THE AMERICAN
UNIVERSITY IN CAIRO

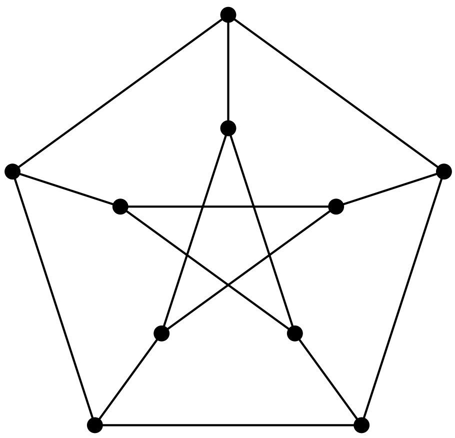
AUC

DEPARTMENT OF MATHEMATICS

SPRING TERM 2026

Graph Theory

Lecture Notes



Lecturer:
Isabel Müller

Term:
Spring 2026

"An die Professorin, der ich meine Wertschätzung nicht zeigen konnte,
und an die Professorin, der ich es niemals vergelten kann."

Contents

Chapter 1

Graphs

1. The Basics

1.1 Recall

1. A **set** is merely an accumulation of objects. These objects are called **elements** of the set. If an object x is an element of S, we write $x \in S$. The set of all elements with a certain property P is denoted via $\{x \mid x \text{ has property } P\}$.

2. An n**-ary relation** R on a set A is a subset of the power set of A^n, i.e., $R \subseteq \mathcal{P}(A^n)$. If $n = 2$, we call the relation **binary**.

   A binary relation R on a set A is called:

   (i) **symmetric** if $R(a, b)$ implies $R(b, a)$ for all $a, b \in A$.

   (ii) **asymmetric** if $R(a, b)$ implies $\neg R(b, a)$ for all $a, b \in A$.

   (iii) **antisymmetric** if $R(a, b) \wedge R(b, a)$ implies $a = b$ for all $a, b \in A$.

   (iv) **reflexive** if $R(a, a)$ for all $a \in A$.

   (v) **irreflexive** if $\neg R(a, a)$ for all $a \in A$.

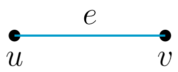
   (vi) **transitive** if $R(a, b) \wedge R(b, c)$ implies $R(a, c)$ for all $a, b, c \in A$.

Definition 1.2

A **graph** $G = (V, E)$ is a pair of sets V and E such that E consists of subsets of V of size two. V is called the set of **vertices** and E the set of **edges**. A graph G is called **finite** if V is a finite set. The **order** $|G|$ of a graph $G = (V, E)$ is the cardinality of its vertex set, so $|G| = |V|$. The **size** $\|G\|$ of G is the cardinality of its edge set, $\|G\| = |E|$.

1.3 Visualisation

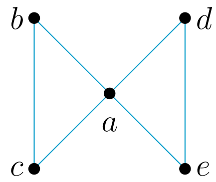
Let $G = (V, E)$ be a graph. We visualise vertices $u, v \dots \in V$ by dots and edges $e = \{u, v\} \in E$ by the diagram:



Example 1.4. *Bowtie Graph* Let $G = (V, E)$ be the graph with $V = \{a, b, c, d, e\}$ and

$$E = \{\{a, b\}, \{a, c\}, \{a, d\}, \{a, e\}, \{b, c\}, \{d, e\}\}.$$

The graph G has order 5 and size 6. It can be visualized via:



This visualisation motivates its name: **bowtie graph**.

1.5 Notation

1. For a graph $G = (V, E)$ we may denote its vertex set by $V(G)$ or V_G for clarity.
2. Similarly, we often denote E by $E(G)$ or E_G.
3. We denote an edge $\{u, v\}$ simply by uv.
4. Edges are often called $e, e_1, e_2, f \dots$, while vertices are called $u, v, x, y, \dots$.

Definition 1.6

Let $G = (V, E)$ be a graph.

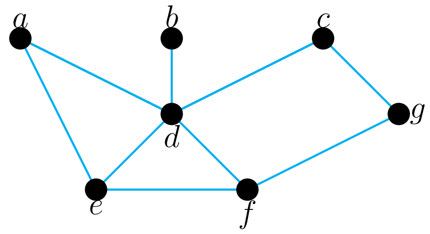
1. If $uv \in E$ is an edge, then we say that u and v are **adjacent** or **neighbours**. If $uv \notin E$, we call u and v **nonadjacent**.
2. If $e = uv \in E$, we say that u and v are the **end vertices** of e or that they are **incident** with e.
3. The **neighborhood** $N(v)$ of a vertex $v \in V$ is the set of all vertices adjacent to v, i.e., $N(v) = \{u \in V \mid uv \in E\}$. The **closed neighborhood** $N[v]$ of v is $N[v] := N(v) \cup \{v\}$.
4. The **neighborhood** $N(S)$ of a set of vertices is defined as $N(S) := \bigcup_{v \in S} N(v)$. Similarly, the **closed neighborhood** $N[S]$ is set to be $N[S] := N(S) \cup S (= \bigcup_{v \in S} N[v])$.
5. The **degree** $\deg(v)$ of $v \in V$ is the number of edges incident with v, i.e., $\deg(v) := |\{e \in E \mid v \in e\}| = |N(v)|$.
6. The **maximum degree** $\Delta(G)$ of G is defined as

$$\Delta(G) := \max\{\deg(v) \mid v \in V\}.$$

Similarly, $\delta(G) := \min\{\deg(v) \mid v \in V\}$ is the **minimum degree** of G.

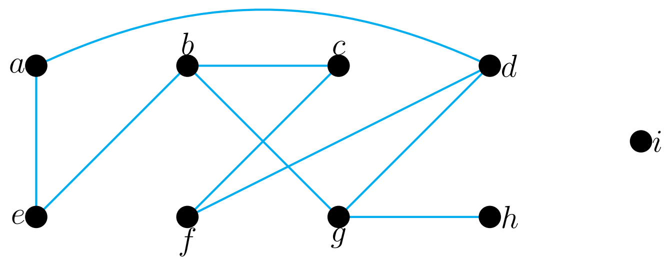
7. The **degree sequence** of a graph G is the sequence containing all degrees of the vertices of G (with repetition) in decreasing order.

Example 1.7. Consider G given by:



Then $\Delta(G) = 5$, $\delta(G) = 1$. $N(e) = \{a, d, f\}$, $N[b] = \{b, d\}$. $N[a, g] = \{a, c, d, e, f, g\}$
Order of G, size of G is 9. Degree sequence $(5, 3, 3, 2, 2, 2, 1)$.

Example 1.8. Consider G given via the diagram:



Then $V(G) = \{a, b, c, d, e, f, g, h, i\}$
$E(G) = \{\{a, d\}, \{a, e\}, \{b, c\}, \{b, e\}, \{b, g\}, \{c, f\}, \{d, f\}, \{d, g\}, \{g, h\}\}$

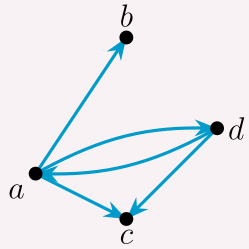
Order $|G| = 9$, size of G is 9, degree sequence $(3, 3, 3, 2, 2, 2, 2, 1, 0)$.
$N(f) = \{c, d\}$, $N[d, e] = \{a, b, c, d, e, f\}$, $\Delta(G) = 3$, $\delta(G) = 0$.

Remark 1.9. A graph can be considered as a set V together with a binary relation E on V which is symmetric and irreflexive.

Definition 1.10 Variants of Graphs

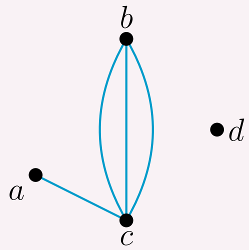
1. If $G = (V, E)$ and we replace E with a set of ordered pairs, then we call G a **directed graph** or **digraph**.

   Ex: $V(G) = \{a, b, c, d\}$, $E(G) = \{(a, b), (a, c), (a, d), (d, a), (d, c)\}$



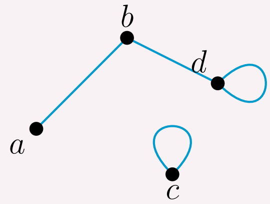
2. If $G = (V, E)$ and we replace E by a multiset (iterations of the same elements are distinguished), then we call G a **multigraph**.

   Ex: $E = [\{a, c\}, \{b, c\}, \{b, c\}, \{b, c\}]$



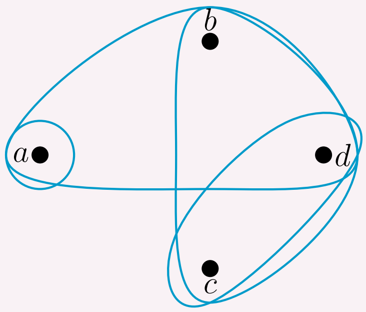
3. If $G = (V, E)$ and we extend E by allowing loops, we call G a **pseudograph**.

   Ex: $E = \{\{a, b\}, \{b, d\}, \{c, c\}, \{d, d\}\}$



4. If we allow edges to be arbitrary sets of vertices instead of 2-elementary ones, we call G a **hypergraph**.

   Ex: $E = \{\{a\}, \{a, b, d\}, \{b, c, d\}, \{c, d\}\}$



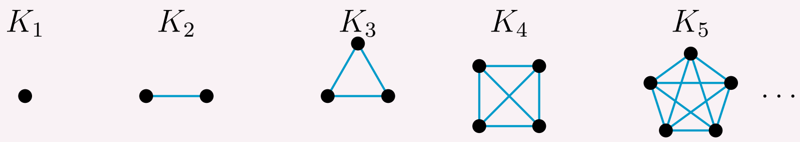
1.11 Setting

In this lecture, unless otherwise stated, by a graph we mean a finite, simple graph with $|V| \geq 1$.

Definition 1.12

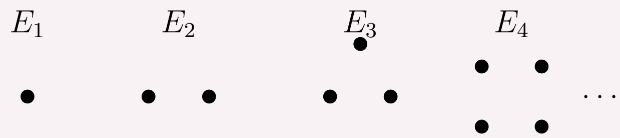
- The **complete graph** K_n for $n \geq 1$ is the graph consisting of n vertices such that any two vertices are adjacent.

  e.g.



- The **empty graph** E_n is the graph consisting of n vertices and no edges.

  e.g.



Theorem 1.13 The Handshaking Lemma

If $G = (V, E)$ is a graph, then

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

Proof. We proceed by induction on $n := |E|$.

<u>**n=0:**</u> If $|E| = 0$, then $\deg(v) = 0$ for any $v \in V$, whence clearly

$$0 = \sum_{v \in V} \deg(v) = 2|E| = 0.$$

<u>**n $\to$ n+1:**</u> Assume $(*)$ holds for any $G' = (V', E')$ with $|E'| = n$ (I.H.) and consider $G = (V, E)$ with $|E| = n + 1 (\geq 1)$ arbitrary. Let $e \in E$ arbitrary and consider $G' = (V, E \setminus \{e\})$. Then, if $e = uv$, we get $|E(G)| = |E(G')| + 1$ and

$$\deg_G(u) = \deg_{G'}(u) + 1 \quad \text{and} \quad \deg_G(v) = \deg_{G'}(v) + 1, \text{ whence}$$

$$\begin{aligned}
2|E(G)| &= 2|E(G')| + 2 \\
&\overset{\text{I.H.}}{=} \sum_{w \in V} \deg_{G'}(w) + 2 \\
&= \sum_{w \in V \setminus \{u,v\}} \deg_{G'}(w) + \deg_{G'}(u) + 1 + \deg_{G'}(v) + 1 \\
&= \sum_{w \in V} \deg_G(w), \quad \text{as desired.}
\end{aligned}$$

$\square$

Corollary 1.14

Any graph G has an even number of vertices of odd degree.

Proof. Exercise. $\square$

Corollary 1.15

For any graph $G = (V, E)$ we have

$$\delta(G) \leq 2\frac{|E|}{|V|} \leq \Delta(G).$$

Proof.

$$|V| \cdot \delta(G) = \sum_{v \in V} \delta(G) \leq \sum_{v \in V} \deg(v) \leq \sum_{v \in V} \Delta(G) = |V|\Delta(G)$$

Using Theorem 1.13, $\sum \deg(v) = 2|E|$. Dividing by $|V|$ yields the result. $\square$

Lemma 1.16

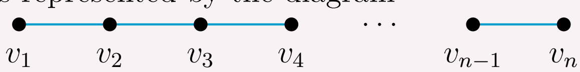
If $|G| \geq 2$, then G contains at least two vertices of the same degree.

Proof. If G has two vertices of degree 0, then we are done. Otherwise, we may assume that G has none. If $|G| = n$, and $v \in V$, then $1 \leq \deg(v) \leq n - 1$. Note that this leaves us with $n - 1$ choices of degrees for n many different vertices. Hence, at least two vertices must have the same degree. $\square$

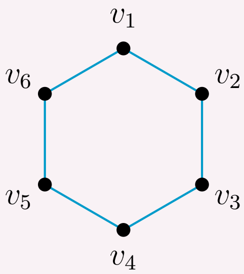
Remark 1.17. The above line of thought is called the **pigeon hole principle**. If there are n many pigeons wanting to fit into $n - 1$ many holes, then at least two of them have to cuddle up in the same hole.

Definition 1.18

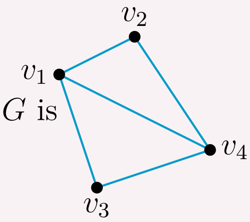
1) The **path** P_n is the graph on n vertices $v_1, \ldots, v_n$ with the edge set $E(P_n) = \{v_i v_{i+1} \mid 1 \leq i < n\}$, i.e. P_n is represented by the diagram

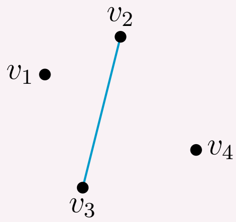


2) The **cycle** C_n is the graph on n vertices with edge set $E(C_n) = \{v_i v_{i+1} \mid 1 \leq i < n\} \cup \{v_n v_1\}$. E.g.

