"SAPIENZA" UNIVERSITY OF ROME

FACULTY OF INFORMATION ENGINEERING,
INFORMATICS AND STATISTICS

DEPARTMENT OF COMPUTER SCIENCE

Graph Theory

Author

Alessio Bandiera

April 18, 2025

Contents

Information and Contacts

Personal notes and summaries collected as part of the *Graph Theory* course offered by the degree in Computer Science of the University of Rome "La Sapienza".

Further information and notes can be found at the following link:
https://github.com/aflaag-notes. Anyone can feel free to report inaccuracies, improvements or requests through the Issue system provided by GitHub itself or by contacting the author privately:

- Email: **alessio.bandiera02@gmail.com**

- LinkedIn: **Alessio Bandiera**

The notes are constantly being updated, so please check if the changes have already been made in the most recent version.

Suggested prerequisites:

- Progettazione degli Algoritmi

Licence:

These documents are distributed under the **GNU Free Documentation License**, a form of copyleft intended for use on a manual, textbook or other documents. Material licensed under the current version of the license can be used for any purpose, as long as the use meets certain conditions:

- All previous authors of the work must be **attributed**.

- All changes to the work must be **logged**.

- All derivative works must be **licensed under the same license**.

- The full text of the license, unmodified invariant sections as defined by the author if any, and any other added warranty disclaimers (such as a general disclaimer alerting readers that the document may not be accurate for example) and copyright notices from previous versions must be maintained.

- Technical measures such as DRM may not be used to control or obstruct distribution or editing of the document.

<div align="right">

1

</div>

Basics of Graph Theory

In the 18th century, in the city of Königsberg (Prussia), a puzzle captured the imagination of the townspeople. Königsberg, nestled along the winding *Pregel River*, was divided into four land masses — two parts of the mainland and two islands, Kneiphof and Lomse. Connecting these regions were **seven bridges**, crisscrossing the river back and forth.

Over time, a curious question arose among the people of Königsberg: was it possible to take a *walk* through the city, crossing each of the seven bridges **exactly once**, without retracing any steps, and ending the walk in the same place where it started? This is known as the Seven Bridges of Königsberg problem.

It seemed simple enough, yet no one had managed to do it. The challenge became a favorite pastime, debated in marketplaces and whispered about in taverns. Some claimed it was possible with the right path, while others remained skeptical.

Figure 1.1: The map of Königsberg in Euler's time, showing the actual layout of the seven bridges [con25].

Word of this peculiar problem reached the brilliant Swiss mathematician Leonhard Euler, a man whose mind was always drawn to patterns and logic. Intrigued, Euler set out to solve the riddle — not by drawing endless maps or walking the streets himself, but by **abstracting** the problem into something entirely new.

Euler realized that the specific layout of the city was *irrelevant*. What truly mattered was the way the landmasses were connected by the bridges. He represented each landmass as a **dot** and each bridge as a **line** between them. In doing so, he stripped away unnecessary details and created a simple, elegant combinatorial structure that we now refer to as **graph**.
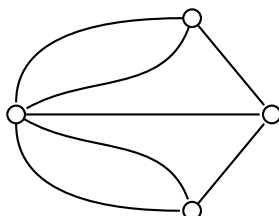

Figure 1.2: The graph drawn by Euler which models the *Seven Bridges of Königsberg* problem.

Through his analysis, Euler discovered a fundamental rule: for a walk to cross each bridge exactly once and return to the starting point, every landmass had to be connected by an **even** number of bridges. In Königsberg's case, however, each landmass had an odd number of bridges, making the task impossible.

Euler's proof, published in 1736, was groundbreaking — not just because he solved the Königsberg puzzle, but because he laid the foundation for an entirely new branch of mathematics: graph theory. His ideas would go on to shape the study of networks, from modern transportation systems to social media connections and even the vast web of the internet itself. And so, from a simple question about bridges in a small Prussian city, a whole new field of mathematics was born — one that continues to shape the world centuries later.

This chapter will discuss the basics of the field of **graph theory**, and will lay the foundation for later chapters.

1.1 Introduction

> **Definition 1.1: Graph**
>
> A **graph** is a pair $G = (V, E)$, where V is the — finite — set of **vertices** of the graph, and E is the set of **edges**.

For now, will assume to be working with **simple** and **undirected** graphs, i.e. graphs in which the set of edges is defined as follows

$$E \subseteq [V]^2 = \{\{x, y\} \mid x, y \in V \wedge x \neq v\}$$

where the notation $\{x, y\}$ will be used to indicate an edge between two nodes $x, y \in V$, and will be replaced with $xy = yx$ directly — the *set* notation for edges is used to highlight that edges have no direction.

