Hence

$$\begin{aligned}
|V(G)| &= \sum_{i=0}^{k} |N_i(z)| \\
&\leq 1 + \sum_{i=1}^{k} \Delta(\Delta-1)^{i-1} \\
&= 1 + \frac{\Delta}{\Delta-2}((\Delta-1)^k - 1) \\
&< \frac{\Delta}{\Delta-2}(\Delta-1)^k.
\end{aligned}$$

**Proposition 7.4** *Let $G$ be a graph, then*

$$g(G) \leq 2d(G) + 1.$$

**Proof:** By contradiction.

- Assume $C$ is a shortest cycle with length $\geq 2d(G) + 2$. Then $C$ has two vertices whose distance in $C \geq d(G) + 1$.
- In $G$, these vertices have lesser distance; any shortest path $P$ between them is therefore not a subgraph of $C$.
- Let $xPy$ be the segment of $P$ between $x$ and $y$ such that $xPy$ contains no vertices of $C$ except $x$ and $y$.
- Together with the shorter of the two $x - y$ paths in $C$, this path $xPy$ forms a shorter cycle than $C$, a contradiction. ∎

# Bipartite Graph Characterization Theorem

**Theorem 1.5.2** (Bipartite Graph Characterization, König, 1936) A graph $G$ is bipartite $\Leftrightarrow$ $G$ contains no odd cycle.

**Proof** $(\Rightarrow)$

- Let $G$ be a bipartite graph with bipartition $\{X, Y\}$, and $C = x_0x_1\cdots x_{k-1}x_0$ be a cycle of length $k$ in $G$.
- Without loss of generality, assume $x_0 \in X$. Since $G$ is bipartite, we have $x_1 \in Y$, $x_2 \in X$, $x_3 \in Y$, $\cdots$. In general, $x_{2i} \in X$ and $x_{2i+1} \in Y$.
- Since $x_0 \in X$, we have $x_{k-1} \in Y$. That is, $k-1$ is odd, so $k$ is even. Therefore, $C$ is an even cycle.

$(\Leftarrow)$

- Without loss of generality, assume $G$ is connected.
- Since $G$ contains no odd cycle, $G$ has no loops. Without loss of generality, assume $|V(G)| \geq 2$. Take any $u \in V(G)$. Let

$$X = \{x \in V(G) : \ d_G(u,x) \text{ is even}\}$$
$$Y = \{y \in V(G) : \ d_G(u,y) \text{ is odd}\}.$$

- Clearly, $u \in X$ and $Y \neq \emptyset$, so $\{X, Y\}$ is a partition of $V(G)$.
- We now prove that $E(G[Y]) = \emptyset$. If $|Y| = 1$, this is self-evident.
- Assume $|Y| \geq 2$, and $y, z \in Y$ with $y \neq z$.
- Let $P$ and $Q$ be the shortest $u, y$-path and $u, z$-path, respectively,
- By the definition of $Y$, the lengths of both $P$ and $Q$ are odd.
- If $yz \in E(G)$, then $P \Delta Q \Delta yz$ (symmetric difference) is an odd circuit in $G$. Since a circuit can be decomposed into a union of

edge-disjoint cycles, $G$ contains an odd cycle, contradicting the hypothesis.

- Similarly, no two vertices in $X$ are adjacent, so $G$ is bipartite. ∎

**Corollary 1.5.2.2** A digraph $D$ is bipartite if and only if its underlying graph is bipartite.

**Proposition 7.5** *If an undirected graph $G$ contains no vertex of odd degree, then $G$ has a balanced orientation.*

**Proof**

- Induct on the number of edges $e(G) \geq 0$.
- When $e(G) = 0$, $G$ is an empty graph and the conclusion holds trivially.

- Assume $e(G) > 0$ and the conclusion holds for all undirected graphs of order $n$ with no odd-degree vertices and fewer than $e(G)$ edges.
- Let $S = \{x \in V(G) : d_G(x) = 0\}$, and set $G_1 = G - S$, then $\delta(G_1) \geq 2$ and $G_1$ has no odd-degree vertex.
- Then $G_1$ contains a cycle; let $C$ be a cycle in $G_1$.
- Orient the edges of $C$ consistently to obtain a directed cycle $C'$.
- Let $G_2 = G - E(C)$; then $|V(G_2)| = n$,

$$e(G_2) = e(G) - e(C) < e(G),$$

  and $G_2$ has no odd-degree vertex.
- By the induction hypothesis, $G_2$ has a balanced orientation $D'$. Then $D = D' \cup C'$ is a balanced orientation of $G$. ∎

**Proposition 7.6** *Any graph $G$ has an almost balanced orientation (that is $|d^+(x) - d^-(x)| \leq 1$ for any $x \in V(G)$).*

**Proof.**

- By the Handshaking Theorem, $G$ has an even number of vertices of odd degree. Hence we can add extra edges (for example red edges) to $G$ such that we obtain a graph $H$ which has no vertices of odd degree. More precisely, pair up the odd-degree vertices arbitrarily and add one new edge for each pair.
- By Proposition 7.5, $H$ has a balanced orientation $D$. Then delete the red edges from $D$, then we get an almost balanced orientation of $G$ (WHY?).

**Proposition 7.7** *Let $G$ be a non-bipartite simple undirected graph of order $n$, with $e(G) > \frac{1}{4}(n-1)^2 + 1$. Then $G$ contains a triangle.*

*Equivalently,* $\mathrm{ex}_{nb}(n, K_3) \leq \frac{1}{4}(n-1)^2+1$, *with equality if and only if* $G = C_5(\lceil \frac{n}{2} \rceil - 2, \lfloor \frac{n}{2} \rfloor - 1, 1, 1, 1)$.

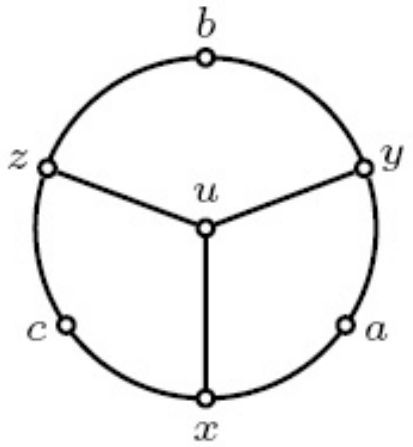
**Proof** (By contradiction)

- Let $G$ be a triangle-free non-bipartite graph. By Corollary 1.5.2.2, $G$ contains an odd cycle.
- Take a shortest odd cycle $C$ in $G$. Let $S = V(C)$; then $s = |S| \geq 5$.
- We first show

$$|(S, \overline{S})| \leqslant 2(n-s).$$

- (see below)
- By the choice of $C$, $G[S] = C$, i.e., $e(G[S]) = s$. Since $G$ is triangle-free, $G[\overline{S}]$ is also triangle-free. By Turán's theorem,

$$\begin{aligned} e(G) = &\quad e(G[S]) + |(S, \overline{S})| + e(G[\overline{S}]) \\ \leqslant &\quad s + 2(n-s) + \tfrac{1}{4}(n-s)^2 \leqslant \tfrac{1}{4}(n-1)^2 + 1, \end{aligned}$$




Otherwise, there exists $u \in \overline{S}$ with $|N_G(u) \cap S| \geq 3$. Let $x, y, z \in N_G(u) \cap S$. Since $G$ is triangle-free, there exist three vertices $a, b, c \in S \setminus \{x, y, z\}$. The cycles $C_1 = (u, x, \cdots, a, \cdots, y, u)$, $C_2 = (u, y, \cdots, b, \cdots, z, u)$ and $C_3 = (u, z, \cdots, c, \cdots, x, u)$ all have length less than $s$, and at least one is an odd cycle, contradicting the choice of $C$.

Figure 8: Example1.5.3

contradicting the hypothesis. Hence $G$ must contain a triangle.∎

# 8 Exercises

1.5.6 Let $G$ be a non-bipartite simple graph of order $n$, $k \geq 2$, and $\delta(G) > \left\lfloor \frac{2n}{2k+1} \right\rfloor$. Prove that $G$ contains an odd cycle of length $\leq 2k-1$.

# 9 Eulerian Graphs

- A trail (or directed trail) containing every edge of a graph is called an Euler trail (or Euler directed trail).
- An Euler trail (or Euler directed trail) with origin $x$ and terminus $y$ is called an Euler $xy$-trail (or Euler $(x, y)$-trail).
- A closed Euler trail (Euler directed trail) is called an Euler circuit (or Euler tour) (Euler directed circuit).
- An undirected graph containing an Euler circuit and a digraph containing an Euler directed circuit are collectively called Eulerian graphs.