C_6 is 

3) Let $G = (V, E)$ be an arbitrary graph. The **complement** $\overline{G}$ of G is the graph $\overline{G} = (V, \overline{E})$, where $\overline{E} = \{uv \mid u, v \in V, uv \notin E\}$.

e.g. If G is 

, then $\overline{G}$ is 

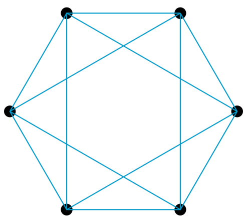
Definition 1.19

We call a graph G **regular** if any of its vertices has the same degree. If this degree is r, we say that G is r**-regular**.

Remark 1.20.

1. A graph G is regular iff $\delta(G) = \Delta(G)$.
2. K_n is $(n-1)$-regular and E_n is 0-regular.
3. An r-regular graph of order n has $\frac{1}{2}nr$ many edges.

Example 1.21. The graph below is 4-regular of order 6.



2. Subgraphs

There are two ways in which one graph can be part of another graph.

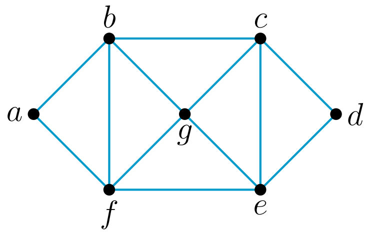
Definition 1.22

1. A graph H is called a **subgraph** of some graph G, written $H \subseteq G$, if $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. We also say G **contains** H.
2. If $H \subseteq G$, we say that H is an **induced subgraph** of G, written $H \sqsubseteq G$, if $E(H) = \{uv \in E(G) \mid u, v \in V(H)\}$.

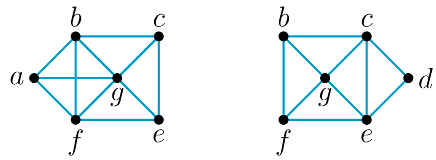
Remark 1.23.

1. $H \subseteq G$ is induced if for any two vertices in H we have: If they are adjacent in G, then they are adjacent in H.
2. Every induced subgraph is a subgraph but not vice versa.
3. If G is a graph and $S \subseteq V(G)$, then there is only one induced subgraph $H \sqsubseteq G$ with vertex set S, i.e. $V(H) = S$. We denote this graph by $\langle S \rangle$ and call it the subgraph of G induced by S.

Example 1.24. Consider G given as



Then



subgraph	✓	✓	✗	✓	✓
induced	✗	✓	✗	✓	✗

3. Walks in Graphs

Definition 1.25

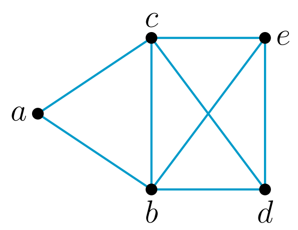
A (v_0, v_k)-**walk** in a graph is a sequence of vertices $(v_0, v_1, \dots, v_k)$ s.t. any two consecutive vertices v_i and v_{i+1} are adjacent. We call the edges $\{v_0v_1, v_1v_2, \dots, v_{k-1}v_k\}$ the **edges of the walk**. We say that the walk is **closed** if $v_0 = v_k$. The **length** of a walk is the number of edges in it (counting repetition).

Definition 1.26

We distinguish the following types of walks:

- A **trail** is a walk whose edges are pairwise distinct.
- A **circuit** is a closed walk whose edges are pairwise distinct.
- A **path** is a walk whose vertices are distinct.
- A **cycle** is a closed walk $(v_0, \dots, v_k = v_0)$ with $k \geq 3$ and whose vertices $v_0, \dots, v_{k-1}$ are pairwise distinct.

Example 1.27. Consider G via



Give examples for a:

- **walk** (d, b, c, d, b, a)
  da-walk, length 5
- **trail** (d, c, a, b, c, e)
  de-trail, length 5
- **path** (d, c, a, b, e)
  de-path, length 4
- **closed walk** (e, b, c, a, b, d, e)
  e-closed walk, length 6
- **circuit** (d, c, a, b, c, e, d)
  d-circuit, length 6
- **cycle** (d, c, a, b, e, d)
  d-cycle, length 5

Lemma 1.28

If $\delta(G) \geq 2$, then G contains a cycle as a subgraph.

Proof. Let $P = (v_0, \dots, v_k)$ be a path in G of maximal length. This exists, as G is finite. Further, as $\delta(G) \geq 2$, we get $k \geq 2$. As $\deg(v_0) \geq \delta(G) \geq 2$, v_0 has at least two neighbors. One of them is v_1. Let us denote the other one by u. If $u \neq v_i$ for all $1 \leq i \leq k$, then $\tilde{P} = (u, v_0, v_1, \dots, v_k)$ is still a path and of greater length than P, contradicting our assumptions. Hence, $u = v_i$ for some $1 \leq i \leq k$. But then the sequence $(v_0, v_1, \dots, v_i = u, v_0)$ is the desired cycle subgraph of G. $\square$

Corollary 1.29 Contrapositive

If G does not contain any cycles, then $\delta(G) \leq 1$.

Theorem 1.30

Every uv-walk in a graph contains a uv-path.

Proof. We proceed by strong induction on the length $n \geq 1$ of the walk. **I.B. n=1.** If the uv-walk is of length one, then it is exactly (u, v), which is also a path. **I.S.** Assume every uv-walk of length at most $n \geq 1$ contains a uv-path (I.H.). Assume there is a uv-walk $W = (u = w_0, w_1, \ldots, w_n, w_{n+1} = v)$ of length $n + 1$. If W is already a path, we are done. Otherwise there are i, j s.t. $0 \leq i < j \leq n + 1$ and $w_i = w_j$. But then the walk $\tilde{W}$ which arises from W by deleting the vertices $w_{i+1}, \ldots, w_{j-1}, w_j$, i.e. $\tilde{W} = (u = w_0, \ldots, w_i, w_{j+1}, \ldots, w_{n+1} = v)$ is still a uv-walk, but of length at most n. Using I.H., we know that $\tilde{W}$ contains a uv-path, whence also W contains (the same) uv-path. $\square$

4. Connectivity

Definition 1.31

A graph is **connected** if there exists an uv-path in G for any vertices $u, v \in V(G)$. Otherwise, it is called **disconnected**.

Intuition

A graph is connected if you could pick it up entirely by just lifting one vertex. If it is not connected, then the subgraph you lift that way is called a connected component.

Definition 1.32

A **connected component** of G is a maximal connected induced subgraph of G. i.e. $C \sqsubseteq G$ is a connected component iff (i) C is connected and (ii) for any $v \in V(G) \setminus V(C)$ the induced subgraph on $V(C) \cup \{v\}$ is **not** connected.

Remark 1.33. G is connected iff it has exactly one connected component.

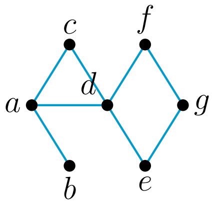
Even among connected graphs, there are different levels of being connected. E.g. the graph K_5  "feels" more connected than the graph . In order to properly describe this intuition, we need more notations.

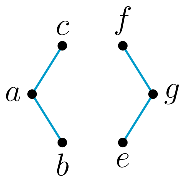
Definition 1.34 Vertex and Edge Deletion

Let G be a graph, $S \subseteq V_G$ and $T \subseteq E_G$.

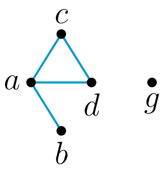
1. By $G - S$ we denote the graph arising from G by removing from V_G all vertices in S and their incident edges.
2. If $S = \{v\}$, we write $G - v$.
3. By $G - T$ we denote the graph arising from G by removing only the edges in T, but no vertices.
4. If $T = \{e\}$, we write $G - e$.

Example 1.35. Consider G as given below. Note that G only has one connected component.

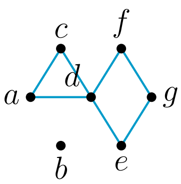
$G =$ 

Then $G - d$ is  and has 2 connected components.

The vertex d is called a **cut vertex**.

Further, $G - \{e, f\}$ is  . It also has 2 connected components.

The set $\{e, f\}$ is called a **cut set**.

Further, $G - ab$ is  Again, it has 2 connected components.

We call the edge ab a **bridge**.

Definition 1.36

Let G be a graph.

1. We call $v \in V_G$ a **cut vertex** if $G - v$ has more connected components than G itself.
2. We call $e \in E_G$ a **bridge** if $G - e$ has more connected components than G itself.
3. We call $S \subseteq V_G$ a **cut set** if $G - S$ is disconnected.
4. A connected graph which does not contain any cut vertices is called **non-separable**.

1.37 Observation

1. If G is connected then v is a cut vertex of G iff $\{v\}$ is a cut set.
2. The vertex v is a cut vertex iff there are vertices u and w, different from v s.t. every uw-path uses v.
3. A graph has no cut sets iff it is a complete graph.

Definition 1.38

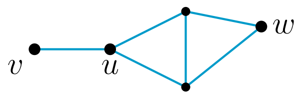
For a non-complete graph G, we define its **connectivity** $\kappa(G)$ as the minimal size of a cut set. For K_n, we set $\kappa(K_n) = n - 1$.

Lemma 1.39

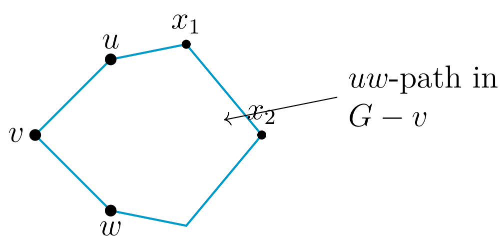
If G is a nonseparable graph of order at least 3, then $\delta(G) \geq 2$ and every vertex of G is contained in a cycle.

Proof. Consider G nonseparable with $|G| \geq 3$. By definition, G is connected, i.e. $\delta(G) \geq 1$.

First we show that $\delta(G) \geq 2$. Otherwise, we have $\delta(G) = 1$, i.e. there is some vertex v s.t. $\deg(v) = 1$. Let u be the unique neighbor of v and w any other vertex of G (which exists as $|G| \geq 3$). Then clearly any vw-path must use the unique neighbor u of v, whence u is a cut vertex. This contradicts the fact that G is inseparable. Hence, $\delta(G) \geq 2$, as desired.

Picture: 

Now, consider $v \in V_G$ arbitrary. We want to show that v is contained in a cycle in G. As $\delta(G) \geq 2$, v has at least 2 neighbors, say u and w. As G is nonseparable, $G - v$ is still connected. In particular, there is a uw-path $(u = x_0, x_1, x_2, \ldots, x_k = w)$ in $G - v$. But then the walk $(x_0 = u, x_1, \ldots, x_k = w, v, x_0 = u)$ is the desired cycle containing v.

Picture: 

∎

Definition 1.40

We say that G is **k-connected** if $\kappa(G) \geq k$, i.e. if G is connected and $G - S$ is still connected for any $S \subseteq V_G$ with $|S| < k$.

Lemma 1.41

The following hold:
1) G is connected iff $\kappa(G) \geq 1$.
2) G is 1-connected iff G is connected.
3) G is 2-connected iff G is connected and has no cut vertices.
4) G is 2-connected iff G is non-separable.
5) If G is 2-connected, then it contains at least one cycle (for $|G| \geq 3$).
6) If G is k-connected, then G is j-connected for all $j \leq k$.
7) $|G| > \kappa(G)$.
8) $\kappa(G) \leq \delta(G)$.

Proof. 1)–6) are easy observations – verify them by yourselves.

7) If $G = K_n$, then $|G| = n > n-1 = \kappa(G)$. Otherwise, assume $\kappa(G) = k$, i.e. ex. $S \subseteq V_G$ s.t. $|S| = k$ and $G - S$ is disconnected. For $G - S$ to be disconnected, it must contain at least 2 vertices, whence

$$|G| \geq |S| + 2 = \kappa(G) + 2 > \kappa(G).$$

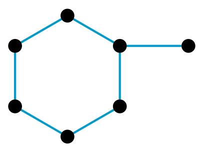
8) Assume $\kappa(G) > \delta(G)$ and let $v \in V_G$ s.t. $\deg(v) = \delta(G)$. Note that $|G| > \kappa(G) > \delta(G) = |N(v)|$, whence $G - N(v)$ contains at least one vertex besides v. But clearly, $G - N(v)$ is disconnected (as $\deg^{G-N(v)}(v) = 0$). Hence, $N(v)$ is a cut set and $\kappa(G) \leq |N(v)| = \delta(G)$, contradicting the assumptions. □

5. Bipartite Graphs

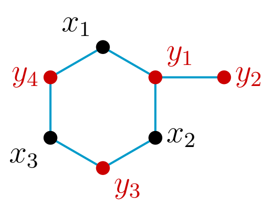
Definition 1.42

A graph G is called **bipartite** if we can partition the vertex set V_G into two disjoint sets $V_G = X \cup Y$ s.t. every edge of G has one end vertex in X and the other in Y.

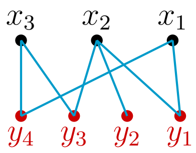
Example 1.43. Consider $G :=$



We can partition the vertices of G into two sets via $X = \{x_1, x_2, x_3\}$ and $Y = \{y_1, y_2, y_3, y_4\}$.



Rearranging the position of the vertices makes it clear that G is bipartite:



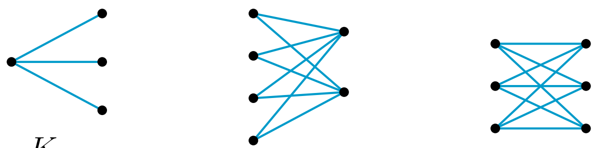
We see that there are no edges between any two vertices in X or in Y.

Remark 1.44. A graph G is bipartite if and only if we can color the vertices of G with two colors s.t. the end vertices of each edge have different colors.