We will indicate with n and m the cardinality of $|V|$ and $|E|$, respectively. Moreover, we will indicate with $V(G)$ and $E(G)$ the set of the vertices and edges of G respectively when there is ambiguity.
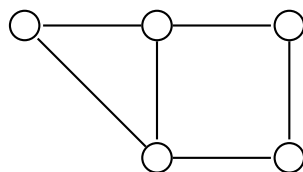
Figure 1.3: A simple graph.

Note that, in this definition, we are assuming that each edge has exactly 2 *distinct* endpoints — i.e. the graphs do not admit **loops** — and there cannot be two edges with the same endpoints — i.e. the graphs do not admit **parallel edges**. In fact, if we drop these assumptions, we obtain what is called a **multigraph**.
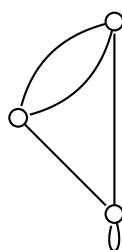
Figure 1.4: A multigraph.

Definition 1.2: Subgraph

Given a graph $G = (V, E)$, a **subgraph** $G' = (V', E')$ of G is a graph such that $V' \subseteq V$ and $E' \subseteq E$. If G' is a subgraph of G, then G is called **supergraph** of G'.

Figure 1.5: This is a subgraph of the graph shown in Figure 1.3.

Definition 1.3: Induced subgraph

Given a graph $G = (V, E)$, and a set of vertices $S \subseteq V$, then $G[S]$ represents the **subgraph induced by** S **on** G, obtained by removing from G all the nodes of $V - S$ and the edges incident on them.

In other words, a subgraph $G' = (V', E')$ of G is **induced** if every edge of G with both ends in V is an edge of V'. We observe that this 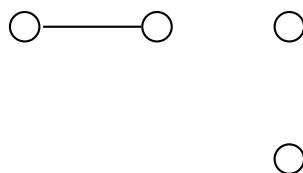definition is *stricter* than the previous one: in fact, the last graph is *not* an example of an induced subgraph, but the following is:
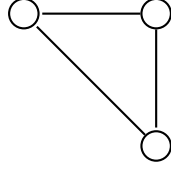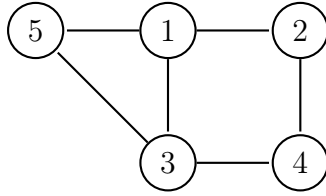

Figure 1.6: This is an *induced* subgraph of the graph shown in Figure 1.3.

Note that every induced subgraph of a graph is **unique** by definition, and we indicate each induced subgraph as follows: suppose that the graph in Figure 1.3 had the following *labeling* on the vertices

then, the induced subgraph in Figure 1.6 would have been referred to as $G[\{1, 3, 5\}]$.

Intuitively, two vertices $x, y \in V$ are said to be **adjacent** if there is an edge $xy \in E$, and we write $x \sim y$. If there is no such edge, we write $x \nsim y$ for non-adjacency and we say that xy is an **anti-edge**. The **neighborhood** of a vertex $x \in V$ is the set of vertices that are adjacent to x, and it will be indicated as follows

$$\mathcal{N}(x) := \{y \in V \mid x \sim y\}$$

Similarly, the neighborhood of a set of vertices will be defined as follows

$$\forall S \subseteq V \quad \mathcal{N}(S) := \bigcup_{v \in S} \mathcal{N}(v)$$

The **degree** of a vertex $x \in V$, denoted with $\deg(x)$, is exactly $|\mathcal{N}(x)|$. We will use the following notation for the **minimum** and **maximum** degree of a graph, respectively

$$\delta := \min_{x \in V} \deg(x) \qquad \Delta := \max_{x \in V} \deg(x)$$

Lemma 1.1: Handshaking lemma

Given a graph $G = (V, E)$, it holds that

$$\sum_{x \in V} \deg(x) = 2\,|E|$$

Proof. Trivially, the sum of the degrees counts every edge in E exactly twice, once for each of the 2 endpoints. ☐

Corollary 1.1

Given a graph $G = (V, E)$, it holds that $|E| \geq \frac{\delta \cdot n}{2}$.

Proof. By the Handshaking lemma, we have that

$$2\,|E| = \sum_{x \in V} \deg(x) \geq \delta \cdot n \iff |E| \geq \frac{\delta \cdot n}{2}$$

☐

Definition 1.4: k-regular graph

A graph G is said to be k-**regular** if every vertex of G has degree k.