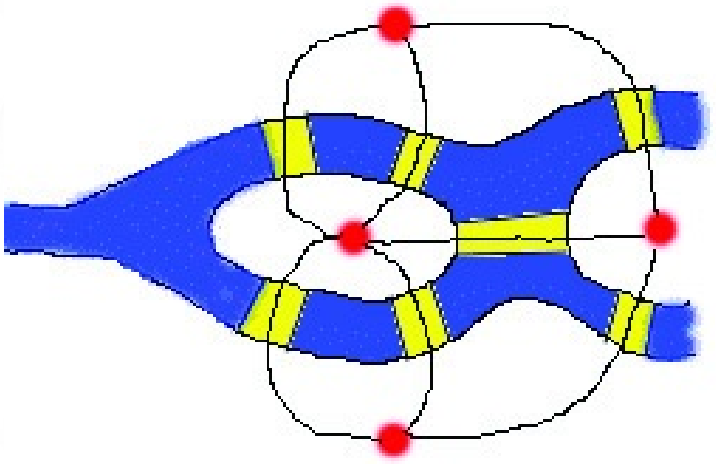


Figure 9: The K̈onigsberg seven bridges problem" reduces to deciding whether a graph is Eulerian.

## Characterization of Eulerian Graphs

The following result was first obtained by Euler in 1736 while solving the Königsberg seven bridges problem; Euler proved the necessity but did not provide a proof of sufficiency, which was given by C. Hierholzer (1873). (In fact, in 1873, Hierholzer had already passed away and it was

his colleague who wrote up the proof Hierholzer had told him before his death, and published it under Hierholzer's name alone.)

**Theorem 9.1 (Euler, 1736)** *$G$ is Eulerian if and only if $G$ is connected and has no vertex of odd degree.*

**Proof** ($\Rightarrow$) Clearly true since if a vertex $v$ appears $k$ times in an Euler tour of $G$, then $d_G(v) = 2k$.

($\Leftarrow$)

- Let
$$T = v_0 e_0 v_1 e_1 v_2 \cdots v_{l-1} e_{l-1} v_l$$
be a longest trail in $G$.

- Since $T$ cannot be extended, it already contains all the edges at $v_l$. Since there are even number of edges at $v_l$, $v_l = v_0$ (WHY). That is $T$ is closed.

- If $T$ contains all the edges of $G$, then we are done.
- Suppose to the contrary that there is an edge $e = uv_i$ outside of $T$ (such an edge exists since $G$ is connected). Then

$$uev_ie_iv_{i+1}\cdots v_{l-1}e_{l-1}v_le_0v_1\cdots v_{i-1}e_{i-1}v_i = T + e$$

  is a trail of length longer than $T$, a contradiction.

**Corollary 9.2** *A nontrivial graph $G$ has an Euler trail $\Leftrightarrow$ $G$ is connected and has at most two odd-degree vertices.*

**Proof** Omitted.

**Theorem 1.6** (Characterization of Directed Eulerian Graphs) (D. König, 1936) A digraph $D$ is Eulerian $\Leftrightarrow$ $D$ is connected and balanced.

**Proof** ($\Rightarrow$)

- Let $D$ be an Eulerian graph. By definition, $D$ is connected.
- Let $C$ be an Euler directed circuit in $D$.
- Since $C$ contains every edge and every vertex of the connected graph $D$, for any $x \in V(C)$, when traversing $C$ in its positive direction, each occurrence of vertex $x$ is entered by one edge and left by one edge. Then $d_D^+(x) = d_D^-(x)$, that is, $D$ is balanced.

($\Leftarrow$) Proof 1: Imitate the undirected case.

Proof 2: By contradiction.

- Let $D$ be a minimal counterexample of a balanced digraph with no directed Euler circuit.
- Since $\delta^+ = \delta^- \neq 0$, Therefore (Exercise 1.6.2), $D$ contains a directed cycle.

- Take a directed cycle $C$ in $D$. Let $D' = D - E(C)$ and

$$D'' = D' - \{x | d^+_{D'}(x) = d^-_{D'}(x) = 0\}.$$

  Then every nontrivial connected component $D_1, \cdots, D_k$ of $D''$ is still a connected balanced digraph with fewer edges than $D$.

- Thus each $D_i$ has a directed Euler circuit $C_i$, and $C, C_1, \cdots, C_k$ are pairwise edge-disjoint. Then $C \cup C_1 \cdots \cup C_k$ is an Euler directed circuit of $D$, contradicting the hypothesis. ∎

**Corollary 1.6.1** Let $D$ be a nontrivial connected digraph. Then $D$ has Euler $(x, y)$-trail $\Leftrightarrow$

$$d^+_D(x) - d^-_D(x) = 1 = d^-_D(y) - d^+_D(y),$$

and for every $u \in V(D) \setminus \{x, y\}$, $d^+_D(u) = d^-_D(u)$.

**Proof** Add to $D$ a directed edge $a$ from $y$ to $x$; call the resulting digraph $D'$.

$D$ has an Euler $(x,y)$-trail $\Leftrightarrow$ $D'$ has an Euler directed circuit.

$\Leftrightarrow$ $D'$ is balanced.

$\Leftrightarrow$ $d_D^+(u) = d_D^-(u),\ \forall\ u \in V(D) \setminus \{x,y\}$

$d_D^+(x) = d_D^-(x) + 1$

$d_D^-(y) = d_D^+(y) + 1.$

∎

## Exercises

1.6.8 Prove: a connected graph $G$ is Eulerian $\Leftrightarrow$ every edge of $G$ is contained in an odd number of cycles.

# 10 Hamiltonian Graphs

- A cycle (or directed cycle) containing every vertex of $G$ (or $D$) is called a Hamilton cycle (or Hamilton directed cycle).
- An undirected graph containing a Hamilton cycle and a digraph containing a Hamilton directed cycle are called Hamiltonian graphs. A graph with no Hamilton cycle is called non-Hamiltonian.
- For example, the complete graph $K_n$ and strongly connected tournaments are Hamiltonian (directed) graphs.
- The problem of deciding whether a given graph is Hamiltonian is called the Hamilton problem.
- The Hamilton problem is NP-complete; algorithmically, there is no polynomial-time algorithm to decide whether a graph is Hamiltonian or to find a Hamilton cycle. Moreover, unlike the Euler problem, we lack a necessary and sufficient condition.





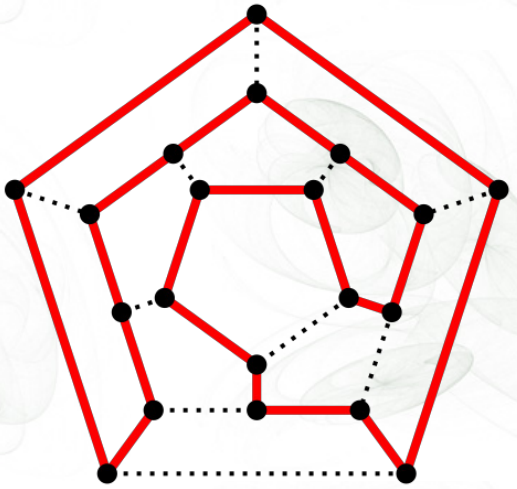



Figure 10: William Rowan Hamilton (1805–1865): Irish physicist, astronomer and mathematician. The Hamilton problem arose from the so-called "around-the-world" game. In 1856, The famous Irish mathematician Sir William Hamilton (1805-1865) designed a game: given 20 cities of the world, the 20 vertices of a dodecahedron (see figure) representing the Earth stand for these 20 cities. Starting from a vertex, travel along the edges of the dodecahedron, visiting each vertex exactly once, and finally return to the starting point. This problem reduces to finding a Hamilton cycle in the graph shown. Such a cycle exists, as indicated by the red edges.

Non-Hamiltonian graphs do exist; the Petersen graph is such an example.

**Example 1.7.1** The Petersen graph is not Hamiltonian.

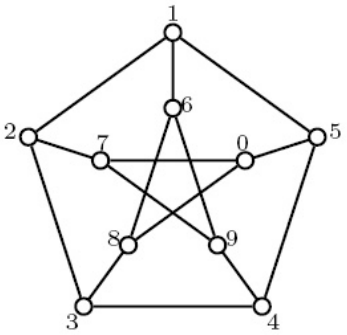



Figure 11: The Petersen graph is not Hamiltonian

Proof:

- Suppose the Petersen graph $G$ (shown in the figure) is Hamiltonian.