Definition 1.45

Let $m, n \in \mathbb{Z}_+$. The **complete bipartite graph** $K_{m,n}$ is the bipartite graph with $X = \{x_1, \dots, x_m\}$, $Y = \{y_1, \dots, y_n\}$, $V_G = X \cup Y$ and $E_G = \{xy \mid x \in X, y \in Y\}$.

Example 1.46. Below are some examples of complete bipartite graphs.



$K_{1,3}$ $K_{4,2}$ $K_{3,3}$

The following theorem helps us decide whether or not a given graph is bipartite.

Theorem 1.47

A graph is bipartite iff it does not contain odd cycles.

Proof. "⇒": Assume G is bipartite and nevertheless there is a cycle of odd length, say $(x_0, x_1, \dots, x_{2k}, x_{2k+1} = x_0)$. By Remark 1.44, we can color V_G in two colors, $C1$ and $C2$, s.t. adjacent vertices have different colors. Then, if x_0 has color $C1$, x_1 has color $C2$ whence x_2 has color $C1$. That way we see that the color of x_i is

$$\begin{cases} C1 \text{ if } i \text{ is even} \\ C2 \text{ if } i \text{ is odd} \end{cases}.$$

Following that logic, the vertex $x_0 = x_{2k+1}$ should have color $C1$ and color $C2$ at the same time, which is a contradiction.

"⇐": Now consider that G does not contain odd cycles. We will show that G is bipartite by providing a partition. We may assume that G is connected as otherwise we work component per component. Pick $v \in V_G$ arbitrary and define

$$X = \{w \in V_G \mid \text{the shortest } vw \text{ path has even length}\} \text{ and}$$

$$Y = \{w \in V_G \mid \text{the shortest } vw \text{ path has odd length}\}.$$

Clearly, X and Y are disjoint. We will show that there are no adjacent vertices in X or Y respectively. Note that $v \in X$.

Aiming for a contradiction, assume that there are vertices $w_1, w_2 \in X$ which are adjacent. Clearly, $w_1 \neq v$, as otherwise the shortest vw_2-path was exactly vw_2 of length 1. Similarly, $w_2 \neq v$. Let $P_1 = (v = x_0, x_1, \ldots, x_{2k} = w_1)$ and $P_2 = (v = y_0, y_1, \ldots, y_{2\ell} = w_2)$ be the shortest vw_1- and vw_2-paths. Suppose that $x_i = y_j$ for some $0 < i \leq 2k$ and $0 < j \leq 2\ell$. If $i < j$, then $(v = x_0, x_1, \ldots, x_i, y_{j+1}, \ldots, y_{2\ell} = w_2)$ is a vw_2 path shorter than P_2, a contradiction. Similarly, $j < i$ is impossible, whence $i = j$, whenever $x_i = y_j$.

Now, pick the largest i s.t. $x_i = y_i$. As $x_0 = v = y_0$, such an i always exists. Then we obtain the following cycle

$$C = (\underbrace{x_i, x_{i+1}, \ldots, x_{2k}}_{2k-i} = w_1, \underbrace{w_2}_{1} = y_{2\ell}, \underbrace{y_{2\ell-1}, \ldots, y_{i+1}}_{2\ell-i}, y_i = x_i).$$

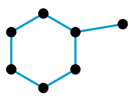
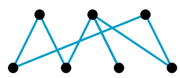
This is a cycle, as P_1 and P_2 were paths and i was maximal s.t. $x_i = y_i$. Further, the length of C is odd, as it equals

$$(2k - i) + 1 + (2\ell - i) = 2(k + \ell - i) + 1.$$

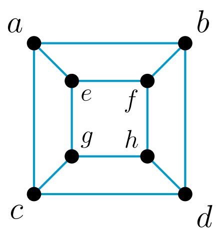
This contradicts our assumption that G does not contain odd cycles. We hence proved that no two vertices w_1 and w_2 from X can be adjacent. The arguments for $v_1, v_2 \in Y$ is analogous (try to write it down). This concludes the proof. $\square$

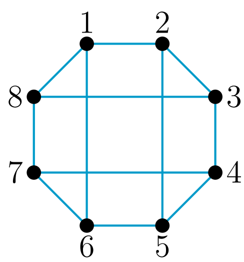
6. Graph Isomorphisms

In Example 1.43, we rearranged the given graph G_1 as G_2.

$G_1 =$  as $G_2 =$ 

We understand G_1 and G_2 as the same, even though on the first glance, they look very similar. Another example is given by

$G =$ 

$H =$ 

We can relabel the vertices of G via $a \mapsto 1, b \mapsto 2, c \mapsto 8, d \mapsto 3, e \mapsto 7, f \mapsto 4, g \mapsto 6$ and $h \mapsto 5$ and obtain H. The aim of this section is to formalise this concept.

Definition 1.48

We say that a graph G is **isomorphic** to a graph H if there exists a bijection $\varphi : V_G \to V_H$ s.t. for any $u, v \in V_G$ we have that $\{u, v\} \in E_G$ if and only if $\{\varphi(u), \varphi(v)\} \in E_H$. Then, the map φ is called an **isomorphism** and we write $G \cong H$.

Remark 1.49. Let $G \cong H$ via $\varphi : V_G \to V_H$. Then:

1) $|V_G| = |V_H|$ and $|E_G| = |E_H|$ and $\overline{G} \cong \overline{H}$.

2) The degree sequence of G equals the degree sequence of H.

3) G is connected iff H is connected.

4) $\deg_G(v) = \deg_H(\varphi(v))$ for all $v \in V_G$.

Chapter 2

Distance in Graphs

1. Introduction

We have a natural understanding of the "distance" between two objects in our physical space. But there are many other ways of defining distances. E.g., the distance between people could be the positive difference of their birth years or the number of acquaintances you need to connect one to the other.

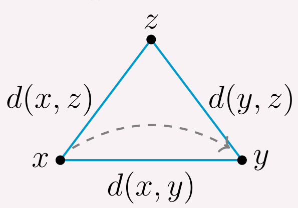
In this chapter we will introduce a notion of distance of vertices in a graph. But first let us note what are the characterising properties that make us call all these concepts "distances".

Definition 2.1

Let X be any set. We call a function $d : X \times X \to \mathbb{R}^{\geq 0} \cup \{\infty\}$ a **metric** if it satisfies for all $x, y, z \in X$:

1) $d(x, y) \geq 0$
2) $d(x, y) = 0$ iff $x = y$
3) $d(x, y) = d(y, x)$
4) $d(x, z) \leq d(x, y) + d(y, z)$ (**Triangle Inequality**)

We then call the pair (X, d) a **metric space**.



Example 2.2. Consider $X = \mathbb{R}$ and $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^{\geq 0}$ via $d(x, y) := |x - y|$. Then $(\mathbb{R}, d)$ is a metric space.

Now we are ready to define a metric on an arbitrary graph.

Definition 2.3

Let G be any graph and $u, v \in V_G$. We define the **distance** $d(u, v)$ between u and v as the length of the shortest uv-path in G, i.e.

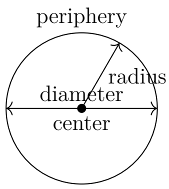
$$d(u, v) := \min\{\text{length}(P) \mid P \text{ is a } uv\text{-path}\}.$$

If there is no such path, we set $d(u, v) := \infty$.
2) If $d(u, v) = k$, then any uv-path of length k is called a **geodesic**.

Remark 2.4.

1) We may write $d_G(u, v)$ to emphasize that we consider the distance in G.
2) While in $(\mathbb{R}, d)$ geodesics are unique, in general this is not the case. Consider for example two opposite poles on a sphere.
3) $d(x, y) = \infty$ iff x and y are in different connected components.
4) (V_G, d) is a metric space for any connected graph G.

We call something eccentric if it is away from the usual. Similarly, in graphs we measure by eccentricity how far a vertex is from the center. Consider the following notions on a cycle:



Definition 2.5

1) The **eccentricity** $ecc(v)$ of a vertex v is its greatest distance to any other vertex, i.e. $ecc(v) = \max\{d(u, v) \mid u \in V_G\}$.
2) The **radius** $rad(G)$ is the smallest possible eccentricity and the **diameter** $diam(G)$ is the largest possible eccentricity.
3) The **center** $C(G)$ is the set $\{v \in V_G \mid ecc(v) = rad(G)\}$ and the **periphery** $P(G)$ is the set $\{v \in V_G \mid ecc(v) = diam(G)\}$.

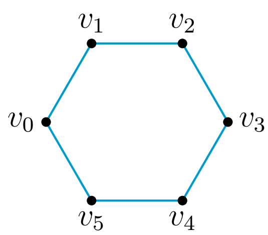
Example 2.6. 1) Consider P_5, the path of length 4, i.e.

Then

$$\begin{aligned}
d(v_1, v_i) &= i - 1, \text{ whence } ecc(v_1) = \max\{0, 1, 2, 3, 4\} = 4.\\
d(v_2, v_i) &= |i - 2|, \text{ whence } ecc(v_2) = \max\{1, 0, 1, 2, 3\} = 3.\\
d(v_3, v_i) &= |i - 3|, \text{ whence } ecc(v_3) = \max\{2, 1, 0, 1, 2\} = 2.\\
d(v_4, v_i) &= |i - 4|, \text{ whence } ecc(v_4) = \max\{3, 2, 1, 0, 1\} = 3.\\
d(v_5, v_i) &= |i - 5|, \text{ whence } ecc(v_5) = \max\{4, 3, 2, 1, 0\} = 4.
\end{aligned}$$

Hence $rad(P_5) = \min\{ecc(v) \mid v \in V\} = \min\{4, 3, 2, 3, 4\} = 2$.
Also $C(P_5) = \{v \in V \mid ecc(v) = rad(P_5)\} = \{v_3\}$.
Further $diam(P_5) = \max\{ecc(v) \mid v \in V\} = \max\{4, 3, 2, 3, 4\} = 4$.
And $P(P_5) = \{v \in V \mid ecc(v) = diam(P_5)\} = \{v_1, v_5\}$.

2) Consider $G := C_6$, the cycle of length 6, i.e. $G =$



Then

$$\begin{aligned}
d(v_0, v_i) &= 3 - |3 - i|, & ecc(v_0) &= \max\{0, 1, 2, 3, 2, 1\} = 3.\\
d(v_1, v_i) &= |3 - |4 - i||, & ecc(v_1) &= \max\{1, 0, 1, 2, 3, 2\} = 3.\\
d(v_2, v_i) &= |3 - |5 - i||, & ecc(v_2) &= \max\{2, 1, 0, 1, 2, 3\} = 3.\\
d(v_3, v_i) &= |3 - i|, & ecc(v_3) &= \max\{3, 2, 1, 0, 1, 2\} = 3.\\
d(v_4, v_i) &= |3 - |1 - i||, & ecc(v_4) &= \max\{2, 3, 2, 1, 0, 1\} = 3.\\
d(v_5, v_i) &= |3 - |2 - i||, & ecc(v_5) &= \max\{1, 2, 3, 2, 1, 0\} = 3.
\end{aligned}$$

Hence, $rad(G) = \min\{ecc(v) \mid v \in V_G\} = \min\{3, 3, 3, 3, 3, 3\} = 3$
whence $C(G) = \{v \in V_G \mid ecc(v) = rad(G)\} = V_G$.
Further, $diam(G) = \max\{ecc(v) \mid v \in V_G\} = \max\{3, 3, 3, 3, 3, 3\} = 3$.
and $P(G) = \{v \in V_G \mid ecc(v) = diam(G)\} = V_G$.

Lemma 2.7

For any graph G we have $rad(G) \leq diam(G) \leq 2rad(G)$.

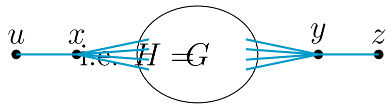
Proof. We have $rad(G) \leq diam(G)$ by definition. For the other inequality, pick $v \in C(G)$ arbitrary and consider $u, w \in V_G$ arbitrary s.t. $d(u, w) = diam(G)$. Then

$$d(u, w) \leq d(u, v) + d(v, w) \leq ecc(v) + ecc(v) = 2rad(G). \qquad \square$$

Theorem 2.8

Every graph G is isomorphic to the graph induced by the center of another graph H, i.e. ex. H s.t. $G \cong \langle C(H) \rangle$.

Proof. Let G be arbitrary. We build a new graph H which contains G as an induced subgraph via: $V_H = V_G \cup \{u, x, y, z\}$, i.e. adding 4 new vertices to G. Further, let $E_H = E_G \cup \{ux, yz\} \cup \{xv, vy \mid v \in V_G\}$.



Now $ecc(v) = 2$ for any $v \in V_G$. Nevertheless, $d(u, z) = 4$ and $d(x, z) = d(y, u) = 3$, whence $ecc(w) > 2$ for all $w \in V_H \setminus V_G$. Thus, $rad(H) = 2$ and $C(H) = V_G$, whence $\langle C(H) \rangle \cong G$. □

Lemma 2.9

A graph G is isomorphic to the graph induced by the periphery of another graph H iff either every vertex has eccentricity 1 or no vertex does.

Proof. "$\Rightarrow$" We use proof by contraposition. Assume ex. $u \in V_G$ s.t. $ecc(u) = 1 < diam(G)$. In particular, $G \neq P(G)$. Now, aiming for a contradiction, assume ex. H s.t. $G \leq H$ and $P(H) = V_G$. As $G \neq P(G)$, we know that $H \neq G$ and $diam(H) \geq 2$. As $u \in V_G = P(H)$, there is some $w \in V_H$ s.t. $d(u, w) = diam(H)$. But then, $w \in P(H) \cong V_G$, and as $ecc(u) = 1$, we also get $d(u, w) = 1 < diam(H)$. Hence, $P(H)$ cannot be V_G.