Note that in a k-regular graph it holds that

$$\sum_{x \in V} \deg(x) = k \cdot n$$

Proposition 1.1

There are no k-regular graphs with k odd and an odd number of vertices.

Proof. By way of contradiction, suppose that there exists a k-regular graph $G = (V, E)$ such that both k and n are odd; however, by the handshaking lemma we would get that

$$2\,|E| = \sum_{x \in V} \deg(x) = k \cdot n$$

but the product of two odd numbers, namely k and n, is still an odd number, while $2\,|E|$ must be even ↯. ☐

1.2 Important structures

Definition 1.5: Path

A **path** is a *graph* with vertex set $x_0, \ldots, x_n$ and edge set $e_1, \ldots, e_n$ such that $e_i = x_{i-1}x_i$.

The **length** of a path is the number of edges between x_0 and x_n, i.e. $|\{e_1, \ldots, e_n\}|$, namely n in this case. A path of length 1 is called *trivial* path.
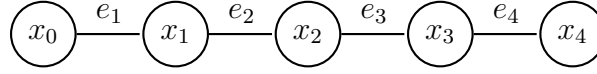
Figure 1.7: A path graph of length 4 that links x_0 and x_4.

Through *paths* we can provide the definition of **distance** between two nodes of a graph.

> **Definition 1.6: Distance**
>
> Given a graph $G = (V, E)$, and two vertices $x, y \in V$, the **distance** between x and y in G, denoted with $\text{dist}_G(x, y)$, is defined as the length of the *shortest* path between x and y in G.

If there is no ambiguity, we will simply write $\text{dist}(x, y)$ instead of $\text{dist}_G(x, y)$. Finally, given a path P and two vertices $u, v \in V(P)$, we will denote with $u P v$ the *subpath* of P between u and v. Now, consider the following definition.

> **Definition 1.7: Walk**
>
> Given a graph $G = (V, E)$, a **walk** is a *sequence* of vertices and edges
>
> $$x_0 \ e_1 \ x_1 \ \ldots \ x_{k-1} \ e_k \ x_k$$
>
> where $x_0, \ldots, x_k \in V$, $e_1, \ldots, e_k \in E$ and $e_i = x_{i-1} x_i$.
>
> The **length** of a walk is the number of edges between x_0 and x_k, i.e. $|\{e_1, \ldots, e_k\}|$, namely k in this case. If $x_0 = x_k$ we say that the walk is **closed**.

If there is a path – or a walk — between two vertices $x, y \in V$, we say that the path — or the walk — **links** x and y, and we write this as $x \to y$. Any vertex that is different from x and y is called *internal node*. For instance, given the previous graph labeled as follows
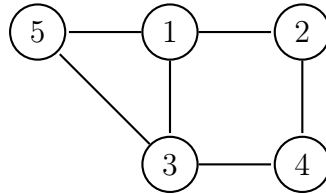
an example of a walk over this graph is given by the following sequence

$$1 \ \{1, 2\} \ 2 \ \{2, 4\} \ 4 \ \{4, 3\} \ 3 \ \{3, 1\} \ 1 \ \{1, 5\} \ 5$$

that *links* 1 and 5, i.e. the walk is of the form $1 \to 5$.

Note that there is a subtle difference between the definitions of **path** and **walk**: the definition of a path implies that this is always a *graph* on its own, while a walk is defined

as a *sequence*. Nonetheless, we will treat *paths* as if they where *sequences* as well. This assumption holds for the following structures that will be discussed as well.

However, by definition of path, not every alternating sequence of vertices and edges is a valid path, in fact:

- in a *walk* it is possible to repeat both vertices and edges

- in a *path* there can be no repetition of vertices nor edges (note that *edge* repetition implies *vertex* repetition)

For instance, the previous example of *walk* is not a valid *path*, because the vertex 1 is repeated.

> **Theorem 1.1: Paths and walks**
>
> Given a graph $G = (V, E)$ and two vertices $x, y \in V$, in G there is a path $x \to y$ if and only if there is a walk $x \to y$.

Proof. By definition, every path is a walk, thus the direct implication is trivially true. To prove the converse implication, consider two vertices x and y for which there is at least one walk $x \to y$ in G. Now, out of all the possible walks $x \to y$ in G, consider the *shortest* one, i.e. the one with the least amount of edges, and let it be the following sequence

$$x \; e_1 \; x_1 \; \ldots \; x_{k-1} \; e_k \; y$$

By way of contradiction, assume that this walk is not a path. Therefore, there must be either one vertex or one edge repeated, but since edge repetition always implies vertex repetition, we just need to take this case into account. Assume that there are two indices $i, j \in [k-1]$ such that $i \neq j$ and $x_i = x_j$; however this implies that

$$x \; e_1 \; \ldots \; x_{i-1} \; e_i \; x_i \; e_{j+1} \; x_{j+1} \; \ldots \; x_{k-1} \; e_k \; y$$

is still a walk $x \to y$ of strictly shorter length, but we chose the original sequence to be the *shortest* possible walk $x \to y$ ⨎. □

> **Proposition 1.2**
>
> The longest path in any graph has a length of at least δ.