- Let the edge subset $T = \{16, 27, 38, 49, 50\}$; then $G - T$ is disconnected. Hence every Hamilton cycle in $G$ must contain an even number of edges in $T$.
- It is easy to see that a cycle containing only two edges of $T$ cannot be a Hamilton cycle. Thus every Hamilton cycle in $G$ must contain 4 edges of $T$.
- Without loss of generality, let $C$ be a Hamilton cycle containing $27, 38, 49, 50$ but not 16,
- Then $C$ must contain the edges $12, 15, 68, 69$, and the edges $23, 45$ must not be in $C$.
- Since vertices 3 and 4 are in $C$, the edge 34 must be in $C$. Then the edge set $\{34, 49, 96, 68, 83\}$ forms a cycle contained in $C$, which is impossible. Therefore, the Petersen graph is not Hamiltonian. ∎

**Theorem 1.7.1** **Let $G$ be a Hamiltonian graph. Then**

$$\omega(G-S) \leqslant |S|, \ \forall \ S \subset V(G).$$

**Proof** Let $C$ be a Hamilton cycle in $G$; then

$$\omega(C-S) \leqslant |S|, \ \forall \ S \subset V(C) = V(G).$$

Since $C-S$ is a spanning subgraph of $G-S$,

$$\omega(G-S) \leqslant \omega(C-S) \leqslant |S|, \ \forall \ S \subset V(G).$$

The theorem is proved. ∎

**Corollary 1.7.1.1** **Every Hamiltonian graph has no cut vertex.**

**Corollary 1.7.1.2** **Odd-order bipartite graphs are non-Hamiltonian.**

# Sufficient Conditions for Hamiltonian Graphs

**Theorem 1.7.2** (O. Ore, 1960, Ore's condition) Let $G$ be a simple graph of order $n$ ($\geq 3$). If for any two non-adjacent vertices $x$ and $y$, $d_G(x)+d_G(y) \geq n$, then $G$ is Hamiltonian.

**Proof Method 1**

- Let $G$ be a maximal non-Hamiltonian graph satisfying the hypothesis, i.e., $G$ is a non-Hamiltonian graph satisfying the hypothesis, but for any two non-adjacent vertices $x$ and $y$ in $G$, $G+xy$ is Hamiltonian.
- Since $G$ is not complete, it has non-adjacent vertices.
- Take two non-adjacent vertices $x$ and $y$ in $G$. By hypothesis $G+xy$ is Hamiltonian, and every Hamilton cycle in $G+xy$ contains the edge $xy$.

- Therefore, $G$ contains a Hamilton path connecting $x$ and $y$

$$P : (x =)x_1x_2\cdots x_{n-1}x_n(= y).$$

- Let

$$N^-(x) = \{x_i \in V(P) :\ xx_{i+1} \in E(G),\ 1 \leqslant i \leqslant n-2\},$$
$$N(y) = \{x_j \in V(P) :\ yx_j \in E(G),\ 2 \leqslant j \leqslant n-1\}.$$

  Since $G$ is simple, $|N^-(x)| = d_G(x), |N(y)| = d_G(y)$.

- Since $y \notin N^-(x) \cup N(y)$, $|N^-(x) \cup N(y)| \leq n-1$.

- We claim $|N^-(x) \cap N(y)| = 0$. Indeed, if some $x_i \in N^-(x) \cap N(y)$, then

$$C = (x, x_{i+1}, \cdots, x_{\upsilon-1}, y, x_i, \cdots, x_2, x)$$

  is a Hamilton cycle in $G$ (see Figure 12). contradicting the assumption that $G$ has no Hamilton cycle.

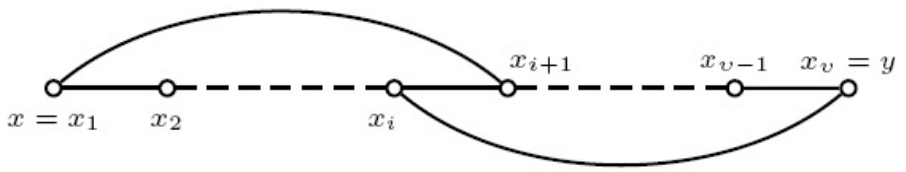



Figure 12: Illustration for the proof of Theorem 1.7.2

- Thus

$$d_G(x) + d_G(y) = |N^-(x)| + |N(y)| = |N^-(x) \cup N(y)| \leqslant n - 1.$$

contradicting the condition assumed in the theorem. Therefore, $G$ is Hamiltonian.

**Method 2** The argument provided below is useful for understanding Theorem 1.7.3 and its proof.

- It follows from the hypothesis that $G$ is connected and contains a cycle. Let $C = (x_1, x_2, \cdots, x_k, x_1)$ be a longest cycle in $G$.
- If $k < n$, let $R = V(G) \setminus \{x_1, x_2, \cdots, x_k\}$. Since $G$ is connected, there exists $y \in R$ adjacent to some vertex of $C$; assume it is $x_k$.
- By the maximality of $C$, $x_1 y \notin E(G)$.
- Let

$$S = \{x_i \in V(C) : \ x_1 x_{i+1} \in E(G), \ 1 \leqslant i \leqslant k-1\},$$
$$T = \{x_j \in V(C) : \ y x_j \in E(G), \ 2 \leqslant j \leqslant k\}.$$

- Then

$$S \cap T = \emptyset$$

because if some $x_i \in S \cap T$, then $G$ contains a cycle

$$C' = (x_1, x_{i+1}, \cdots, x_k, y, x_i, \cdots, x_2, x_1)$$

with $|C'| = |C| + 1$, contradicting the maximality of $C$.

- Hence $|S \cup T| \leqslant k$. (1)

- By equation (1), the number of neighbors of $x_1$ and $y$ on $C$, $d_C(x_1)$ and $d_C(y)$, satisfy:

$$d_C(x_1) + d_C(y) = |S| + |T| = |S \cup T| \leqslant k. \quad (2)$$

- On the other hand, by the maximality of $C$, for any $z \in R$, $z \neq y$, at most one of the edges $x_1 z$ and $yz$ exists.

- So the numbers of neighbors of $x_1$ and $y$ in $R$, $d_R(x_1)$ and $d_R(y)$, satisfy:

$$d_R(x_1) + d_R(y) \leqslant n - k - 1. \quad (3)$$

- By equations (2) and (3), we have

$$d_G(x_1)+d_G(y)\le k+n-k-1=n-1$$

contradicting the hypothesis. Hence $k=n$, i.e., $C$ is a Hamilton cycle in $G$. ∎

**Method 3:** A corollary of Theorem 5.2.

**Corollary 1.7.2** (G. A. Dirac, 1952, Dirac's condition) Let $G$ be a simple graph of order $n$. If $\delta(G)\ge\frac{n}{2}$, $n\ge 3$, then $G$ is Hamiltonian. ∎

(Prove Dirac's theorem without relying on Ore's condition!)

**Theorem 10.1 (A. Ghouila-Houri, 1960)** *Let $D$ be a strongly connected simple digraph of order $n$. If $\delta^0(D)=\min\{\delta^+(D),\delta^-(D)\}\ge\frac{n}{2}$, then $D$ is Hamiltonian.*

**Proof:**

- By contradiction.
- Choose a longest directed cycle $C$ in $D$. Then $|V(C)| \geq \frac{n}{2}$.
- Choose a longest dipath $P = v_1 v_2 \cdots v_k$ in $D - V(C)$. Then $N^-(v_1), N^+(v_k) \subseteq V(C) \cup V(P)$.
- Let $S = N^-(v_1) \cap V(C)$ and $T = N^+(v_k) \cap V(C)$. Then $T \cap S^+ = \emptyset$ and $dist_C(S^+, T) \geq k$, where $dist_C(S^+, T) = \min\{dist_C(s,t) : s \in S^+, t \in T\}$.
- Hence $|T| \leq |V(C)| - |S^+| - k + 1 \leq |V(C)| - \frac{n}{2}$ (WHY?).
- And since $\frac{n}{2} \leq d^+(v_k) \leq |T| + k - 1 \leq |V(C)| - \frac{n}{2} + k - 1$, we have $|V(C)| + k \geq n + 1$, a contradiction.

**Theorem 10.2 (H. Meyniel, 1973)** *Let $D$ be a strongly connected simple digraph of order $n$. If for any two non-adjacent vertices $x$ and $y$ in $D$, $d_D(x)+d_D(y) \geq 2n-1$, then $D$ is Hamiltonian.*

**Proof:** USING Bondy-Thomassen Theorem. (self-study)

**Corollary 1.7.3.4** (P. Camion, 1959) **Every strongly connected tournament is Hamiltonian**

**Corollary 1.7.3.6** (L. Rédei, 1934) **Every tournament contains a Hamilton path.**

**Conjecture 10.3 (Pósa-Seymour conjecture, 1974)** *Let $G$ be a graph on $n \geq 3$ vertices, and let $k$ be a positive integer. If*

$$\delta(G) \geq \frac{k}{k+1} n,$$

*then $G$ contains the $k$-th power of a Hamilton cycle; that is, there exists a Hamilton cycle $H$ in $G$ such that*

$$H^k \subseteq G,$$

*where $H^k$ is the graph on $V(H)$ in which two distinct vertices are adjacent if and only if their distance on $H$ is at most $k$.*

For $k = 1$, this is precisely Dirac's Theorem.

Remark: Komlós, Sárközy and Szemerédi (1998) proved the conjecture for every fixed $k$ and all sufficiently large $n$.