"$\Leftarrow$" If all vertices in G have eccentricity 1 or 0, then G is complete and $G \cong P(G)$. For the second case, assume $rad(G) > 1$. And consider H s.t. $V_H = V_G \cup \{v\}$ contains one new vertex which is connected to everyone else, i.e. $E_H = E_G \cup \{vx \mid x \in V_G\}$. Then, as $ecc(x) \geq 2$ for all $x \in V_G$,

$$ecc_H(x) = \begin{cases} 2 & \text{if } x \in V_G \\ 1 & \text{if } x = v \end{cases}.$$

Hence, $diam(H) = 2$ and $\langle P(H) \rangle = G$, as desired. □

2. Adjacency Matrices

We saw the visual benefits of studying graphs by their diagram. This is very useful to illustrate ideas and study small graphs. In applications on the other hand, when studying e.g. correlations of weather phenomena or social links, graphs tend to have thousands of vertices. Here, it is no longer practical to use neither the set- nor the diagram representation of graphs. The way computers store and analyze graphs is by using adjacency matrices.

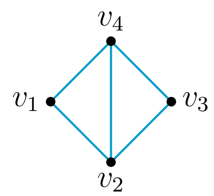
Definition 2.10

Let G be a graph of order n with vertices $V_G = \{v_1, v_2, \ldots, v_n\}$. The **adjacency matrix** of G is the matrix $A_G = (a_{ij}) \in M_{n\times n}$ defined via

$$a_{ij} = \begin{cases} 1 & \text{if } v_i v_j \in E \\ 0 & \text{otherwise.} \end{cases}$$

We also write $A(i,j)$ for a_{ij}.

Example 2.11. Consider G given by



Then $A_G \in M_{4\times 4}$

$$A_G = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

is the adjacency matrix of G.

Remark 2.12. If $A_G = (a_{ij})$ is an adjacency matrix of a graph G, then

1) $a_{ii} = 0$ for all $1 \le i \le |G|$
2) A is symmetric.
3) $\sum_{j=1}^{|G|} a_{ij} = \deg(v_i)$ and thus $\sum_{i,j=1}^{|G|} a_{ij} = \sum_{i=1}^{|G|} \deg(v_i) = 2|E|$.
4) A_G is only unique up to reordering the vertices.

Example 2.13. Let revisit the graph G from 2.11. The fact that $A_G(2,3) \neq 0$ means that v_2 and v_3 are adjacent. And $A(1,3) = 0$ says that v_1 and v_3 are not. Now consider

$$A_G^2 = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 2 & 1 \\ 1 & 3 & 1 & 2 \\ 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 3 \end{pmatrix}.$$

Let's interpret the values of A_G^2. Now, $A_G^2(1,3) = 2$. How did we compute it? $A_G^2(1,3) = \sum_{j=1}^{4} a_{1j}a_{j3}$. Now $a_{1j}a_{j3} = 1$ iff v_1v_j and v_jv_3 are edges iff (v_1, v_j, v_3) is a walk of length 2 from v_1 to v_3. Hence, $A_G^2(1,3) = \sum a_{1j}a_{j3}$ is the number of walks from v_1 to v_3 of length 2. This generalises and provides a strong tool to study graphs.

Theorem 2.14

Let G be a graph with $V_G = \{v_1, \dots, v_n\}$ and A_G the corresponding adjacency matrix. Then the entry $A_G^k(i,j)$ is the number of possible walks from v_i to v_j of length k.

Proof. We proceed by induction on the power k. (Note that $k = 0$ works too). $\underline{k = 1}$: We get that $A(i,j) = \begin{cases} 0 & \text{iff } v_iv_j \notin E_G \text{ iff there are } 0\ v_iv_j\text{-walks of length } 1 \\ 1 & \text{iff } v_iv_j \in E_G \text{ iff there is } 1\ v_iv_j\text{-walk of length } 1 \end{cases}$.

$\underline{k \to k+1}$: Assume that $A^k(i,j)$ gives exactly the number of v_iv_j-walks of length exactly k. Let's denote $A^k = (b_{ij})$ and $A = (a_{ij})$. Note that there is a v_iv_j-walk of length $k+1$ iff there ex. a vertex v_ℓ s.t. there is a v_iv_ℓ-walk of length k and an $v_\ell v_j$-walk of length one. Hence

$$\begin{aligned} |\{v_iv_j\text{-walk of length } k+1\}| &= \sum_{\ell \mid v_\ell \in N(v_j)} |\{v_iv_\ell\text{-walk of length } k\}| \\ &\overset{\text{I.H.}}{=} \sum_{\ell \mid v_\ell \in N(v_j)} b_{i\ell} = \sum_{\ell=1}^{n} b_{i\ell}a_{\ell j} \\ &= \sum_{\ell=1}^{n} A^k(i,\ell) \cdot A(\ell,j) = A^{k+1}(i,j). \end{aligned}$$

$\square$

Corollary 2.15

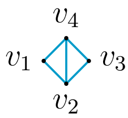
Let G be a graph with $V_G = \{v_1, \dots, v_n\}$ and A_G the adjacency matrix. Then $d(v_i, v_j) = \min\{k \mid A^k(i,j) \neq 0\}$. (Recall that $A_G^0 = I_n$).

Definition 2.16

Let G be a graph with adjacency matrix A. For every $k \in N$ we define the **Stoll matrix** S_k via

$$S_k = \sum_{i=0}^{k} A^k = I_n + A + A^2 + \cdots + A^k.$$

Remark 2.17. As $S_k(i,j) = \sum_{i=0}^{k} A^i(i,j)$, we get that $S_k(i,j)$ is the number of v_iv_j-walks of length at most k.

Example 2.18. Recall the graph $G =$ 

with $A = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$.

$A^2 = \begin{pmatrix} 2 & 1 & 2 & 1 \\ 1 & 3 & 1 & 2 \\ 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 3 \end{pmatrix}$ and $A^3 = \begin{pmatrix} 2 & 5 & 2 & 5 \\ 5 & 4 & 5 & 5 \\ 2 & 5 & 2 & 5 \\ 5 & 5 & 5 & 4 \end{pmatrix}$.

Then $S_0 = I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, $S_1 = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$, $S_2 = \begin{pmatrix} 3 & 2 & 2 & 3 \\ 2 & 4 & 2 & 3 \\ 2 & 2 & 3 & 2 \\ 2 & 3 & 2 & 4 \end{pmatrix}$.

$S_3 = \begin{pmatrix} 5 & 7 & 4 & 8 \\ 7 & 8 & 7 & 8 \\ 4 & 7 & 5 & 7 \\ 7 & 8 & 7 & 8 \end{pmatrix}$. This means there are for example 4 v_1v_3 walks of length at most 3, namely (v_1, v_2, v_3), (v_1, v_4, v_3), (v_1, v_2, v_4, v_3) and (v_1, v_4, v_2, v_3).

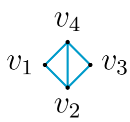
Theorem 2.19

Let G be a graph with $V_G = \{v_1, \dots, v_n\}$, adjacency matrix A and Stoll matrices S_k. Then the following hold.

1) $d(v_i, v_j)$ is the least k s.t. $S_k(i,j) \neq 0$.
2) $ecc(v_i)$ is the least k s.t. the i-th row of S_k has no zero entries.
3) $rad(G)$ is the least k s.t. S_k contains at least one row without zero entries (or ∞ otherwise).
4) $diam(G)$ is the least k s.t. S_k does not contain any zero entries.
5) G is disconnected iff S_{n-1} contains a zero.

Definition 2.20

Let G be a graph with $V_G = \{v_1, \dots, v_n\}$. The **distance matrix** of G is the matrix $D \in M_{n \times n}$ s.t. $D(i,j) = d(v_i, v_j)$.

Example 2.21. Back to our example $G =$ 

. Then the distance matrix D is

$$\begin{pmatrix} 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

Example 2.22. **Erdős Number** Paul Erdős - Hungarian Mathematician, published over 1500 papers. Consider G with $V_G =$ all mathematicians, $E_G = \{xy \mid x \text{ and } y \text{ published together}\}$. Then $\deg(\text{Erdős}) > 500$ and the Erdős number of x is $d(\text{Erdős}, x)$.

Chapter 3

Trees

1. Introdction

The intuition for graph theoretic trees comes from actual trees in nature. Here, the stem splits into several branches that afterwards never rejoin.

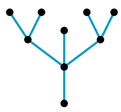
Definition 3.1

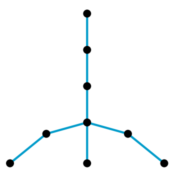
A graph which does not contain cycles is called **acyclic**. We call a graph G a **tree** if it is connected and acyclic. An arbitrary acyclic graph is called a **forest**. In a forest, any vertex of degree 1 is called a **leaf**.

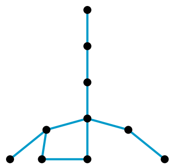
Remark 3.2.

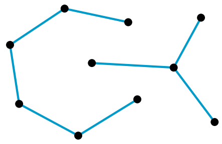
1) The graphs P_n, K_1, K_2 and $K_{1,n}$ are trees for any $n \in \mathbb{N}$.
2) Every tree is a forest.
3) Every connected component in a forest is a tree.
4) Every subgraph of a forest is a forest.

Example 3.3.

1)  **is a tree**.

2)  **is a tree**.

3)  **is not a tree**.

4)  **is not a tree**, but it **is a forest**.

Lemma 3.4

Any tree of order at least 2 has at least two leaves.

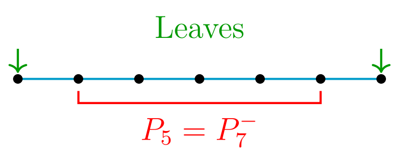
Proof. Let T be a tree with $|T| \geq 2$. In particular, T is connected. Consider a path of maximal length $P = (v_0, v_1, \dots, v_n)$ in T. As $|T| \geq 2$, we know that $v_0 \neq v_n$. We claim that v_0 and v_n are leaves, i.e. $\deg(v_0) = \deg(v_n) = 1$. We execute the argument for v_0. As usual, we know that $N(v_0) \subseteq \{v_1, v_2, \dots, v_n\}$. Let $u \in N(v_0)$ arbitrary, i.e. $u = v_i$ for some $i \geq 1$. But then $(v_0, v_1, \dots, v_i, v_0)$ is a closed walk which is a cycle for all $i \geq 2$. As T does not contain cycles, we conclude that $i = 1$ and v_1 is the only neighbour of v_0. Hence $\deg(v_0) = 1$ and v_0 is a leaf. The argument for v_n is analogous. □

Definition 3.5 Tree Pruning

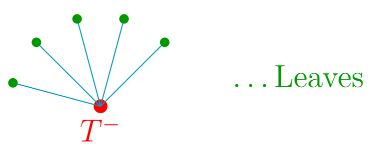
Let T be a tree of order at least 3. We denote by T^- the induced subgraph of T obtained by deleting all leaves of T.

Example 3.6.

1) If $T = P_7$ the path of length 6, then $T^- = P_7^- = P_5$ is the path of length 4:



2) If $T = K_{1,n}$ the complete bipartite graph, then $T^- = K_1 = E_1$ consists of one vertex only:



Lemma 3.7

Any tree of order n has exactly $n - 1$ edges.

Proof. We proceed by induction on $|T|$. If $|T| = 1$, then $T = K_1$, which has zero edges and the claim holds. Now assume we know that any tree of order n has exactly $n-1$ many edges and consider T of order $n+1$ arbitrary. By Lemma 3.4, T has a leaf u. Clearly, $T - u$ is still connected and of order n, whence $T - u$ has exactly $n - 1$ edges. As u was a leaf in T, T has exactly one edge more than $T - u$, whence

$$\|T\| = n = (n+1) - 1, \text{ as desired.}$$ □

Corollary 3.8

A forest of order n, consisting of k-many connected components, has exactly $n - k$ many edges.

We will now see that given a graph G is connected, Lemma 3.7 is not only a necessary, but even a sufficient condition for G to be a tree.

Theorem 3.9

A graph G of order n is a tree iff it is connected and has exactly $n - 1$ many edges.

Proof. "$\Rightarrow$" Clear by definition of a tree and Lemma 3.7.

"$\Leftarrow$" Assume G is connected of order n and contains exactly $n - 1$ many edges. If G contains a cycle, take any edge e_1 within the cycle and consider $G - e_1$. Then $G - e_1$ is still connected and of order n. If $G - e_1$ still contains a cycle, we proceed likewise and after $k \leq n-1$ many steps we obtain a graph $G - \{e_1, e_2, \ldots, e_k\}$ which is of order n, connected and without cycles, whence it is a tree. But $G - \{e_1, \ldots, e_k\}$ has $(n - 1) - k < n - 1$ many edges, contradicting Lemma 3.7. □

Theorem 3.10

A graph of order n is a tree iff it is acyclic and has $n - 1$ many edges.

Proof. "$\Rightarrow$" Clear.

"$\Leftarrow$" Assume G is of order n with $n - 1$ many edges and acyclic, i.e. G is a forest. But by Corollary 3.8, if G has k-many connected components then $\|G\| = n - k = n - 1$, whence $k = 1$ and G is connected and hence a tree. □

Corollary 3.11 Summary

Let G be a graph of order n. Then TFAE:

1) G is connected and acyclic (i.e. a tree).
2) G is connected and has $n - 1$ many edges.
3) G is acyclic and has $n - 1$ many edges.

3.12 Homework

Every edge in a tree is a bridge.

Lemma 3.13

For any two vertices $u, v \in V_T$ in a tree T, there is a unique uv-path.

Proof. As T is connected, there clearly is a uv-path for any $u, v \in V_T$. Now assume that $P_1 = (u = x_0, x_1, \ldots, x_k = v)$ and $P_2 = (u = y_0, y_1, \ldots, y_\ell = v)$ are two distinct uv-paths. Then $P_1 \cup P_2$ is again a tree. Let i be minimal s.t. $x_i \neq y_i$. Then $(P_1 \cup P_2) - y_i y_{i-1}$ is still connected, contradicting the fact that every edge in a tree is a bridge. □

Corollary 3.14

Let T be a tree and $v \in V_T$. Then $ecc(v)$ is the length of the longest path starting from v.

Lemma 3.15

Let T be a tree of order at least 2. Consider $u, v \in V_T$ s.t. $ecc(v) = d(u, v)$. Then u is a leaf.