Proof. Consider a graph $G = (V, E)$, and let P be a longest path in G, labeled as follows

$$x_0 \; e_1 \; x_1 \; \ldots \; x_{k-1} \; e_k \; x_k$$

and assume that its length is k. Since P is a longest path in G, x_k cannot have neighbors outside P itself, otherwise P would not have been the longest path of G — it could have been extended by one of x_k's neighbors. This implies that

$$\mathcal{N}(x_k) \subseteq \{x_0, \ldots, x_{k-1}\}$$

and since $\delta \leq \deg(x_k) := |\mathcal{N}(x_k)|$ by definition of δ, this implies that

$$\delta \leq |\{x_0, \ldots, x_{k-1}\}| = k$$

$\square$

Definition 1.8: Cycle

A **cycle** is a *graph* with vertex set $x_1, \ldots, x_n$ and edge set $x_1 x_2, x_2 x_3, \ldots, x_{n-1} x_n, x_n x_1$.

The **length** of a cycle is the number of edges between x_1 and x_n, namely n in this case. A cycle of length n is denoted as C_n.
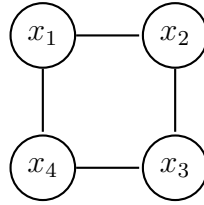
Figure 1.8: A cycle graph of length 4.

A graph that does not admit cycle subgraphs — or *cycles*, for short — is said to be **acyclic**.

Proposition 1.3

Every graph with $\delta \geq 2$ has a cycle of length at least $\delta + 1$.

Proof. Consider the proof of [Proposition 1.2](); by applying the same reasoning, we know that x_k cannot have neighbors outside P itself. However, since $\delta \geq 2$, and $x_k \sim x_{k-1}$, there must be at least one vertex in x_k's neighborhood that lies in P. Therefore, let x_i be the first vertex of P — w.r.t. our labeling of P — that is adjacent to x_k; hence, we have

$$\mathcal{N}(x_k) \subseteq \{x_i, \ldots, x_{k-1}\} \implies \delta \leq |\{x_i, \ldots, x_{k-1}\}|$$

which implies that $x_i, \ldots, x_{k-1}, x_k, x_i$ is a cycle of length at least $\delta + 1$. $\square$

Definition 1.9: Connected graph

An undirected graph $G = (V, E)$ is said to be **connected** if and only if for each vertex pair $x, y \in V$ there is a path $x \to y$.

All the graphs that we presented so far are *connected*, thus the following figure provides an example of an **disconnected** graph.
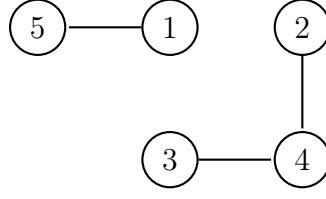
Definition 1.10: Component

Given a graph G, a **component** of G is a maximal connected subgraph of G.

For instance, the graph of the previous example is made up of 2 components, namely the following two subgraphs

$$C_1 = (\{1, 5\}, \{\{1, 5\}\})$$
$$C_2 = (\{2, 3, 4\}, \{\{2, 4\}, \{4, 3\}\})$$

Proposition 1.4

If G is a connected graph, and C is a cycle in G, then for any edge $e \in C$ it holds that $G - \{e\}$ is still connected.

Proof. Consider a graph $G = (V, E)$ that has a cycle C, and any two vertices $x, y \in V$; in particular, since G is connected, there must be a path $x \to y$ in G, and let this path be

$$P = x \ e_1 \ x_1 \ \ldots \ x_{k-1} \ e_k \ y$$

Consider an edge $e \in C$; if P does not traverse e, trivially $G - \{e\}$ will still contain P.