# Exercises

1.7.7 Let $G$ be a simple graph with no cut vertex and at least $n \geq 3$ vertices. Prove: if for any two vertices $x$ and $y$ at distance 2, $\max\{d_G(x), d_G(y)\} \geq \frac{1}{2}n$ (called the Fan condition), then $G$ is Hamiltonian

[See Geng-Hua Fan, New sufficient conditions for cycles in graphs J. Combin. Theory Ser. B, 37 (1984), pp. 221-227].

# 11 P and NP

- The class of problems solvable in polynomial time is usually denoted by $P$.
- The class $NP$ consists, roughly speaking, of all those questions with the property that for any input that has a positive answer, there is a 'certificate' from which the correctness of this answer can be derived in polynomial time. Clearly, $P \subseteq NP$.
- For instance, the question:

  'Given a graph $G$, is $G$ Hamiltonian?'

  belongs to NP. (WHY)
- In particular, NP does not mean: "non-polynomial time", the letters NP stand for "nondeterministic polynomial-time".

# NP-complete

- A problem $\Pi$ is *NP-complete* if $\Pi \in NP$ and every problem in NP can be reduced to $\Pi$, in polynomial time.
- It implies that if one NP-complete problem can be proved to be solvable in polynomial time, then each problem in NP can be solved in polynomial time. In other words: then $P = NP$ would follow.
- A problem $\Pi$ is NP-hard if every problem in NP can be reduced in polynomial time to $\Pi$.
- For example, the halting problem:

  "Given a program and its input, will it run forever?"

  belongs to NP-hard but not belong to NP-complete.

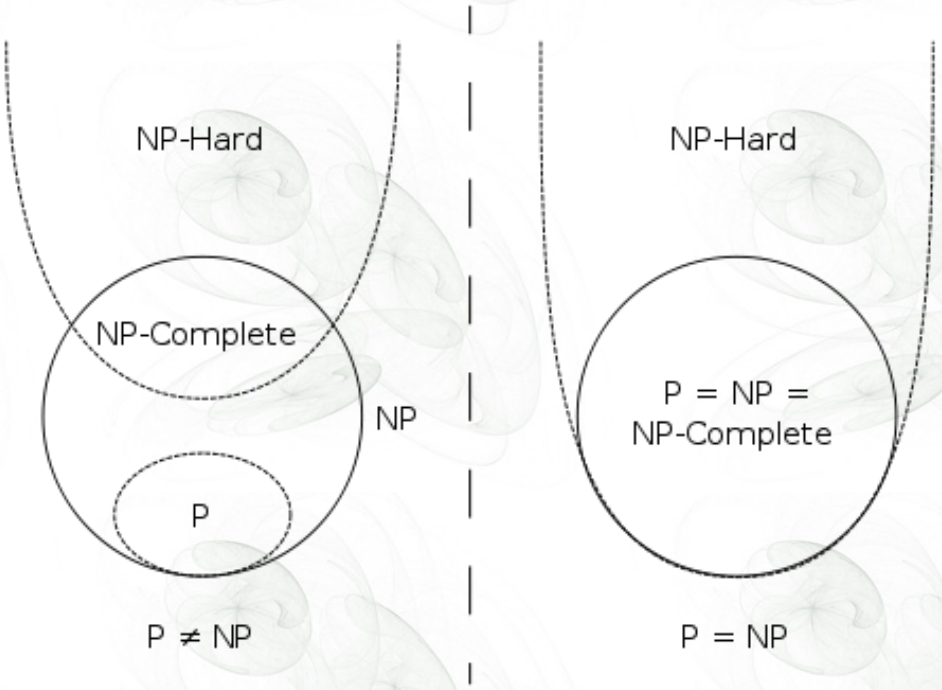



Figure 13: Relationship between P, NP, NP-Complete and NP-Hard

## Well Known NP-Complete problems

- SAT: Given a Boolean expression $f(x_1, \ldots, x_n)$, if there is $\alpha_1, \ldots, \alpha_n \in \{0, 1\}$ such that $f(\alpha_1, \ldots, \alpha_n) = 1$. SAT is the first

problem that was proven to be NP-complete

- 3-SAT: Determining the satisfiability of a Boolean expression in conjunctive normal form where each clause is limited to at most three literals is called 3-SAT.
- 3-SAT$\trianglelefteq$ Hamiltonian Cycle Problem.

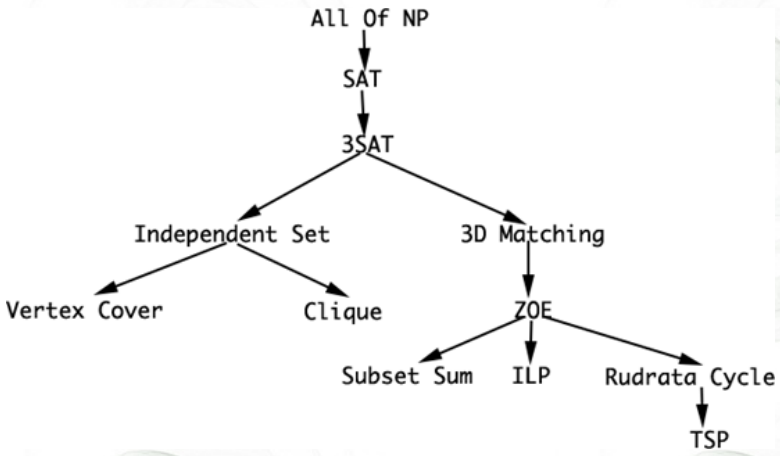



Figure 14: Reductions between NPCs

# A proof of 3-SAT to HCP

- Construct $n$ paths $P_1, \ldots, P_n$ corresponding to the $n$ variables.
- Each path $P_i$ consists of $2k$ vertices $v_{i,1}, \ldots, v_{i,2k}$, where $k$ is the number of clauses in the Boolean expression.
- Add edges from $v_{i,j-1}$ to $v_{i,j}$ (resp. $v_{i,j}$ to $v_{i,j-1}$) on $P_i$ to correspond to the assignment $x_i = 1$ (resp. $x_i = 0$).
- Add edges from $\{v_{i,1}, v_{i,2k}\}$ to $\{v_{i+1,1}, v_{i+1,2k}\}$.
- Add source $s$ and target $t$, and add edges from $s$ to $v_{1,1}, v_{1,2k}$, and from $v_{n,1}, v_{n,2k}$ to $t$, and from $t$ to $s$.
- Add $k$ vertices $c_1, \ldots, c_k$ to correspond to the $k$ clauses in the Boolean expression.
- If a clause $c_j$ contains the variable $x_i$, add the following directed paths:

1. if $c_j$ contains $x_i$, add

$$v_{i,2j-1} \to c_j \to v_{i,2j};$$

2. if $c_j$ contains $\overline{x_i}$, add

$$v_{i,2j} \to c_j \to v_{i,2j-1}.$$

- For example,

  $f(x_1, x_2, x_3, x_4) = (x_1 \vee x_2 \vee \bar{x_3}) \wedge (\bar{x_2} \vee x_3 \vee \bar{x_4}) \wedge (\bar{x_1} \vee x_2 \vee \bar{x_4})$

  and its corresponding graph are shown in Fig 15.

- Any Hamilton cycle in the constructed graph $G$ traverses $P_i$ either from right to left ($x_i = 0$) or left to right ($x_i = 1$).

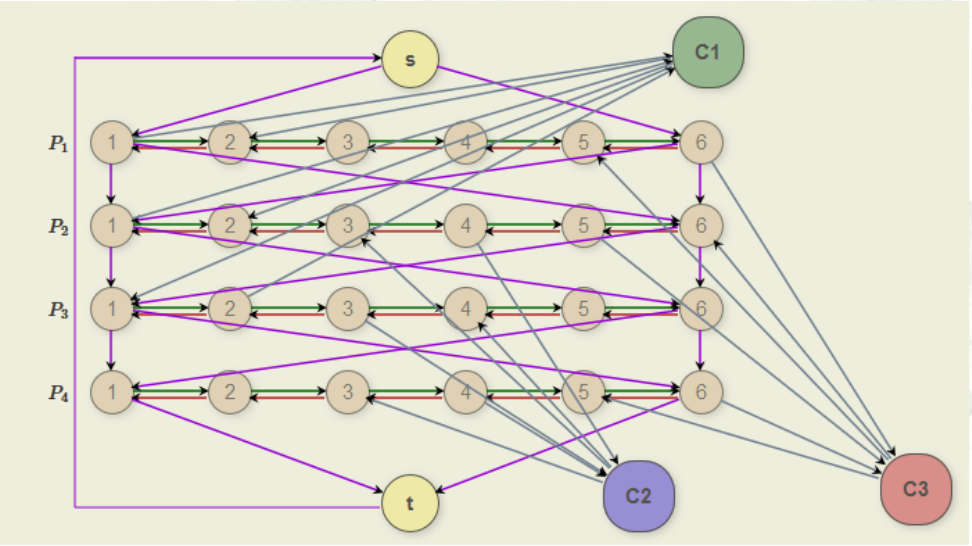



Figure 15: 3SAT to HCP

- Fig. 16 shows a corresponding between a HC and an assignment of variables $x_1, x_2, \ldots, x_n$.