Proof. Let $P = (v = x_0, x_1, \ldots, x_k = u)$ be the unique vu path. If u were not a leaf, then it had at least one neighbour $w \notin P$. But then $(v = x_0, x_1, \ldots, x_k, w)$ would be a path starting in v and longer than P, contradicting Corollary 3.14. □

Lemma 3.16

Let T be a tree of order at least 3. Then $C(T) = C(T^-)$.

Proof.

1) Show that $C(T) \subseteq T^-$, i.e. $C(T)$ contains no leaf. To this end, let u be a leaf and v its unique neighbour. As $|T| \geq 3$, v is not a leaf itself and $d(u, w) = d(v, w)+1$ for any $w \in V_T \setminus \{u\}$, whence $ecc(u) > ecc(v)$ and hence $u \notin C(T)$.
2) Show that $ecc_{T^-}(v) = ecc_T(v) - 1$ for every non-leaf $v \in V_T$. To that end, consider an arbitrary non-leaf $v \in V_T$ and pick $u \in V_T$ s.t. $d(v, u) = ecc(v)$. By 3.15, u is a leaf. Let P be the unique vu-path in T and note that u is the only leaf on P. Hence only u will be deleted from P in T^-. As this holds for all paths in T starting in v of length $ecc(v)$, we obtain that $ecc_{T^-}(v) = ecc_T(v) - 1$, as desired.
3) We conclude from 1) + 2) that for any vertex $v \in T^-$, $ecc_{T^-}(v) = ecc_T(v) - 1$, whence $v \in C(T)$ iff $v \in C(T^-)$ (and $rad(T^-) = rad(T) - 1$).

□

Lemma 3.17

Let T be a tree. Then $C(T)$ is either K_1 or K_2.

Proof. We do induction on $|T|$. If $|T| = 1$, then $T = K_1$ is its own center and we are done. Similarly for $|T| = 2$, where $T = K_2$. Now assume that the claim holds for all trees of order $n \geq 3$ and consider a tree T with $|T| = n + 1$ arbitrary. By 3.16, we know that $C(T) = C(T^-)$. By 3.4 we know that T contains at least two leaves, whence $|T^-| \leq |T| - 2 < n$. Hence, by I.H., $C(T) = C(T^-)$ is either K_2 or K_1 as desired. □

Lemma 3.18

Let T be a tree of order n and G an arbitrary graph s.t. $\delta(G) \geq n - 1$. Then G contains T as a subgraph.

Proof. We use induction on $|T|$. If $|T| = 1$, then $T = K_1$ is a subgraph of any graph G. Now assume we proved the claim for all trees of order at most n. Consider T with $|T| = n+1$ and G with $\delta(G) \geq n$ arbitrary. Let u be a leaf of T and denote by $T' := T - u$. Then $|T'| = n$, whence T' can be seen as a subgraph of G. Let v be the unique neighbour of u in T. Then $\deg_G(v) \geq \delta(G) \geq n$, but as $|T'| = n$ and v cannot be its own neighbour, there exist some $u' \in G$ adjacent to v and not contained in T'. Hence, the subgraph $(V_{T'} \cup \{u'\}, E_{T'} \cup \{vu'\})$ is the desired subgraph of G isomorphic to T. □

Summary

1) A tree of order n contains exactly $n - 1$ edges.
2) Any tree of order at least two contains at least two leaves.
3) A graph of order n is a tree iff it is connected of size $n - 1$.
4) A graph of order n is a tree iff it is acyclic and of size $n - 1$.
5) A graph is a tree iff for any vertices u, v there is a unique uv-path.
6) The centre of any tree is either K_1 or K_2.
7) Any graph G contains any tree of order at most $\delta(G) + 1$ as a subgraph.

2. Spanning Trees

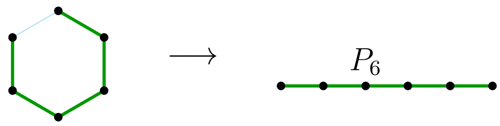
Definition 3.19

Let G be any graph. We call a subgraph $T \subseteq G$ a **spanning tree** for G if it is a tree and contains all vertices of G.

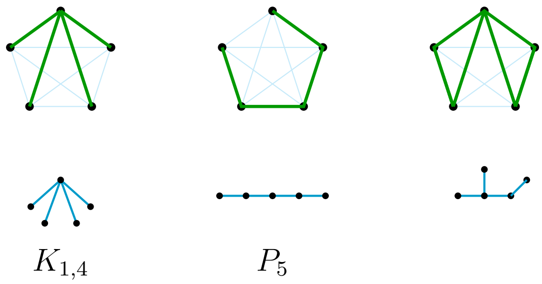
Remark 3.20. From the previous chapter it is clear that a spanning tree of a graph G of order n has n many vertices and $n - 1$ many edges.

Example 3.21. Consider the following graphs and spanning trees.

1) $G = C_6$, a possible spanning tree:



2) $G = K_5$, possible spanning trees:



Lemma 3.22

Every connected graph contains at least one spanning tree.

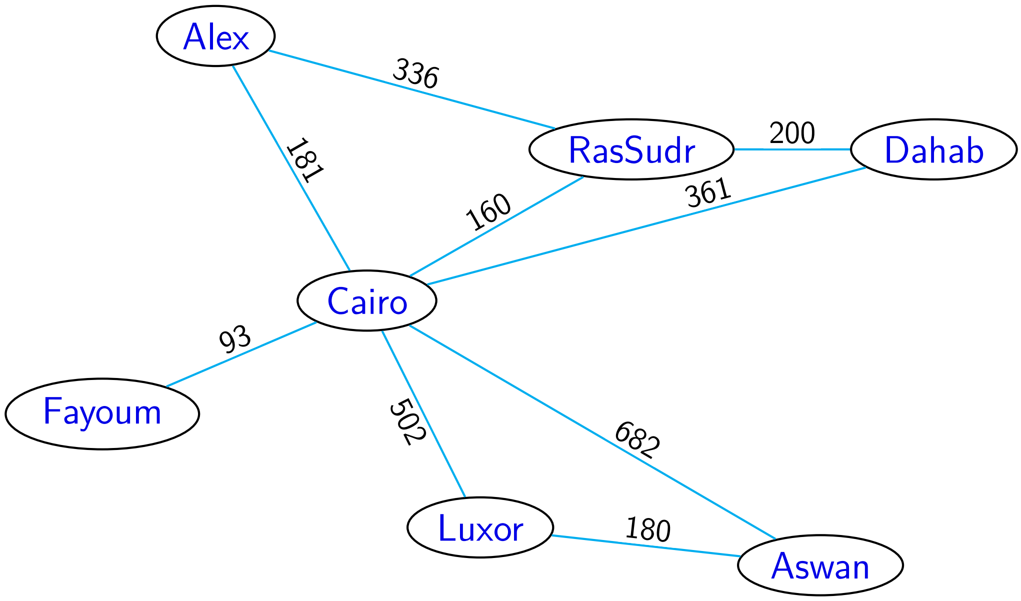
Proof. Assume G is connected and let T be a subgraph of G of maximal order s.t. T is a tree. We need to show that $V_T = V_G$. Otherwise, as G is connected, there is some vertex $u \in V_G \setminus V_T$ which is adjacent to some vertex $v \in V_T$. Now, consider the new subgraph $\hat{T} = (V_T \cup \{u\}, E_T \cup \{uv\})$. As $\deg_{\hat{T}}(u) = 1$, u is not contained in any cycles in $\hat{T}$, whence $\hat{T}$ is still a tree. As this contradicts maximality of $|T|$, we conclude that T must contain all vertices of G, whence it is a spanning tree for G. $\square$

Definition 3.23

A function $w : E_G \to \mathbb{R}$ is called a **weight function** on G. A graph G together with a weight function (i.e. the triple (V_G, E_G, w)) is called a **weighted graph**.

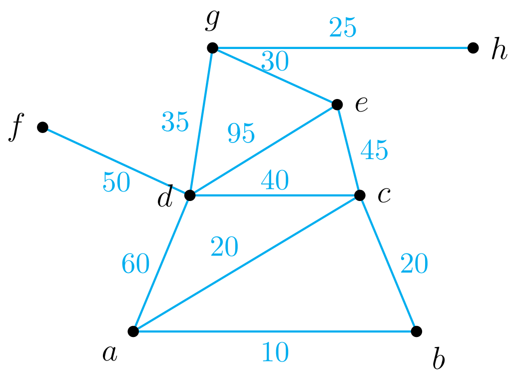
Example 3.24. *Visualisation*

We visualise the weighting of a graph by denoting the weight $w(e)$ on top of the edge e, e.g.

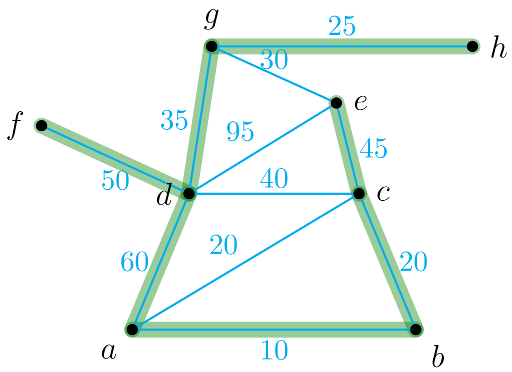


Here the weight function of an edge $e = uv$ is given by the (birds eye) distance between u and v.

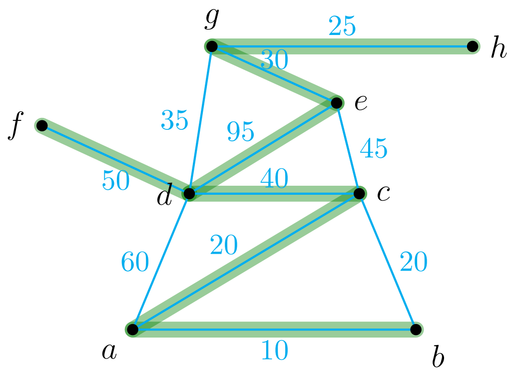
Example 3.25. Consider the following weighted graph.



We can find several spanning trees. Let's name some and compute their weight.



Total weight:
$45 + 20 + 10 + 60 + 50 + 35 + 25$
$= 245$



Total weight:
$10 + 20 + 40 + 50 + 95 + 30 + 25$
$= 270$

Definition 3.26

Let (G, w) be a connected weighted tree. A **minimum-weight spanning tree** T is a spanning tree of G s.t. the sum of the weights of its edges is minimal among all possible spanning trees of G, i.e. if $\tilde{T}$ is another spanning tree, then $\sum_{e \in E_T} w(e) \leq \sum_{e \in E_{\tilde{T}}} w(e)$.

Now how can we find a minimal spanning tree effectively? Consider the following algorithm:

3.27 Kruskal's Algorithm (1956)

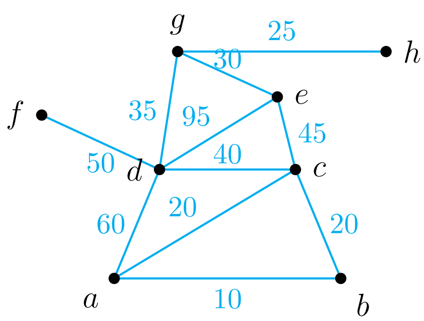
Consider the set of vertices as a forest $F = (V_G, \emptyset)$ where each vertex is a maximal subtree of F. Let $E := E_G$.

While (F is not a tree $\wedge$ $E \neq \emptyset$)

- Pick $e \in E$ of minimal weight. Let $E := E \setminus \{e\}$.
- If e connects two trees in F, let $E_F = E_F \cup \{e\}$.
- (i.e. $F + e$ is still acyclic)

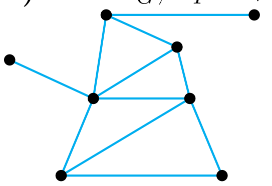
This algorithm stops after at most $|E_G|$ many repetitions.

Example 3.28. We apply the algorithm on the following weighted graph:

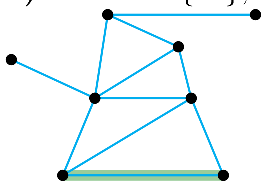


Let us mark edges we add to F green and the ones we disregard, red.

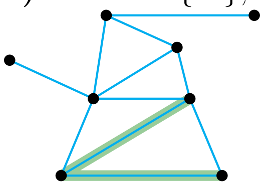
1) $E = E_G, E_F = \emptyset$



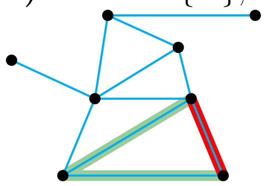
2) $E = E - \{ab\}, E_F = \{ab\}$



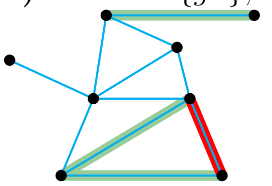
3) $E = E - \{ac\}, E_F = E_F \cup \{ac\}$



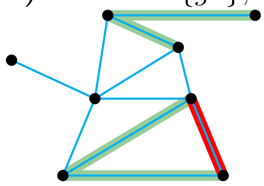
4) $E = E - \{bc\}, E_F = E_F$



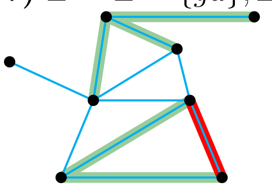
5) $E = E - \{gh\}, E_F = E_F \cup \{gh\}$



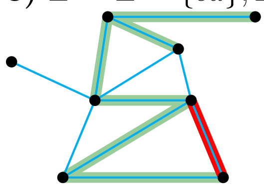
6) $E = E - \{ge\}, E_F = E_F \cup \{ge\}$



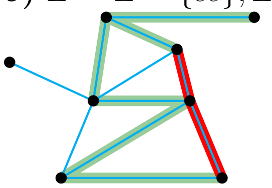
7) $E = E - \{gd\}, E_F = E_F \cup \{gd\}$



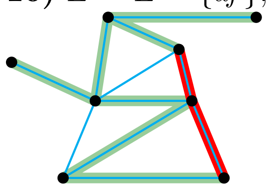
8) $E = E - \{cd\}, E_F = E_F \cup \{cd\}$



9) $E = E - \{ce\}, E_F = E_F$



10) $E = E - \{df\}, E_F = E_F \cup \{df\}$



Here the algorithm stops, as $F = (V_G, E_F)$ with $E_F = \{ab, ac, gh, eg, dg, cd, df\}$ is a single tree whence the conditions in the while loop are violated.
(The output is the spanning tree F. Note that the second condition in the while loop was still valid, as $E = \{ad, de\} \neq \emptyset$).