Now let

$$C = z_1 \ f_2 \ z_2 \ \ldots \ z_{l-1} \ f_l \ z_l \ f_{l+1} \ z_1$$

and without loss of generality assume that $e = f_2 = z_1 z_2 = x_i x_{i+1} = e_{i+1}$ for some $i \in [k-1]$. Thus we can construct the following walk

$$x \ e_1 \ x_1 \ \ldots \ x_i \ f_{l+1} \ z_l \ \ldots \ f_3 \ z_2 \ e_{i+2} \ x_{i+2} \ \ldots \ x_{k-1} \ e_k \ y$$

from x to y, and by Theorem 1.1 we have that there is a path from $x \to y$, which proves that $G - \{e\}$ is still connected. $\square$

1.2.1 Trees

Definition 1.11: Tree

A **tree** is a connected acyclic graph. Usually, but not necessarily, there is a fixed vertex called **root**, and any vertex that has degree 1 in the tree is called **leaf**.
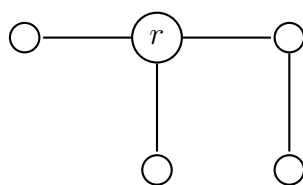
Figure 1.9: A tree with tree leaves, rooted in r.

A **forest** is an disconnected graph in which each component is a *tree*, as in the following example
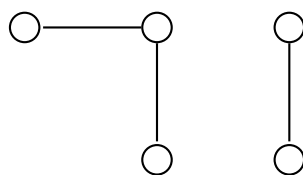
Figure 1.10: A forest.

Given a tree T rooted in some node $r \in V(T)$, and two vertices $x, y \in V(T)$, consider the paths P_x of the form $x \to r$ and P_y of the form $y \to r$, respectively. The first vertex of P_y that is encountered by tracing P_x from x to r is called **lowest common ancestor** (LCA) of x and y.
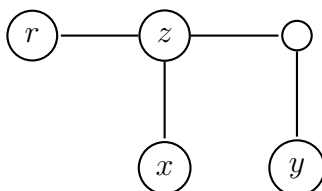
Figure 1.11: For instance, in this tree — rooted in r — the LCA of x and y is the vertex labeled with z.

Note that the LCA between any two vertices of a tree is *always defined*, since in the "worst case" it is the root r itself.

> **Theorem 1.2: Alternative definitions of tree**
>
> Given a graph $T = (V, E)$, the following statements are equivalent:
>
> 1. T is a tree
>
> 2. every vertex pair of T is connected by a unique path
>
> 3. T is *minimally connected*, i.e. T is connected and $\forall e \in E$ it holds that $T - \{e\}$ is disconnected
>
> 4. T is *maximally acyclic*, i.e. T is acyclic and $\forall x, y \in V$ such that $x \nsim y$, it holds that $T \cup \{xy\}$ has a cycle

Proof. We will prove the statements cyclically.

- 1 $\implies$ 2. By contrapositive, assume that in T there exist two vertices $x, y \in V$ for which there are two distinct paths P and Q of the form $x \to y$. If P and Q are edge-disjoint, then $P \cup Q$ is a cycle, which implies that T is not a tree by definition.

  Otherwise, assume that P and Q are not edge-disjoint. If we start say in x, and we follow Q edge by edge since P and Q are distinct, at some point we will encounter an edge $\{u, v\}$ such that $u \in P \cap Q$ and $v \in Q - P$ — possibly, $u = x$ itself. Moreover, since both paths lead to y, if we keep following Q we will encounter a vertex $z \in P \cap Q$ — possibly, $z = y$ itself — from which the two paths will coincide. Let Q' be the subpath of P starting with u and ending in z; then, $Q' \cup (Q - P)$ is a cycle in T, which implies that T is not a tree by definition.
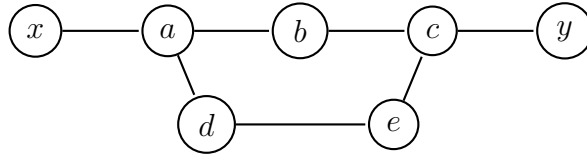

Figure 1.12: For instance, applying the argument of the proof in this graph we would get that $P = \{x, a, b, c, y\}$, $Q = \{x, a, d, e, c, y\}$, $Q - P = \{d, e\}$, $u = a$, $z = c$ and $Q' = \{a, b, c\}$, in fact $Q' \cup (Q - P) = \{a, b, c, e, d\}$ which is a cycle.

- 2 $\implies$ 3. Consider an edge $xy \in E$; this edge itself is a path $x \to y$, and if we assume statement 2 this implies that it is the *only* path from x to y. This implies that $T - \{xy\}$ cannot contain a path from x to y, therefore $T - \{xy\}$ is disconnected.