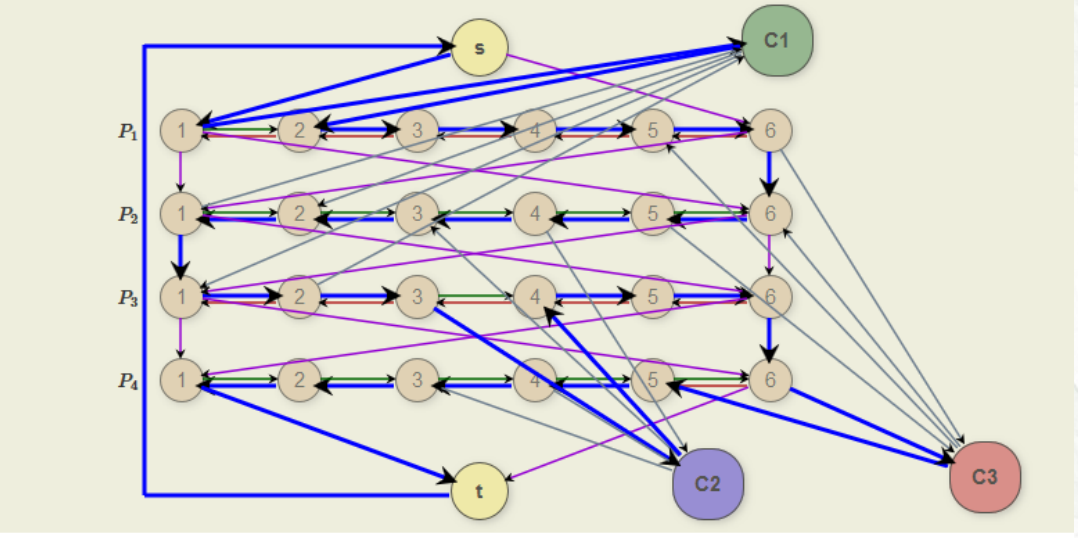



Figure 16: Hamiltonian Cycle and its corresponding assignment

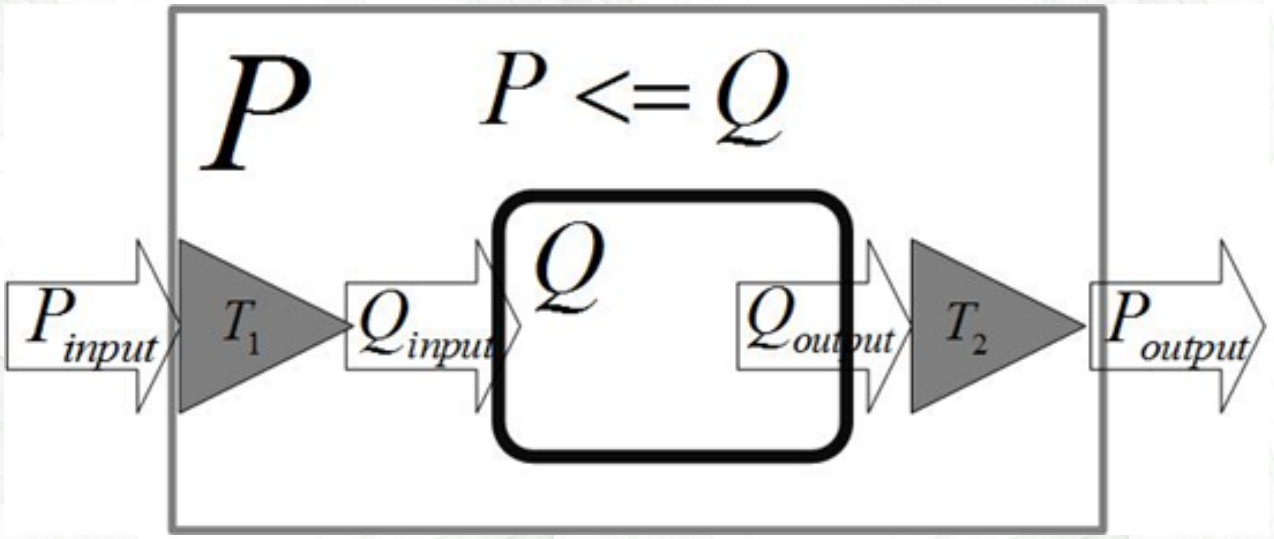



Figure 17: Reduction procedure

# 12 Distance and Diameter

We recall the concepts of distance and diameter.

- For a digraph $D$, in general, $d_D(x,y) \neq d_D(y,x)$.

  $$d(D) = \max\{d_D(x,y) : \forall\ x,y \in V(D)\}$$

  is called the **diameter** of $D$.

- For an undirected graph $G$, $d_G(x,y) = d_G(y,x)$.

  $$d(G) = \max\{d_G(x,y) : \forall\ x,y \in V(G)\}.$$

- For a directed path $P_n$ on $n$ vertices, its diameter $d(P_n) = \infty$. Clearly, $d(D) = 1 \Leftrightarrow K_n^* \subseteq D$.

- For an undirected path $P_n$ on $n$ vertices, its diameter $d(P_n) = n-1$; and likewise $d(G) = 1 \Leftrightarrow K_n \subseteq G$.

# Lower Bounds

**Proposition 12.1** *Let $D$ be a strongly connected digraph. If its order $n \geq 2$ and maximum degree $\Delta = \max\{\Delta^+, \Delta^-\} \geq 1$, then*

$$d(D) \geqslant \begin{cases} n-1, & \text{if } \Delta = 1; \\ \lceil \log_\Delta(n(\Delta-1)+1) \rceil - 1, & \text{if } \Delta \geq 2. \end{cases}$$

**Proof**

- Since $G$ is strongly connected, its diameter $d(G)$ is well-defined. Let $d(G) = d$.
- For a fixed vertex $x$, the number of vertices at distance 1 from $x$ is at most $\Delta$. In general, the number at distance $i$ is at most $\Delta^i$. Therefore,

$$n \leqslant 1 + \Delta + \Delta^2 + \cdots + \Delta^{d-1} + \Delta^d$$

$$= \begin{cases} d+1, & \text{if } \Delta = 1; \\ \dfrac{\Delta^{d+1}-1}{\Delta-1}, & \text{if } \Delta \geq 2. \end{cases} \tag{4}$$

- When $\Delta = 1$, equation (4) gives $n \leq d+1$, i.e., $d(G) = d \geq n-1$. On the other hand, clearly, $d(G) \leq n-1$. Hence $d(G) = n-1$.
- When $\Delta \geq 2$, equation (4) gives $(\Delta-1)n \leq \Delta^{d+1}-1$, i.e.,

$$d(G) = d \geqslant \lceil \log_{\Delta}(n(\Delta-1)+1) \rceil - 1.$$

as required. ∎

**Proposition 12.2** *Let $G$ be a connected undirected graph. If its order $n \geq 3$ and maximum degree $\Delta \geq 2$, then*

$$d(G) \geqslant \begin{cases} \lfloor \frac{1}{2} n \rfloor, & \text{if } \Delta = 2; \\ \left\lceil \log_{(\Delta-1)} \dfrac{n(\Delta-2)+2}{\Delta} \right\rceil, & \text{if } \Delta \geq 3. \end{cases}$$

**Proof**

- Since $G$ is connected, its diameter $d(G) = d < \infty$.
- For a fixed vertex $x$, the number of vertices at distance 1 from $x$ is at most $\Delta$; at distance 2 at most $\Delta(\Delta-1)$. In general, at distance $i$ at most $\Delta(\Delta-1)^{i-1}$. Hence

$$\begin{aligned} n &\leqslant 1 + \Delta + \Delta(\Delta-1) + \cdots + \Delta(\Delta-1)^{d-1} \\ &= \begin{cases} 2d+1, & \text{if } \Delta = 2; \\ \dfrac{\Delta(\Delta-1)^d - 2}{\Delta - 2}, & \text{if } \Delta \geq 3. \end{cases} \end{aligned} \tag{5}$$

- When $\Delta = 2$, equation (5) gives $n \leq 2d+1$, i.e., $d(G) = d \geq \lfloor \frac{1}{2}n \rfloor$.
- When $\Delta \geq 3$, equation (5) gives $(\Delta-2)n \leq \Delta(\Delta-1)^d - 2$, i.e.,

$$d(G) = d \geqslant \left\lceil \log_{(\Delta-1)} \frac{n(\Delta-2)+2}{\Delta} \right\rceil.$$

as required. ∎

## $(\Delta, d)$-Moore Graphs

- In fact, equations (4) and (5) give an upper bound on the order of a (directed) graph with given maximum degree and diameter, called the $(\Delta, d)$-**Moore bound**. Denote it by

$$M(\Delta, d) = \max\{|V(G)| \mid \Delta(G) = \Delta, d(G) = d\}.$$

- A digraph attaining the $(\Delta, d)$-Moore bound (4) is called a $(\Delta, d)$-**Moore digraph**.
- An undirected graph attaining the $(\Delta, d)$-Moore bound (5) is called a $(\Delta, d)$-**Moore undirected graph**.
- For example, the directed cycle $C_{d+1}$ is a $(1, d)$-Moore digraph; the complete digraph $K^*_{\Delta+1}$ is a $(\Delta, 1)$-Moore digraph.
- The undirected cycle $C_{2d+1}$ is a $(2, d)$-Moore undirected graph; the complete undirected graph $K_{\Delta+1}$ is a $(\Delta, 1)$-Moore undirected graph.
- The Petersen graph is a $(3, 2)$-Moore undirected graph.
- In fact, few $(\Delta, d)$-Moore graphs are known. In Example 1.9.4, We will prove: for given $\Delta \geq 2$ and $d \geq 2$, no $(\Delta, d)$-Moore digraph exists.