Theorem 3.29

Kruskal's algorithm is correct, i.e. it always terminates and its output is a minimum-weight spanning tree.

Proof. 1) Termination: As after $|E_G|$-many steps the condition $E \neq \emptyset$ is violated, the algorithm always terminates.

2) The output F is a spanning tree: As $V_F = V_G$, it clearly contains all vertices of G. Further, in each step the regarded edge e either connects two disconnected trees into one larger tree, or, if it would connect two vertices of the same subtree in F, is disregarded. Hence after each step, F is still a forest, i.e. acyclic. It remains to show that F is connected. If the algorithm stops because F is a tree, then it is clearly connected. If it stops because we went through all the edges, then any edge of G not contained in F would connect two vertices of the same connected component. Thus F has as many connected components as G, which is one, as G is connected.

3) F is a minimum-weight spanning tree. Aiming for a contradiction, assume this is not the case. Let $\{e_1, \ldots, e_{n-1}\}$ be all the edges in F, enumerated in the order they were added to F by the algorithm. Among all possible minimum-weight spanning trees, let T be one that agrees with F on the largest initial segment of $(e_1, \ldots, e_{n-1})$, i.e. if k is the smallest index s.t. $e_{k+1} \notin T$, then there is no minimum-weight spanning tree which contains $\{e_1, \ldots, e_{k+1}\}$. As by assumption F is not minimum-weight, we have $k < n-1$. As T is a spanning tree which does not contain e_{k+1}, we know that $T + e_{k+1}$ contains a cycle C. As F did not contain cycles, there is one edge $e \in C \subseteq T$ which is not in F. Now $T + e_{k+1} - e$ is a connected graph of order n and size $n-1$, whence still a spanning tree. It contains the edges $\{e_1, \ldots, e_k, e_{k+1}\}$, hence it can no longer be of minimum weight. This means that $w(e_{k+1}) > w(e)$. But as $e \notin F$ and in particular $e \notin \{e_1, \ldots, e_k\}$ this means e was available at the step of the algorithm after we added e_k and of less weight than e_{k+1}. This contradicts the assumption that the algorithm chooses the edge of minimal weight which keeps F acyclic. □

Lemma 3.30

If G is a connected weighted graph s.t. distinct edges have distinct weights, then there is a unique minimum-weight spanning tree.

Proof. Homework. □

Chapter 4

Euler and Hamilton

1. Euler

Imagine

A salesperson with their wagon wants to pass by every street in his neighbourhood to sell their goods. Of course, they want to minimize efforts, so they would like to avoid passing the same street twice. These type of problems are considered when discussing **Eulerian graphs**.

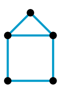
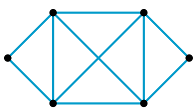
Then, as only few people buy, they switch to their car and only visit a central place in each city of the area. Again, to improve efficiency, they only want to visit each city once. This type of problem is studied when discussing **Hamiltonian graphs**.

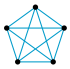
How do these two problems differ? Let's find out!

Definition 4.1

We call a trail in a graph G an **Eulerian trail** if it contains every edge of G. We call it an **Eulerian circuit** if it is a closed Eulerian trail. Finally, the graph G itself is called an **Eulerian graph** iff it contains an Eulerian circuit.

Example 4.2.

1)  **is not** Eulerian, but has an Eulerian trail.
2)  **is indeed** Eulerian.
3) Any cycle C_n is clearly Eulerian.
4) Any path of length $n \geq 1$ is **not** Eulerian but has an Eulerian trail.

5)  K_5 **is indeed** Eulerian, but

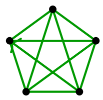
6)  K_4 **does not even contain** an Eulerian trail.

7) Generally, every K_{2n+1} is Eulerian and every K_{2n+2} does not even contain an Eulerian trail for $n \geq 1$.

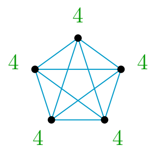
4.3 Observation

Consider the graph K_5. We observe the following properties:

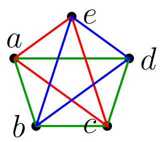
1) K_5 is Eulerian: $(a, b, c, d, e, a, d, b, e, c, a)$ is an Eulerian circuit.



2) Every vertex of K_5 has an even degree. (All 4).

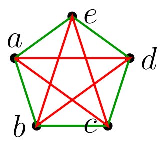


3) We can partition E_{K_5} into cycles (i.e. find mutually edge-disjoint cycles that together use all edges of K_5).



$C_1 = (a, b, c, d)$
$C_2 = (a, c, e)$
$C_3 = (b, d, e)$

or



$C_1 = (a, b, c, d, e, a)$
$C_2 = (a, c, e, b, d, a)$

These three properties do not appear together by coincidence. It turns out, they are equivalent to each other.

Lemma 4.4 Auxiliary Lemma

Let G be a connected graph with $|G| \geq 2$. If $\deg(v)$ is even for all $v \in V_G$, then G contains a cycle C. Moreover, $G - C$ still contains a cycle or is $E_{|G|}$.

Proof. Assume G is as above. If G would not contain a cycle, it was a tree. But then it had to contain a leaf v. But then $\deg(v) = 1$ is not even $\lightning$. For the "moreover" part, observe that

$$\deg^{G-C}(v) = \begin{cases} \deg^G(v) - 2 & \text{if } v \in V_C \\ \deg^G(v) & \text{else} \end{cases}$$

hence still even. Then each connected component of $G - C$ still contains a cycle (whence so does $G - C$), or is of order 1. $\square$

Theorem 4.5 Euler-Hierholzer-Veblen

Let G be a connected graph. The following are equivalent:
1) G is Eulerian.
2) Every vertex of G is of even degree.
3) The edge set of G can be partitioned into a set of edge-disjoint cycles.

Corollary 4.6

A graph contains an Eulerian trail iff either each vertex has even degree or there are exactly two vertices of odd degree.

Proof. Proof of Theorem 4.5: As all three clearly hold for $|G| = 1$, we may assume that $|G| > 1$.

1) $\Rightarrow$ 2): Assume G is Eulerian. Let Q be an Eulerian circuit of G. Now consider $v \in V_G$ arbitrary. Without loss of generalisation, we may assume that Q does not start with v. Now, every appearence of v in Q corresponds to two distinct edges involving v, the one leading *into* v and the one leading *away* from v. As Q is Eulerian, it uses all edges incident with v whence in total there is an even number of edges incident with v and $\deg(v)$ is even. ✓

2) $\Rightarrow$ 3): Assume G only contains vertices of even degree. By Lemma 4.4, G contains at least one cycle C. We proceed by induction on the number n of cycles in C. **n=1**: If G contains only one cycle, then $G = C_{|G|}$ and hence the desired partition of edges is just the cycle G itself. **n $\rightarrow$ n+1**: Now assume every graph containing at most n-many cycles allows a partition into edge-disjoint cycles. Consider any connected G with $(n+1)$-many cycles. Pick an arbitrary cycle C in G. Then as in 4.4, in $H := (V_G, E_G - E_C)$, every vertex still has even degree. Now, every connected component of H contains at most n-many cycles. By induction hypothesis, we can partition each connected component of H, and hence H itself, into edge-disjoint cycles. Once we add C to this partition, we obtain the desired partition of G. ✓(Note that this gives you a cooking recipe of how to find cycles).

3) $\Rightarrow$ 1): Assume the edge set of G can be partitioned into k-many sets $S_1, S_2, \ldots, S_k$ s.t. the edges of each S_i form a cycle. Let Q be a circuit of maximal length in G s.t. the edges of Q equals the union of some sets S_i, i.e. such that there is $I \subseteq \{1, \ldots, k\}$ with $E_Q = \bigcup_{i \in I} S_i$. As the S_i are pairwise disjoint, we know that Q contains either no edge from S_i or all edges from S_i for every $i \leq k$. Now, if $E_Q = E_G$, then G is Eulerian and we are done. Otherwise, there is some edge not contained in Q, but incident with a vertex v in Q. The edge must be contained in exactly one S_ℓ with $\ell \notin I$. Note that Q and S_ℓ have no common edges, but they share the vertex v. Hence we may glue the circuit Q

and the cycle S_ℓ at v and obtain a new circuit Q' longer than Q with $E_{Q'} = \bigcup_{i \in I \cup \{\ell\}} S_i$, contradicting our choice of Q. Hence Q contained all edges of G and hence G is Eulerian. ✓ □

2. Hamilton

Sir William Rowan Hamilton (1805–1865)

- Irish pure mathematician
- Contributions to optics, mechanics and algebra.
- Also invented a game (The Icosian Game) build on graph theory (bought by Jaques and Son, huge failure).

Definition 4.7

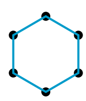
Let G be a graph. A **Hamiltonian path** is a path in G which uses all vertices of G. A **Hamiltonian cycle** is a cycle in G which uses all of V_G. We call G **traceable** if it contains a Hamiltonian path and we call it **Hamiltonian** if it contains a Hamiltonian cycle.

Remark 4.8.

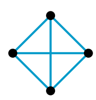
1) Every Hamiltonian graph is traceable but not vice versa.
2) Traceable graphs are connected.
3) If $|G| = n$, then G is Hamiltonian iff it contains C_n as a subgraph and it is traceable iff it contains P_n as a subgraph.

Example 4.9.

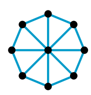
1. C_6 **is Hamiltonian** via $v_1 \dots v_6 v_1$.
   All vertices have even degree.



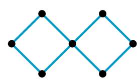
2. K_4 **is Hamiltonian** via $v_1 v_2 v_3 v_4 v_1$.
   All vertices have odd degree.



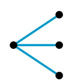
3. The graph G_1 **is Hamiltonian**.
   There are vertices of even and odd degree.



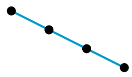
4. The graph G_2 is **not** Hamiltonian.
   Every vertex has even degree.



5. The graph $K_{1,3}$ is **not** Hamiltonian.
   All vertices have odd degree.



6. The path P_4 is **not** Hamiltonian.
There are vertices of even and odd degree.



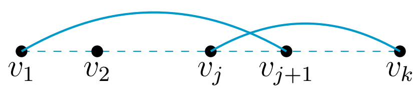
Remark 4.10. While it is rather easy to decide whether a graph is Eulerian (P-TIME, $O(|G|^2)$), it is surprisingly **hard** to do the same for Hamiltonian graphs. This problem is known to be **NP-complete** and still we did not manage to find an equivalent condition for Hamiltonianity (other than containing $C_{|G|}$ as a subgraph, which is basically the definition).

We hence see, even though the Eulerian graph problem and the Hamiltonian graph problem seem so similar, their resolution requires very different levels of efforts. The best we can do at the moment is give some **sufficient** criteria.

Theorem 4.11 Dirac

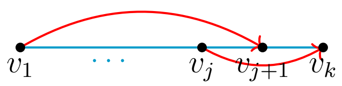
Let G be s.t. $|G| \geq 3$. If $\delta(G) \geq \frac{n}{2}$, then G is Hamiltonian.

Proof. Consider G arbitrary s.t. $|G| = n \geq 3$ and $\delta(G) \geq \frac{n}{2}$.
Then G is necessarily connected (think why).
Consider a path $P = (v_1, v_2, \ldots, v_k)$ of maximal length in G.
We claim that there is some $j < k$ s.t. $v_{j+1} \in N(v_1)$ and $v_j \in N(v_k)$. i.e.



is a subgraph of P. Note that as usual, as P is of maximal length, all neighbours of v_1 and v_k must be on P. As $\delta(G) \geq \frac{n}{2}$, v_k has at least $\frac{n}{2}$ many neighbours v_j in P. Aiming for a contradiction, assume for every neighbour $v_j \in N(v_k)$, $v_{j+1} \notin N(v_1)$. Then there are at least $\frac{n}{2}$ many vertices in P which are **not** neighbours of v_1. This now yields the desired contradiction, as all neighbours of v_1 are on P and thus $\deg(v_1) \leq (k-1) - \frac{n}{2} \leq (n-1) - \frac{n}{2} = \frac{n}{2} - 1 < \frac{n}{2}$, contradicting $\delta(G) \geq \frac{n}{2}$. Hence there is some j s.t. $v_1 v_{j+1}$ and $v_j v_k$ are edges, which leads to the existence of a cycle

$$C = (v_1, v_2, \ldots, v_j, v_k, v_{k-1}, \ldots, v_{j+1}, v_1).$$



Finally, we claim that C is indeed a Hamiltonian cycle, i.e. it contains all vertices of G. Otherwise, as G is connected, there is a vertex u in $G \setminus C$ which is adjacent to one vertex v_i in C. But as C is a cycle, we can form a new path starting in $u, v_i \ldots$ and then traveling through all $k-1$ many vertices of C. This path is longer than P, contradicting our choice of P. Hence, C indeed contains all vertices of G whence it is a Hamiltonian cycle and G is Hamiltonian. $\square$

Theorem 4.12 Fact - Ore, 1960

Let G be a graph of order $n \geq 3$. Suppose for every pair of non-adjacent vertices u, v we have that $\deg(u) + \deg(v) \geq n$. Then G is Hamiltonian.

Note that now Dirac's Theorem is a mere corollary of Ore's theorem.
We want to achieve yet another sufficient criterion for Hamiltonicity. This leads us to the so-called independence number.

Definition 4.13

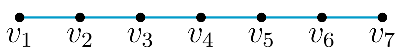
Let G be a graph. A set $S \subseteq V_G$ of vertices is called an **independent set** if any two vertices in S are nonadjacent. The **independence number** $\alpha(G)$ of G is the maximal size of an independent set.