- 3 $\implies$ 4. Since statement 3 implies that T is connected, by Proposition 1.4 we have that T is acyclic. Now, pick $x, y \in V$ such that $x \nsim y$; by connectivity of T there must be a path $x \to y$ in T, and let this path be P. Lastly, since $x \nsim y$, we have that $P \cup \{xy\}$ is a cycle in T.

- 4 $\implies$ 1 By contrapositive, we want to prove that if T is not a tree, then T is not maximally acyclic. Note that if T is not a tree, we have two options:

  - if T is connected but contains a cycle, then T is clearly not maximally acyclic

  - if T is acyclic but disconnected, then by definition there must be two vertices $x, y \in V$ such that there is no path $x \to y$, which implies that $T \cup \{xy\}$ still does not contain any cycle

$\square$

Lemma 1.2

Every tree with at least 2 vertices has a leaf.

Proof. By way of contradiction, assume T is a tree with at least 2 vertices that does not contain any leaves; then $\delta \geq 2$ in T, which implies that T contains a cycle of length at least $\delta + 1$ by Proposition 1.3. $\square$

Lemma 1.3

Given a tree T, and a leaf v of T, it holds that $T[V(T) - \{v\}]$ is still a tree.

Proof. Since T is acyclic by definition, $T[V(T) - \{v\}]$ is still acyclic, we just need to prove that $T[V(T) - \{v\}]$ is still connected. By way of contradiction, assume that in $T[V(T) - \{v\}]$ there exist two vertices x and y such that there is no path between them. However, since T is connected, there is a path P of the form $x \to y$ in T.

Note that, if by removing v from T we disconnect x and y, it must be that v lies in P. Moreover, since v is in T but not in $T[V(T) - \{v\}]$, while both x and y are also in $T[V(T) - \{v\}]$, it must be that v is an *internal* node of P, i.e. $v \neq x, y$, which implies that $\deg(v) \geq 2$ by definition of path, contradicting the hypothesis for which v was a leaf $\natural$. $\square$

Proposition 1.5

If T is a tree, then $m = n - 1$.

Proof. We will prove the statement by induction on n

Base case. When $n = 1$, there are no edges in the tree, and $0 = m = 1 - 1$.

Inductive hypothesis. Assume that for a tree that has $n = k - 1$ nodes the statement holds.

Inductive step. We will prove the statement for a tree T that has $n = k$ nodes. Note that, since $n = 1$ is the base case, we can assume that $n = k \geq 2$, hence by Lemma 1.2 T contains at least one leaf. Let this leaf be v; then, by Lemma 1.3 it holds that $T[V(T) - \{v\}]$ is still a tree, and clearly $T[V(T) - \{v\}]$ has $k - 1$ nodes, which implies that we can apply the inductive hypothesis on $T[V(T) - \{v\}]$, i.e.

$$|E(T[V(T) - \{v\}])| = |V(T[V(T) - \{v\}])| - 1 = k - 1 - 1 = k - 2$$

However, note that v is a leaf, concluding that

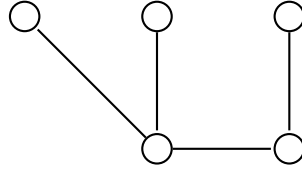$$m = |E(T)| = |E(T[V(T) - \{v\}])| + 1 = k - 2 + 1 = k - 1 = n - 1$$

$\square$

Definition 1.12: Spanning tree

Given a graph $G = (V, E)$, a **spanning tree** T of G is a subgraph of G such that

- T is a tree
- $V(T) = V(G)$, i.e. T *spans* every vertex of G

For instance, given the graph in Figure 1.3, a possible spanning tree is the following:

Lemma 1.4

Any connected graph has a spanning tree.

Proof. Consider a connected graph G, and keep removing edges from $E(G)$ — and their relative endpoints from $V(G)$ — one by one, as long as G is still connected. If no other edge can be removed from G without violating connectivity, we will end up with a graph that must be a tree by statement 3 of Theorem 1.2. $\square$

Thanks to this last proposition, we can actually prove a stronger version of the Proposition 1.5, which is the following.

Theorem 1.3

T is a tree if and only if T is connected and $m = n - 1$.