# An Upper Bound

**Proposition 12.3** *Let $G$ be a connected undirected graph. If its order is $n$ and its minimum degree is $\delta$, then*

$$d(G) \leqslant \frac{3n}{\delta+1}.$$

**Proof**

- Since for any connected undirected graph $G$ of order $n$, its diameter $d(G) \leq n-1$ and $n \geq \delta+1$, we may assume $\delta \geq 3$ and $d(G) = d \geq 4$.
- Choose $x, y \in V(G)$ with $d_G(x,y) = d = d(G)$, and let $P = (x_0, x_1, \cdots, x_{d-1}, x_d)$ be a shortest $(x,y)$-path in $G$, where $x_0 = x$ and $x_d = y$.
- By the minimality of $P$, for any two vertices $x_{3i}$ and $x_{3j}$ $(i \neq j)$, $N_G(x_{3i}) \cap N_G(x_{3j}) = \emptyset$.

- Let $h = \lfloor \frac{d}{3} \rfloor$. Then

$$\begin{aligned} n &\geqslant \sum_{i=0}^{h} |N_G(x_{3i}) \cup \{x_{3i}\}| \geq (h+1)(\delta+1) \\ &= \left( \left\lfloor \frac{d}{3} \right\rfloor + 1 \right) (\delta+1) \geqslant \frac{d}{3}(\delta+1). \end{aligned}$$

This proves the proposition. ∎

**Proposition 12.4 (O. Ore, 1968)** *Let $G$ be a connected undirected graph of order $n$ and diameter $d$. Then*

$$e(G) \leqslant d + \frac{1}{2}(n-d+4)(n-d-1).$$

**Proof**

- Let $x$ and $y$ be vertices of $G$ with $d_G(x,y)=d$, and $P$ a shortest $xy$-path.
- Let $Z=V(G)\backslash V(P)$. Then for any $z\in Z$, $|N_G(z)\cap V(P)|\leq 3$.
- Therefore,

$$\begin{aligned} e(G) &= e(P)+e(G[Z])+|E_G(Z,P)| \\ &\leqslant d+\tbinom{n-d-1}{2}+3(n-d-1) \\ &= d+\tfrac{1}{2}(n-d-1)(n-d-2)+3(n-d-1) \\ &= d+\tfrac{1}{2}(n-d+4)(n-d-1) \end{aligned}$$

This proves the proposition. ∎

- Remarks: (1) We have not characterized the extremal graphs here; can one characterize extremal graphs attaining the upper bound? Explore lower bounds on edges and extremal graphs for given diameter and order; explore extremal problems for given radius and order.

(2) The result is an induced version of Turán Problem: $\mathrm{ex}(n, IP_d) = \frac{1}{2}(n-d+4)(n-d-1)+d$.

(3) What's the value of $\mathrm{ex}(n, IC_{\geq d}) = ?$

## Average Distance

- Let $G$ be a nontrivial connected undirected graph or strongly connected digraph of order $n$. The average distance (or mean distance) of $G$, denoted $m(G)$, is defined by

$$m(G) = \frac{1}{n(n-1)} \sum_{x,y \in V} d_G(x,y).$$

- Since $m(G)$ is the arithmetic mean of all nonzero distances, $m(G) \geq 1$, with equality if and only if $G$ is complete.

- Let
$$\sigma(G) = \sum_{x,y \in V} d_G(x,y).$$
denotes the sum of distances between all distinct pairs of vertices in $G$.

- For example, for the star $K_{1,n-1}$ of order $n$, one easily computes $\sigma(K_{1,n-1}) = 2(n-1)^2$, $m(K_{1,n-1}) = 2(n-1)/n$.

- Consider the directed cycle $C_n$ $(n \geq 3)$ and a fixed vertex $x_0$ in $C_n$. Since the sum of distances from $x_0$ to every other vertex is $1+2+\cdots+n-1 = \frac{1}{2}n(n-1)$,
$$\sigma(C_n) = n^2(n-1)/2, \qquad m(C_n) = n/2.$$

The following example gives a lower bound for $\sigma(G)$; the undirected version is due to R. C. Entringer, D. E. Jackson, P. J. Slater (1977); the digraph version is due to C. P. Ng, H. H. Teh (1978).

**Example 1.8.5** Suppose $G$ have $n$ vertices and $m$ edges.

(a) If $G$ is an undirected graph, then $\sigma(G) \geq 2n(n-1) - 2m$;

(b) If $G$ is a digraph, then $\sigma(G) \geq 2n(n-1) - m$.

Moreover, equality holds when $d(G) \leq 2$.

**Proof**

- Let the undirected graph $G$ be obtained from the complete graph $K_n$ by deleting $p$ edges. Thus,

$$m = e(G) = \frac{1}{2} n(n-1) - p. \tag{6}$$

- Let the deleted edges be $x_1y_1, x_2y_2, \cdots, x_py_p$. Then

$$\sigma(G) = 2 \sum_{xy \in E(G)} d_G(x,y) + 2 \sum_{i=1}^{p} d_G(x_i, y_i). \tag{7}$$

- Clearly, if $xy \in E(G)$, then $d_G(x,y) = 1$; otherwise, $d_G(x,y) \geq 2$. Therefore, from equation (7),

$$\sigma(G) \geqslant 2m + 4p. \tag{8}$$

  Combining equations (6) and (8) yields conclusion (a).

- Similarly, considering the digraph $D$, $m = e(D) = e(K_n^*) - p = n(n-1) - p$.

- Thus,

$$\begin{aligned} \sigma(D) &= \sum_{(x,y)\in E(D)} d_D(x,y) + \sum_{i=1}^{p} d_D(x_i,y_i) \\ &\geqslant m + 2p = 2n(n-1) - m. \end{aligned}$$

- As can be seen from the proof, the last conclusion is clearly true.

∎

As for an upper bound on $m(G)$, M. Kouider and P. Winkler (1997) obtained the following result:

$$m(G) \leqslant \frac{n}{\delta + 1} + 2.$$

But the proof is somewhat difficult; interested readers may refer to the original paper.

## Line Graphs

- The line graph of a nonempty digraph $D$, denoted $L(D)$, is the graph with vertex set $E(D)$, and for $a_i, a_j \in E(D)$, then

$$(a_i, a_j) \in E(L(D)) \Leftrightarrow \text{ the head of } a_i \text{ is the tail of } a_j \text{ in } D.$$

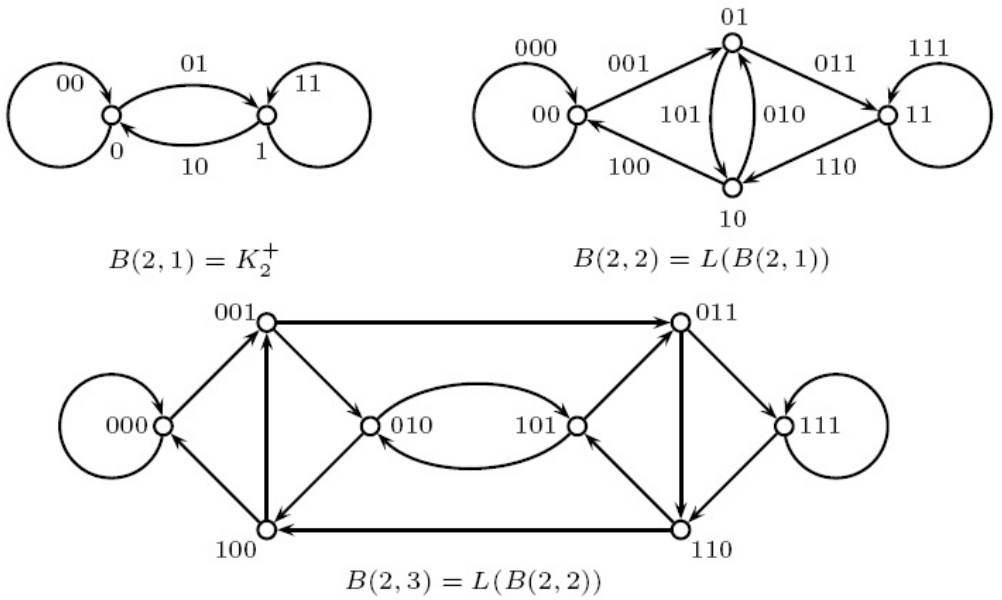



Figure 18: A digraph $D$ and its line graph $L(D)$

- Let $L(D)$ be the line graph of $D$. If $L(D)$ is nonempty with no isolated vertex, then $L(D)$ has a line graph $L(L(D))$. Recursively define the $n$-th iterated line graph

$$L^n(D) = L(L^{n-1}(D)),$$

where $L^0(D)$ and $L^1(D)$ are $D$ and $L(D)$, respectively, assuming $L^{n-1}(D)$ is nonempty with no isolated vertex.