Example 4.14.

- $\alpha(E_n) = n$, as V_{E_n} is an independent set.
- $\alpha(K_n) = 1$, as any two vertices are adjacent. Actually, the converse also holds, i.e. $\alpha(G) = 1$ iff G is complete.
- $\alpha(K_{n,m}) = \max\{n, m\}$, as any set is independent iff it is contained in one of the parts.

4.15 Notation

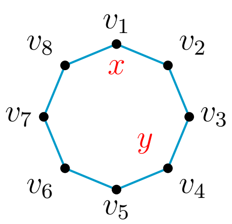
If P is a path and x, y are two vertices on P, then we denote by $P[x, y]$ the subpath on P from x to y. E.g. for



we have $P[v_6, v_3] = (v_6, v_5, v_4, v_3)$.

Similarly, if C is a cycle and $x, y \in C, x \neq y$, then we denote by $C^+[x, y]$ the xy-path on C in clockwise direction and by $C^-[x, y]$ the xy-path on C in counter-clockwise direction.

E.g. if C is



then $C^+[x, y] = (v_1, v_2, v_3, v_4)$ and $C^-[x, y] = (v_1, v_8, v_7, v_6, v_5, v_4)$.

Finally, for sequences $s = (x_1, \dots, x_\ell), t = (y_1, \dots, y_k)$ we define $s\hat{}t := (x_1, \dots, x_\ell, y_1, \dots, y_k)$ to be the concatenation of both.

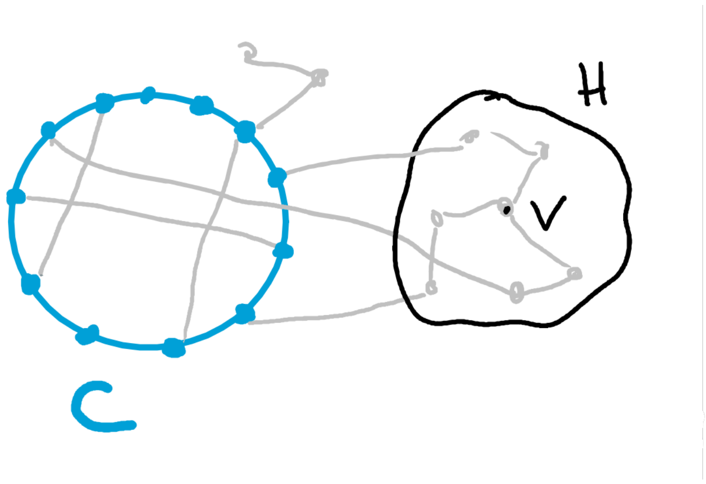
Theorem 4.16 Chvátal, Erdős, 1972

Let G be a graph of order at least 3. If $\kappa(G) \geq \alpha(G)$, then G is Hamiltonian.

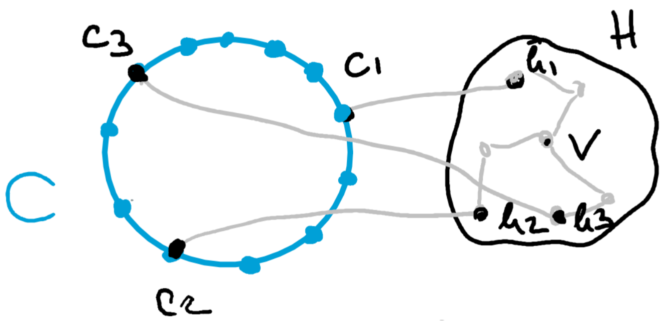
Proof. Let G be as above, i.e. $|G| \geq 3, \kappa(G) \geq 1, \kappa(G) \geq \alpha(G)$.

- First we argue that $\kappa(G) \geq 2$. Otherwise $\kappa(G) = \alpha(G) = 1$, whence G is a complete graph. As further $\kappa(K_n) = n - 1$, G would be K_2, contradicting $|G| \geq 3$.
- Hence now we know that $\kappa(G) \geq 2$. By 1.41(8), we know that $\delta(G) \geq \kappa(G) \geq 2$, whence by 1.39, G contains a cycle.

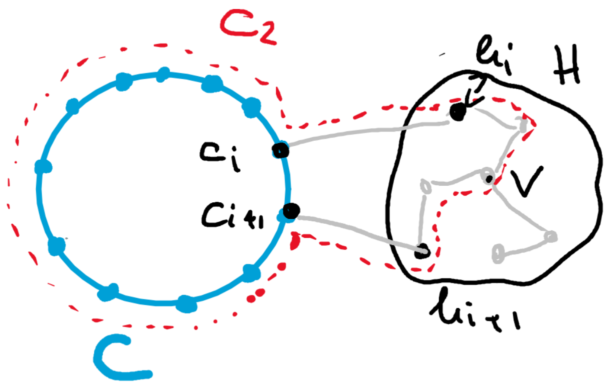
Now consider a cycle C of maximal length in C. We claim that C is Hamiltonian. Aiming for a contradiction, assume C is not Hamiltonian, i.e. there is some vertex $v \notin C$. Let H be the connected component of v in $G \setminus C$.



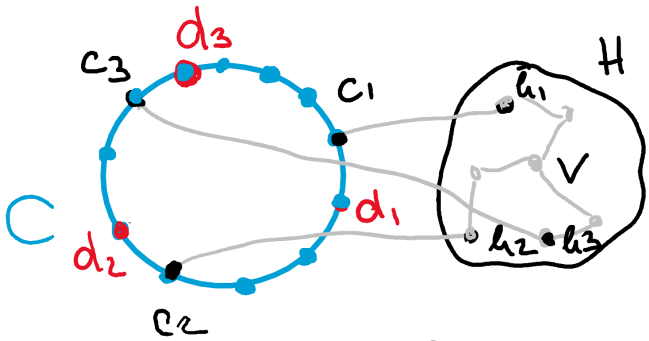
Now, we list all elements of C which are connected to some vertex in H in clockwise order: $\{c_1, c_2, \dots, c_r\}$ (s.t. $c_j \in C^+[c_{j-1}, c_{j+1}]$), i.e. where each c_i is adjacent to some $h_i \in H$.



Claim 1: No two c_i's are consecutive vertices in C. Proof: Otherwise assume there is an i s.t. c_{i+1} is the clockwise successor of c_i. Let P be a path from h_i to h_{i+1} in H. Then $C^+[c_{i+1}, c_i]\hat{}(c_i h_i)\hat{}P\hat{}(h_{i+1} c_{i+1})$ is a cycle strictly longer than C, contradicting our assumptions. $\lightning$

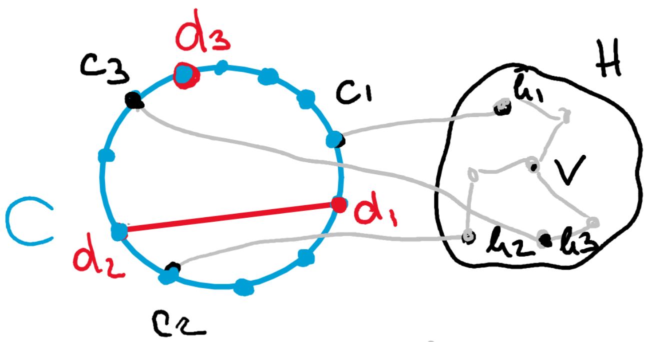


Now that no two c_i and c_j are clockwise successors, we can define the set $D = \{d_1, d_2, \ldots, d_r\}$ where each d_i is the clockwise successor of c_i in C and we get that $\{c_1, \ldots, c_r\} \cap \{d_1, \ldots, d_r\} = \emptyset$.

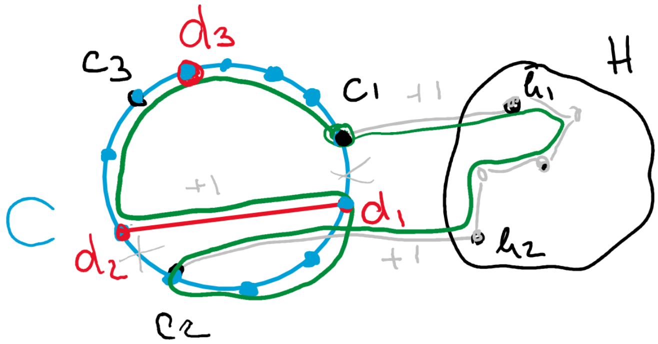


Claim 2: $\{c_1, \ldots, c_r\}$ is a cut set for G. Proof: This is clear as any path from v to a vertex in C has to pass through one of the vertices in $\{c_1, \ldots, c_r\}$, so $G - \{c_1, \ldots, c_r\}$ is disconnected. Consequently, as $\kappa(G)$ is the size of a smallest cut set, we obtain that $r \geq \kappa(G) \geq 2$.

Claim 3: There are d_i and d_j which are adjacent. Proof: Consider the set $X := \{d_1, d_2, \ldots, d_r, v\}$, and recall that there is no edge between any d_i and v. As $|X| = r + 1 \geq \kappa(G) + 1 > \alpha(G)$, X cannot be an independent set, whence at least one pair d_i, d_j must be adjacent.



Now we are ready for our final contradiction: We produce a cycle $\hat{C}$ longer than C. Assume $d_i d_j$ is an edge and $i < j$. Let $q_{h_i h_j}$ be a path in H from h_i to h_j. Now define $\hat{C} = (c_i)\hat{}q_{h_i h_j}\hat{}(c_j)\hat{}C^-[c_j, d_i]\hat{}(d_i d_j)\hat{}C^+[d_j, c_i]$.



Note that $\hat{C}$ uses all edges of C except the edges $c_i d_i$ and $c_j d_j$. Instead it uses at least the three additional edges $c_i h_i, h_j c_j$ and $d_i d_j$. Hence, $\hat{C}$ is strictly longer than the cycle C, which contradicts our choice of C. Conclusively, C must contain all vertices of G and hence is Hamiltonian. $\square$

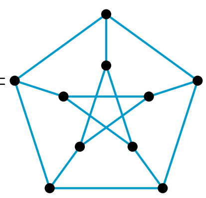
So far, we have encountered sufficient criteria for Hamiltonian graphs using the degree of vertices and the independence number. We will conclude the chapter by providing a last sufficient criterium using a new concept - **forbidden subgraphs**.

Definition 4.17

Let H and G be graphs. We say that G is ***H*-free** if H is not (isomorphic to) an induced subgraph of G. Moreover, if S is a collection of graphs, then we call G ***S*-free** iff G is H-free for any $H \in S$.

Example 4.18. The Petersen graph

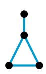
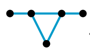
is **indeed** C_3-free
is **not** C_5-free
is **not** E_4-free
is **indeed** E_5-free

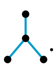
$G =$ 

Recall:
$\kappa(G) = 3$
$\alpha(G) = 4$

G is hence also $\{C_3, E_5\}$-free.

4.19 Notations

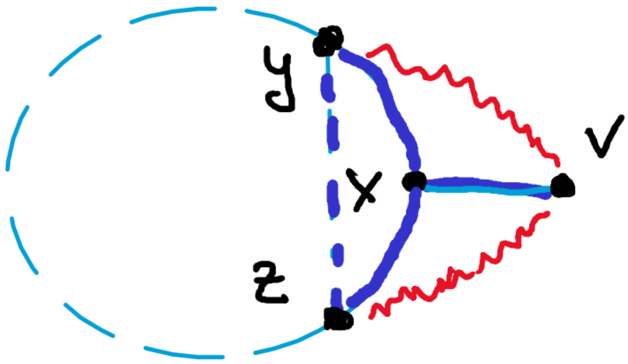
Let Z_1 be the graph $Z_1 =$  and $N =$ .

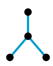
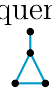
Further, we call the graph $K_{1,3}$ the **claw**, based on its shape: $K_{1,3} =$ , or also $K_{1,3} =$ .

Theorem 4.20 Goodman, Hedetniemi, 1974

Let G be 2-connected and $\{K_{1,3}, Z_1\}$-free, then G is Hamiltonian.

Proof. As $\delta(G) \geq \kappa(G) \geq 2$, we get that G contains a cycle. Consider such a cycle C of maximal length. We claim that C is Hamiltonian.
Otherwise, as G is connected, there was a vertex $v \in V_G$ not on C but adjacent to some vertex x on C, i.e.



Denote by y and z the neighbours of x on C.
Note that yv is not an edge as otherwise replacing the subsequence (y, x) in C by (y, v, x) would yield a cycle longer than C. Similarly, vz is not an edge. Consequently, the induced subgraph on $S = \{x, y, z, v\}$ is either $\langle S \rangle = K_{1,3}$  or $\langle S \rangle = Z_1$,  both of which contradict our assumptions on G. □

The following final condition is now easy to verify:

Theorem 4.21 Duffus, Gould, Jacobson, 1980

Let G be a $\{K_{1,3}, N\}$-free graph.
i) If G is connected, it is traceable.
ii) If G is 2-connected, it is Hamiltonian.

Note that neither of these are necessary for G to be Hamiltonian. Indeed, for any graph H there is a Hamiltonian graph G which contains H as an induced subgraph.

Chapter 5

Planarity

1. Planar Graphs

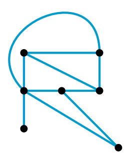
(lat. planaris - flat, level)

Definition 5.1

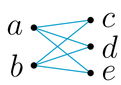
A **graph** $G = (V, E)$ is called **planar** if it can be drawn on a plane s.t. its edges at most intersect in their end vertices. Any such drawing is then called a **planar representation** or a **planar embedding**. If G is not planar, it is called **nonplanar**.

Example 5.2.