Proof. The direct implication is proved in Proposition 1.5, so we just need to prove the converse implication. Consider a connected graph T such that $m = n - 1$; by Lemma 1.4 T must have a spanning tree T', and by Proposition 1.5 itself it holds that $|E(T')| = |V(T')| - 1$. However, T' is a spanning tree of T, therefore

$$V(T) = V(T') \implies |V(T)| = |V(T')| \implies |E(T')| = |V(T)| - 1 = |E(T)|$$

which implies $E(T) = E(T')$ because T' is a subgraph of T, therefore $T = T'$, concluding that T must be a tree since T' is a tree. $\square$

1.2.2 Bipartite graphs

Definition 1.13: Bipartite graph

A graph $G = (V, E)$ is said to be **bipartite** if there exists a set $X \subseteq V$ such that every edge of G has exactly one endpoint in X and one in $V - X$. If such a set X exists, we say that $(X, V - x)$ is a **bipartition** of G. If G has the maximum number of edges that preserve the bipartition, we say that is **complete bipartite**.
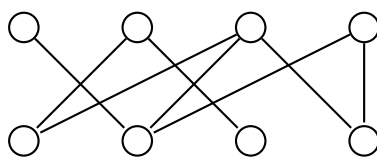
Figure 1.13: An example of a *bipartite graph*. In particular, if we call the uppermost set of nodes A and the lowermost one B, then (A, B) is a bipartition of the graph.

There are various types of graphs that can be bipartitioned. For example, any **tree** T can be bipartitioned through a bipartition $(X, V(T) - X)$ by considering the following set of vertices

$$X := \{v \in V(T) \mid \operatorname{dist}(r, v) \text{ even}\}$$

where r is T's root.
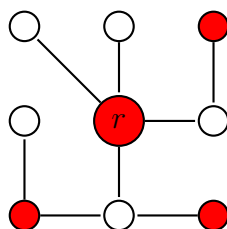
Figure 1.14: The set of red vertices X defines a bipartition $(X, V(T) - X)$ of this tree T.

However, *not every type* of graph can be bipartitioned. For instance, consider the following type.

> **Definition 1.14: Clique**
>
> A **clique** is a *graph* in which each vertex is adjacent to any other vertex of the graph. A clique that has n vertices is denoted as K_n.
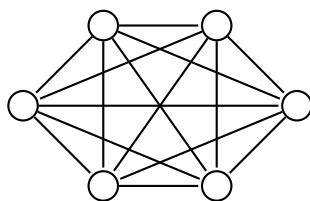
Figure 1.15: The clique K_6.

It is easy to see that no clique K_n can be bipartitioned, since there is an edge between any pair of vertices of the graph. However, this is not the only type of graph that cannot be bipartitioned.

Lemma 1.5

If G is a bipartite graph, and H is a subgraph of G, then H must be bipartite.

Proof. Given a bipartite graph $G = (V, E)$, assume that $(X, V - X)$ is a bipartition of G, and let H be a subgraph of G; then, it is easy to see that $(X \cap V(H), V(H) - X)$ is a bipartition for H. □

Note that this lemma implies that G is bipartite if and only if every connected component of G is bipartite: in fact, the direct implication follows from this lemma, and the following figure provides an intuition for the converse implication.
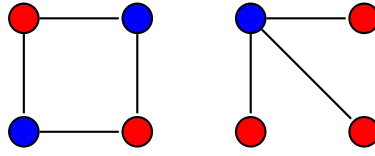
Figure 1.16: Consider the following disconnected graph G made of these two connected components, C_1 and C_2 respectively. Say that C_1 has a bipartition $(X_1, V(C_1) - X_1)$ where X_1 is the red set, and C_2 has a bipartition $(X_2, V(C_2) - X_2)$ where X_2 is the red set; thus, $(X_1 \cup X_2, V(C_1 \cup C_2) - X_1 - X_2)$ is clearly a bipartition of G. This process may be repeated for all the connected components of any disconnected graph.

Theorem 1.4

G is bipartite if and only if G has no odd-length cycle.

Proof.

Direct implication. We will prove the contrapositive, i.e. if G has an odd-length cycle, then G cannot be bipartitioned. Consider a graph G with an odd-length cycle C_{2k+1} of vertices $x_1, \ldots, x_{2k+1}$; by way of contradiction, assume that G is bipartite through a bipartition $(X, V(G) - X)$ for some $X \subseteq V(G)$. Without loss of generality assume that $x_1 \in X$; then, since X defines a bipartition of G it must be that $x_2 \notin X$, and $x_3 \in X$ and so on and so forth. In particular, for any odd value of i we will have that $x_i \in X$, but this implies that both x_1 and x_{2k+1} must be inside X, which means that the edge $x_1 x_{2k+1}$ violates the bipartition induced by X ↯.