- For example, de Bruijn digraphs and Kautz digraphs

$$B(d,n) = L^{n-1}(B(d,1)) = L^{n-1}(K_n^+);$$

$$K(d,n) = L^{n-1}(K(d,1)) = L^{n-1}(K_{d+1}^*).$$

- A graph $G$ is claw-free if $G$ does not contain a $K_{1,3}$ as an induced subgraph of $G$.

**Proposition 12.5** *The line graph $L(G)$ of any graph $G$ is claw-free.*

**Proof:** Homework.

**Theorem 1.8.1** (M. Aigner, 1967) Let $D$ are nontrivial strongly connected digraphs; let $L$ be its line graph. Then

$$d(D) \leqslant d(L) \leqslant d(D) + 1.$$

Moreover, $d(D) = d(L) \Leftrightarrow D$ is a directed

**Proof**

- Since $D$ is strongly connected, so is $L$; hence $d(D)$ and $d(L)$ are both well-defined.
- Choose $x, y \in V(D)$ with $d_D(x, y) = d(D)$, and let $P$ be a shortest $(x, y)$-path in $D$.
- Let $a \in E_D^-(x)$, $b \in E_D^-(y) \cap E(P)$, and let $Q = a + P$.
- Then $L(Q)$ is a shortest $(a, b)$-path in $L$. Thus,

$$d(L) \geqslant e(L(Q)) = e(P) = d(D). \tag{9}$$

- On the other hand, choose $a, b \in V(L)$ with $d_L(a,b) = d(L)$. There exist $x, y, z, u \in V(D)$ with $a = (z,x), b = (y,u)$. Thus,

$$d(D) \geqslant d_D(x,y) = d_L(a,b) - 1 = d(L) - 1. \tag{10}$$

- From (9) and (10), $d(D) \leq d(L) \leqslant d(D) + 1$.

We now prove: $d(D) = d(L) \Leftrightarrow D$ is a directed cycle.

- If $D$ is a directed cycle, then $L \cong D$, so $d(D) = d(L)$.
- It suffices to prove: if $D$ is not a directed cycle, Then $d(L) = d(D) + 1$.
- To this end, let $d = d(D)$, choose $x, y \in V(D)$ with $d_D(x,y) = d$. Let $P$ be a shortest $(x,y)$-path in $D$.
- Since $D$ is strongly connected, $d_D^-(x) \geq 1$ and $d_D^+(y) \geq 1$.

- If there exist $x', y' \in V(D)$ with $a = (x', x)$, $b = (y, y')$ and $a \neq b$, Then $d_L(a, b) = d + 1$.
- Otherwise, there exists $c \in E(D)$ with $c = (y, x)$. Then $P \cup \{c\}$ is a directed cycle in $D$, written $C = (x_0, x_1, \cdots, x_d, x_0)$, where $x_0 = x$ and $x_d = y$.
- Since $D$ is strongly connected but not a directed cycle, there exist $x_i \in V(C)$ and $z \in V(D)$ with $(x_i, z) \in E(D)$ ($z$ may be some $x_j$ $(0 \leq j \leq d)$).
- Choose such $x_i$ with $i$ $(0 \leq i \leq d)$ as large as possible. Then $d_D(x_{i+1}, x_i) = d$. Let $a = (x_i, x_{i+1})$ and $b = (x_i, z)$. Then $d_L(a, b) = d + 1$.
- Thus, $d + 1 \geqslant d(L) \geqslant d_L(a, b) = d + 1$. This means $d(L) = d + 1$. ∎

**Remark.** Line graphs are characterized by nine forbidden subgraphs and can be recognized in linear time.

## Exercises

1.8.1 Let $x, y, z$ be three distinct vertices of a strongly connected digraph $D$. Prove: $d_D(x,z) \leqslant d_D(x,y) + d_D(y,z)$.

# 13 Matrix Representation of Graphs

Graphs are often stored in computers via their adjacency matrices.

- The adjacency matrix is the $n \times n$ matrix:

$$\mathbf{A} = (a_{ij})_{n\times n}, \quad \text{where } a_{ij} = \mu(x_i, x_j),$$

  where $\mu(x_i, x_j)$ is the number of directed edges in digraph $D$ from $x_i$ to $x_j$, or the number of edges joining $x_i$ and $x_j$ in undirected graph $G$.

- The adjacency matrices of digraph $D$ and undirected graph $G$ are denoted $\mathbf{A}(D)$ and $\mathbf{A}(G)$.

- $\mathbf{A}(G)$ is symmetric, while $\mathbf{A}(D)$ is generally not symmetric.

- The incidence matrix of a graph is the $n \times m$ matrix:

$$\mathbf{M} = (m_{ij})_{n\times m},$$

where $v_i \in V, e_j \in E$, and for a loopless digraph $D$,

$$m_{ij} = \begin{cases} 1, & v_i \text{ is the tail of } e_j, \\ -1, & v_i \text{ is the head of } e_j, \\ 0, & \text{otherwise;} \end{cases}$$

while for undirected graph $G$,

$$m_{ij} = \begin{cases} 1, & v_i \text{ is an endpoint of } e_j, \\ 0, & \text{otherwise.} \end{cases}$$

The incidence matrices of digraph $D$ and undirected graph $G$ are denoted $\mathbf{M}(D)$ and $\mathbf{M}(G)$.

- The adjacency matrices of any two isomorphic graphs are permutation-similar, while their incidence matrices are permutation-equivalent.

- Let $\mathbf{A}$ be the adjacency matrix of a loopless undirected graph $G$ with vertex set $V = \{x_1, x_2, \cdots, x_n\}$; let $\mathbf{B}$ be the $n \times n$ diagonal

matrix with diagonal entries $b_{ii} = d_G(x_i)$. The matrix $\mathbf{B} - \mathbf{A}$ is called the Laplacian matrix of $G$. It is easy to verify

$$\mathbf{B} - \mathbf{A} = \mathbf{M}\mathbf{M}^T.$$

Here $M$ is the incidence matrix of any orientation of $G$.

- Matrix representations of graphs build a bridge between graph theory and matrix theory. With such representations, problems in graph theory can be studied using matrix theory and analytical methods. Conversely, problems in matrix theory can be studied using graph theory and analytical methods.

$$\mathbf{A}(D) = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{A}(G) = \begin{pmatrix} 0 & 2 & 1 & 1 \\ 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}.$$

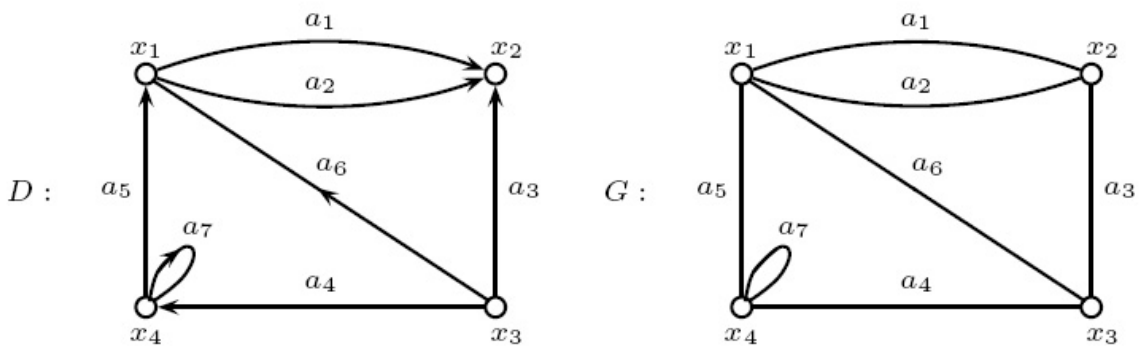



Figure 19: A digraph $D$ and an undirected graph $G$

$$\mathbf{M}(D - a_7) = \begin{pmatrix} 1 & 1 & 0 & 0 & -1 & -1 \\ -1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 1 & 0 \end{pmatrix},$$

$$\mathbf{M}(G) = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 2 \end{pmatrix}.$$

**Theorem 1.9** (Adjacency Matrix Theorem) Let $\mathbf{A}$ be the adjacency matrix of digraph $D$, then $(\mathbf{A}^k)_{ij}$ is the number of $(x_i, x_j)$-walks of length $k$ in $D$.