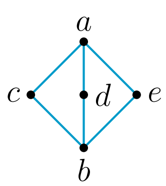
1) Clearly, trees, cycles and empty graphs are planar.
2) The graph



is planar and this is a planar representation:

3) The graph $K_{2,3}$ 

is planar, but this is not a planar representation. One planar representation is given by



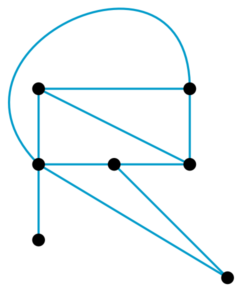
In order to prove that a graph is planar, we "just" have to provide one planar representation. But these can be very hard to find. In order to prove that a graph is <u>not</u> planar, we would have to check "all" possible representations. This is not feasible.

But we can help ourselves by throwing some math on the problem. To this end, we need some more terminology.

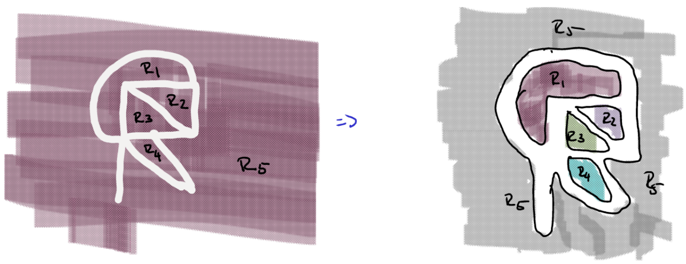
Definition 5.3

A **region** in a planar rep. is a maximal area of the plane s.t. any two points within can be connected by a curve which does not intersect or touch any part of the graph. Regions which are completely bounded (= surrounded) by graph edges are called **interior regions**. The unique non-interior region is called **exterior region**.

Example 5.4. Consider the planar representation



We can find the regions by considering this representation as a cookie cutter which we use to part the dough on our table surface. So the regions are:

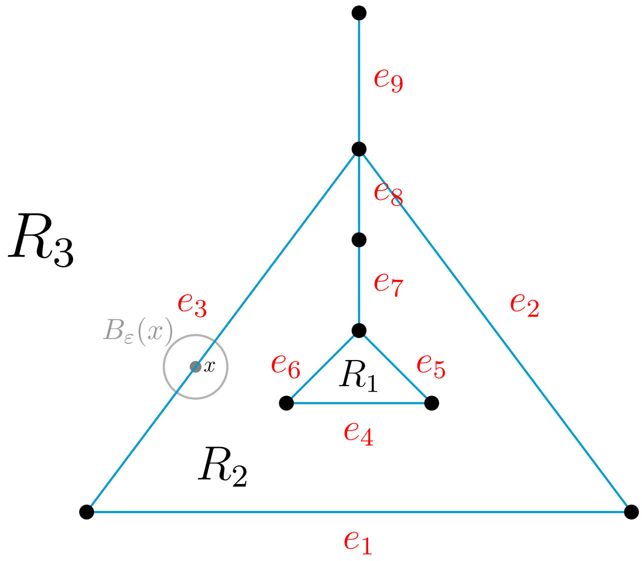


where $R_1 - R_4$ are interior and R_5 is the exterior region.

Definition 5.5

Let Γ be a planar representation of G in $\mathbb{R}^2$. We say that an edge e is **incident** with a region R iff every point x on the edge e and $\varepsilon > 0$ there is a point y in R s.t. $d(x, y) < \varepsilon$ (or $B_\varepsilon(x) \cap R \neq \emptyset$). We say that e is a **bound** for R, if it is incident with R and at least one other region. We denote the number of bounds for a given region R by $b(R)$ and call it its **boundary degree** while $B(R) = \{e \in E(G) \mid e \text{ is a bound for } R\}$ is called the **boundary** of R.

Example 5.6. Consider G given by its planar representation



Then e_6 is a bound of R_2, as it is incident with R_1 and R_2, but e_7 is not a bound of R_2, as it is only incident with R_2. Further, $B(R_2) = \{e_1, e_2, e_3, e_4, e_5, e_6\}$ and $b(R_2) = 6$. Also, $B(R_3) = \{e_1, e_2, e_3\}$ and $B(R_1) = \{e_4, e_5, e_6\}$.

Remark 5.7.

- A very proper treatment needs a good fusion of real analysis with topology and exceeds our time frame (nice thesis topic).
- Any edge is incident with either 1 or 2 regions, whence

$$\sum b(R) = \sum |B(R)| = 2|\{e \mid e \text{ is a bound for some } R\}| \leq 2|E|.$$

- Any region either has no bounds or at least three.
- A region has no bound iff it is the only region and G is a forest.

5.8 Fact

Assume Γ is the planar representation of a graph G with regions $R_1, R_2, \ldots, R_r$. Then

1) Either G is a forest and $r = 1$ and $b(R_1) = 0$ or
2) G contains a cycle and $r \geq 2$ with

$$3r \leq \sum_{i=1}^{r} b(R_i) \leq 2|E|.$$

5.9 Observation

Let e be the edge of a planar graph G. Then e is a bound for some region R iff e is part of a cycle in G.

We now will prove that the relation discovered in the intermezzo holds for any planar representation.

Theorem 5.10 Euler's Formula

Let Γ be any planar representation of a connected graph G. Then for $|G| = n, \|G\| = m$ and r the number of regions in Γ, we have

$$\boxed{n - m + r = 2.}$$

Theorem 5.11 Corollaries

1) As consequently $r = \|G\| - |G| + 2$, we get that for a planar graph G, the number of regions is independent from the chosen planar representation. We can hence say that G has r-many regions.
2) If G is a planar graph on k-many connected components $C_1 \dots C_k$ then each C_i is planar. If C_i has r_i-many regions, note that G has $\sum_{i=1}^{k} r_i - (k-1)$ many regions as all components share the common exterior region in a joint embedding. Thus the number of regions r of G is

$$r = (\sum r_i) - (k-1) = \sum(\|C_i\| - |C_i| + 2) - (k-1)$$

$$= \|G\| - |G| + 2k + 1 = \|G\| - |G| + k + 1.$$

Hence, for arbitrary planar G with $|G| = n, \|G\| = m$ and $\rho(G) = r$, we get

$$\boxed{n - m + r = k + 1}$$

Proof of Theorem 5.11. Let G be a connected planar graph in any given planar representation with regions $R_1, R_2, \dots, R_r$. Let $|G| = n$. We prove the theorem by induction on $\|G\| = m$.

$m = 0$: If G has no edges, then $G = K_1$. Then clearly $n = 1, r = 1$ and thus $n - m + r = 1 - 0 + 1 = 2$, as desired.

$m - 1 \to m$: Assume we established the claim for all graphs with less than m many edges. Consider G with m many edges. If G is a tree, then we know that $n = m + 1$ and $r = 1$ (whence $n - m + r = m + 1 - m + 1 = 2$, as desired). Otherwise, G contains at least one cycle. Let $e \in E(G)$ be one edge on that cycle. Then by Observation 5.10, e is a bound for two regions R_i and R_j. Note that in $G - e$ the regions R_i and R_j merge together to one new region R', whence $G - e$ has one less region than G. Thus, by I.H. we get

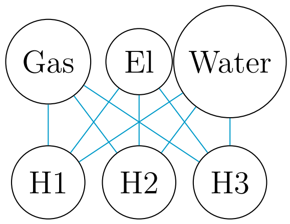
$$|G - e| - \|G - e\| + (r - 1) = n - (m - 1) + (r - 1) = n - m + r, \text{ as desired.} \qquad \square$$

Remark 5.12.

1) Every subgraph of a planar graph is planar.
2) If G is planar and R is any region, then in the induced planar representation for $B(R)$ (i.e. deleting all drawings apart from $B(R)$), R is still a region and $B(R)$ its boundary.
3) For any region we have $b(R) \in \{0\} \cup \mathbb{N}_{\geq 3}$. In particular, if $b(R) = 3$, then $B(R) \cong C_3$.

2. Nonplanar Graphs

Imagine you want to connect three houses each to gas, electricity and water, s.t. the pipes don't intersect. This question boils down to asking: Is the utility graph $K_{3,3}$ a planar graph?



(This is why $K_{3,3}$ is also called the **utility graph**.)

Theorem 5.13

The utility graph $K_{3,3}$ is not planar.

Proof. Aiming for a contradiction, assume $K_{3,3}$ was planar. Then it needed to have $r = \|K_{3,3}\| - |K_{3,3}| + 2 = 9 - 6 + 2 = 5$ regions. On the other hand, as $K_{3,3}$ is bipartite and thus does not contain odd cycles, every region has at least four bounds by Remark 5.13(3). Hence

$$\sum_{i=1}^{5} b(R_i) \geq 5 \cdot 4 = 20 > 2 \cdot 9 = 2 \cdot \|K_{3,3}\|,$$

contradicting Fact 5.9. Thus, $K_{3,3}$ cannot be planar. □

Theorem 5.14

Let G be planar of order at least 3. Then $\|G\| \leq 3(|G| - 2)$. Further, if equality holds then every region is bounded by exactly three edges.

Proof. Assume G consists of k connected components. The equation clearly holds for forests as then $\|G\| = |G| - k \leq 3|G| - 6$ iff $6 - k \leq 2|G|$. Otherwise, $r \geq 2$ and G contains a cycle. Recall that

$$(*) 3 \cdot r \leq \sum b(R_i) \leq 2\|G\|.$$

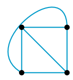
Hence, by Euler's generalised formula, we get $r = \|G\| - |G| + k + 1 \geq \|G\| - |G| + 2$. This yields

$$2\|G\| \geq 3r \geq 3(\|G\| - |G| + 2), \text{ whence } 3(|G| - 2) \geq \|G\|, \text{ as desired.}$$

Finally, for equality to hold we need in particular that $2\|G\| = 3r$, whence also $3r = \sum b(R_i)$. Thus, every region has boundary degree 3 as desired. □

Corollary 5.15

The complete graph K_n is planar iff $n \leq 4$.

Proof. First note that $K_4 =$  is planar, and hence so are its subgraphs K_1, K_2 and K_3. Now, for K_5, by Lemma 5.15 we get $10 = \|K_5\| \leq 3(\|K_5\| - 2) = 3(3) = 9$, a contradiction. Hence, K_5 is not planar and so is neither of the K_n for $n \geq 5$, as they contain K_5 as a subgraph. □

Theorem 5.16

If G is planar, then $\delta(G) \leq 5$.

Proof. Let G be a planar graph and set $n = |G|$ and $m = \|G\|$. Aiming for a contradiction, assume $\delta(G) > 5$. Then in particular, $n \geq 6$. Then

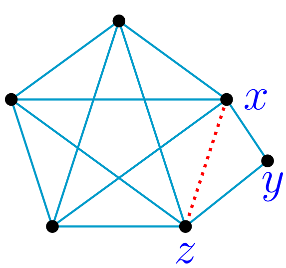
$$6 \cdot n \leq \sum_{v \in V_G} \deg(v) = 2m \overset{(5.15)}{\leq} 2(3(n-2)) = 6n - 12$$

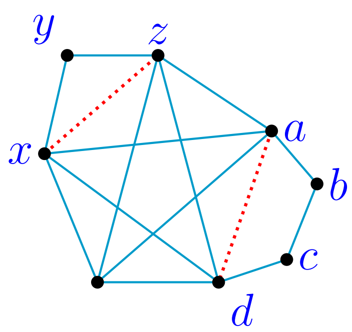
yields a contradiction. □

3. Kuratowski's Theorem

We have seen that the graphs $K_{3,3}$ and K_5 are not planar. It turns out that these two graphs are the major obstruction for any graph to be planar. This section gives an introduction to Kuratowski's theorem. But first, we want to introduce a new notation.

Example 5.17. Consider

$G =$ 

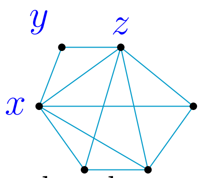
$H =$ 

Note that neither of G or H contains K_5 (or $K_{3,3}$) as a subgraph. Nevertheless, they both look so similar to K_5 that they should not be planar. Indeed, G arises from K_5 by replacing the edge xz by $\{xy, yz\}$ and H arises from G by replacing the edges $\{xz, ad\}$ by $\{xy, yz, ab, bc, cd\}$. This process is called subdivision.

Definition 5.18

1) Let G be a graph and $e = xy \in E(G)$. An **edge subdivision** of e is the replacement of e by a finite path of length ≥ 2 starting in x and ending in y.
2) Let G and H be graphs. Then H is called a **subdivision** of G iff H can be obtained through a finite sequence of edge subdivisions.

Example 5.19.

1) The graphs G, H from 5.18 are edge subdivisions of K_5.
2) The graph 

is **not** an edge subdivision of K_5, as the edges $\{xy, yz\}$ were added rather than replacing the edge xz.

Lemma 5.20

A graph G is planar iff every subdivision of G is planar.

Proof. "$\Rightarrow$" Consider a planar representation Γ of G. Let H be a subdivision of G where a sequence of n-many edge subdivisions were performed. We do induction on n.

If $n = 0$, then $H = G$ is clearly planar.

Now assume we proved the claim for n and consider H arising from G through $n+1$ many subdivisions. Let H_0 be the graph arising from G through the first n-many subdivisions. By I.H. H_0 is planar and H can be obtained from H_0 through exactly one edge-subdivision. Let Γ_0 be a planar drawing of H_0 and $e = xy$ be the edge which is subdivided to obtain H, say by replacing it by the path $P = (x = x_0, x_1, \ldots, x_k = y)$ for $k \geq 2$. Let $\vec{xy}$ be the geodesic from x to y in $\mathbb{R}^2$. Then we draw in the vertex x_i at the point $x + \frac{i}{k}\vec{xy}$ for any i. This yields a planar representation of H, as desired. $\square$

Corollary 5.21

If G contains a subdivision of $K_{3,3}$ or K_5 as a subgraph, then G is not planar. (Or: If G planar $\Rightarrow$ no subdivision of $K_5, K_{3,3} \subseteq G$)

Theorem 5.22 Kuratowski's Theorem

A graph is planar if and only if it does not contain a subdivision of $K_{3,3}$ or K_5 as a subgraph.

Remark 5.23. The left over, hard direction is the backwards direction, i.e. if G does not contain a subdivision of $K_{3,3}$ or K_5, then G is planar.