**Proof** By the matrix multiplication formula,

$$a_{ij}^{(k)} = \sum_{1 \le i_1, \cdots, i_{k-1} \le n} a_{ii_1} a_{i_1 i_2} \cdots a_{i_{k-1} j} \tag{11}$$

the result follows immediately. ∎

## Primitive Matrices and Primitive Exponents

- A nonnegative square matrix $\mathbf{A}$ of order $n$ is called **primitive**, if there exists a positive integer $k$ such that every entry of $\mathbf{A}^k$ is positive (denoted $\mathbf{A}^k > 0$).
- The smallest $k$ for which $\mathbf{A}^k > 0$ is called the **primitive exponent**, denoted by $\gamma(\mathbf{A})$.
- A digraph $D$ is called primitive, if its adjacency matrix $\mathbf{A}$ is primitive. $\gamma(\mathbf{A})$ is called the **primitive exponent** of $D$, denoted by $\gamma(D)$.

**Theorem 1.10.1** (D. Rosenblatt, 1957) **a digraph $D$ is primitive $\Leftrightarrow$ $D$ is strongly connected, and the gcd of all distinct directed cycle lengths $\ell_1, \ell_2, \cdots, \ell_c$ in $D$ is 1, i.e., g.c.d$(\ell_1, \ell_2, \cdots, \ell_c) = 1$.**
Proof: Omitted.

Remark: From Theorem 1.10.1, the direct cycle $C_3$ is primitive, but $C_4$ is not.

From Theorem 1.9 and Corollary 1.5.1, we immediately have

**Corollary 1.9** **Let $n$ ($\geqslant 5$) strongly connected tournament $T_n$, its adjacency matrix $\mathbf{A}$ satisfies $\mathbf{A}^{d+3} > 0$, where $d = d(T_n)$.**

The adjacency matrix and incidence matrix of a graph play very important roles in analyzing graph properties and structure. We give an example below.

**Theorem 13.1 (The Perron-Frobenius Theorem)** *Suppose $A$ is a real nonnegative $n \times n$ matrix whose underlying directed graph $D$ is strongly connected. Then:*

*(a) $\rho(A)$ is a simple eigenvalue of $A$. If $x$ is an eigenvector for $\rho$, then no entries of $x$ are zeros, and all have the same sign.*

*(b) Suppose $A_1$ is a real nonnegative $n \times n$ matrix such that $A - A_1$ is nonnegative. Then $\rho(A_1) \leq \rho(A)$, with equality if and only if $A_1 = A$.*

*(c) If $\theta$ is an eigenvalue of $A$ and $|\theta| = \rho(A)$, then $\theta/\rho(A)$ is an $m$th root of unity and $e^{2\pi ir/m}\rho(A)$ is an eigenvalue of $A$ for all $r$. Further, all cycles of $D$ have length divisible by $m$.*

**Example 1.9.4** (J. Plesnik & Š. Znám (1974); W. G. Bridges & S. Touge, (1980)) For given $\Delta \geq 2$ and $d \geq 2$, no $(\Delta, d)$-Moore digraph exists. Note: $\Delta = \max\{\Delta^+, \Delta^-\}$.

**Proof**

- Let $D$ be a $(\Delta, d)$-Moore digraph. Then its order is

$$n = 1 + \Delta + \Delta^2 + \cdots + \Delta^d$$

as in equation (4).

- By the properties of a $(\Delta, d)$-Moore digraph (Exercise 1.8.5): (i) $D$ is a $\Delta$-regular simple digraph; (ii) it has no directed cycle of length $\leq d$; (iii) for any $x, y \in V(D)$, there is a unique $(x, y)$-path of length $\leq d$ in $D$.

- Let $\mathbf{A} = \mathbf{A}(D)$ be the adjacency matrix of $D$. Then $\Delta$ is a simple eigenvalue of $\mathbf{A}$ (Perron–Frobenius theorem, Exercise 1.9.8(d)).

- From (iii) and Theorem 1.9, it is easy to obtain,

$$\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^d = \mathbf{J}. \tag{12}$$

  Equation (12) means $\mathbf{J}$ is a polynomial in $\mathbf{A}$. Therefore, $\mathbf{A}$ and $\mathbf{J}$ have the same eigenvectors.

- Let $r$ be an eigenvalue of $\mathbf{A}$ distinct from $\Delta$, and $X$ an eigenvector of $\mathbf{A}$ for eigenvalue $r$ (i.e., $X \neq \mathbf{1}$). Since 0 is an $(n-1)$-fold eigenvalue of $\mathbf{J}$,

$$\mathbf{A}X = rX, \qquad \mathbf{J}X = 0.$$

Equation (12) yields the relation

$$1+r+r^2+\cdots+r^d=0. \tag{13}$$

Equation (13) shows $r$ is a $(d+1)$-th root of unity, i.e., $r^{d+1}=1$.

- Let $r_1, r_2, \cdots, r_{n-1}$ be the $n-1$ eigenvalues of $\mathbf{A}$ distinct from $\Delta$. By (ii) and Theorem 1.9, all diagonal entries of $\mathbf{A}^i$ $(1 \le i \le d)$ are 0. Therefore, their sum

$$\mathrm{Tr}\mathbf{A}^i = 0, \quad i = 1, 2, \cdots, d.$$

This means the sum of all eigenvalues of $\mathbf{A}^i$

$$\Delta^i + \sum_{j=1}^{n-1} r_j^i = 0, \quad i = 1, 2 \cdots, d. \tag{14}$$

- Since $r_j \overline{r}_j = |r_j|^2 = 1 = r_j^{d+1}$, $r_j^{-1} = \overline{r}_j = r_j^d$, where $\overline{r}_j$ is the complex conjugate of $r_j$.

- Taking $i = 1$ and $i = d$ in equation (14),

$$-\Delta = \sum_{j=1}^{n-1} r_j, \qquad -\Delta^d = \sum_{j=1}^{n-1} r_j^d.$$

Taking complex conjugates in the first equation and noting $\overline{r}_j = r_j^{-1} = r_j^d$, gives

$$-\Delta = \sum_{j=1}^{n-1} r_j^{-1} = \sum_{j=1}^{n-1} r_j^d = -\Delta^d.$$

This holds $\Leftrightarrow$ either $d = 1$ or $\Delta = 1$, contradicting $d, \Delta \geq 2$. The theorem is proved. ∎

**Remark:** Let $\overrightarrow{n}(\Delta, d)$ denote the maximum order of a digraph with maximum degree $\Delta$ and diameter $d$. Noting that the Kautz digraph

$K(\Delta, d)$ has $\Delta^d + \Delta^{d-1}$ vertices. Thus, for $\Delta \geq 2$ and $d \geq 2$ we have the following inequality:

$$\Delta^{d-1} + \Delta^d \leqslant \overrightarrow{n}(\Delta, d) \leqslant \Delta + \Delta^2 + \cdots + \Delta^{d-1} + \Delta^d. \qquad (15)$$

From equation (15), $\overrightarrow{n}(\Delta, 2) = \Delta + \Delta^2$; the digraph attaining this bound is $K(\Delta, 2)$.

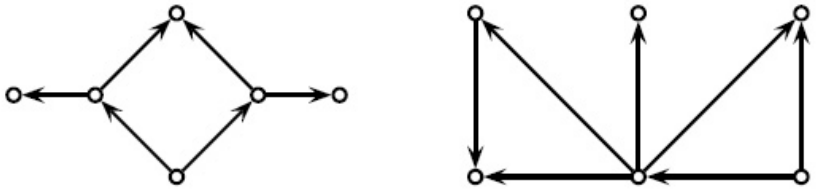
## Exercises

1.9.8 Let $\mathbf{A}$ be the adjacency matrix of a (directed or undirected) graph $D$, and $\mathbf{I}$ the $n \times n$ identity matrix. The polynomial

$$P_D(\lambda) = \det(\lambda \mathbf{I} - \mathbf{A}) = \lambda^n + c_1 \lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n$$

is called the characteristic polynomial of $D$; $P_D(\lambda)$ are called the characteristic roots of $D$; the characteristic roots together with their multiplicities form the spectrum.

(a) Find the characteristic polynomials (both are $\lambda^6$) and characteristic roots of the following two graphs.



Note: graphs with the same spectrum as in the figure above are called cospectral graphs. Clearly, a graph uniquely determines its spectrum; the converse is one of the difficult open problems in graph theory.