Converse implication. Again, we will prove the contrapositive, i.e. if G is not bipartite it must contain an odd-length cycle. By the previous observation, G is not bipartite if and only if at least one connected component of G is not bipartite, and let this component be $\overline{G}$. Note that, since $\overline{G}$ is connected, it must contain a spanning tree T by Lemma 1.4. Moreover, as previously described, we can always define a bipartition on a tree, namely $(X, V(T) - X)$ where

$$X := \{v \in V(T) \mid \operatorname{dist}_T(r, v) \text{ even}\}$$

for some root node $r \in V(T)$.

Now, since T is a spanning tree of $\overline{G}$, which is now bipartite by hypothesis, there must be an edge $xy \in V(\overline{G})$ such that either $x, y \in X$ or $x, y \in V(T) - X$, i.e. the edge xy must have both endpoints in the same set of T's bipartition. Let z be the LCA between x and y in T, and P_x and P_y be the paths of the form $x \to r$ and $y \to r$, respectively. Note that, since xy has both endpoints in the same set, it must be that the lengths of P_x and P_y have the same parity by definition of X. Lastly, by statement 2 of Theorem 1.2 we have that $r \; P_x \; z = r \; P_y \; z$, which implies that the lengths of $z \; P_x \; x$ and $z \; P_y \; y$ must have the same parity. This concludes that

$$z \; P_x \; x \cup z \; P_y \; y \cup xy$$

is an odd-length cycle of G.

$\square$

1.2.3 Eulerian tours

At the start of this chapter, we introduced the *Seven Bridges of Königsberg* problem, which led to the emergence of graph theory as a branch of combinatorics. Over time, as the field developed, this problem was formalized into the following definition.

> **Definition 1.15: Eulerian tour**
>
> An **Eulerian tour** over a graph G is a closed walk that traverses every edge of G exactly once.

For instance, consider the following graph
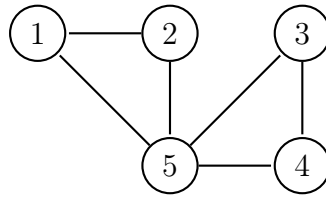
After some trial an error, it is easy to find an Eulerian tour over this graph, for instance

$$1 \; \{1, 2\} \; 2 \; \{2, 5\} \; 5 \; \{5, 3\} \; 3 \; \{3, 4\} \; 4 \; \{4, 5\} \; 5 \; \{5, 1\} \; 1$$

Note that this is a valid Eulerian tour because there is no edge repetition, and vertex repetition is allowed by definition. On the counter side, the graph — or, more precisely, the *multigraph* — that the bridges of Königsberg define, which is the following
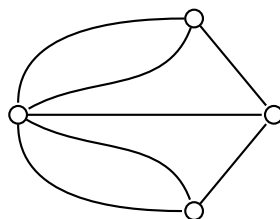
does not admit any Eulerian tour. But, given a graph, how can we determine with certainty whether it contains an Eulerian tour? The following theorem, proved by Euler himself in his original paper [Eul41], answers this question. Note that this theorem holds both for graphs and multigraph.

> **Theorem 1.5: Eulerian tours**
>
> A graph G admits an Eulerian tour if and only G is connected and every vertex of G has even degree.

Proof.

Direct implication. Consider the contrapositive of the direct implication, and assume that G contains at least one odd-degree vertex v. Recall that an Eulerian tour is a closed walk that does not allow edge repetition, which implies that the *starting* point of the walk is actually not relevant. Therefore, we can assume without loss of generality that any possible Eulerian tour defined on G starts on v itself, but then to be *closed* it must end on v as well. Moreover, each time any Eulerian tour passes through v it must use 2 distinct edges, however v is an odd-degree vertex, therefore at the end of the Eulerian tour there is no way we can come back to v and close the walk.

Converse implication. Consider a graph G in which every vertex has even degree, and let W be the *longest* walk of G that does not repeat edges, and let it be labeled as follows

$$x_0 \ e_1 \ x_1 \ \ldots \ x_{k-1} \ e_k \ x_k$$

Claim: $x_0 = x_k$, i.e. W is closed.

Proof of the Claim. By way of contradiction, assume that W $x_0 \neq x_k$. If this is the case, then $x_k = v$ for some vertex v that is not x_0. Note that W is a *walk*, therefore v may be repeated multiple times inside W, i.e.

$$x_0 \ e_1 \ x_1 \ \ldots \ v \ \ldots \ v \ \ldots \ x_{k-1} \ e_k \ v$$

If l is the number of times v occurs in W *without counting* x_k, then clearly in W there are $2l + 1$ edges incident to v, namely e_k and 2 edges each other time v appears in W. However, since $2l + 1$ is odd and we assumed that G has no odd-degree vertices, there must be at least one edge $vu \in E(G)$ such that $u \notin V(W)$. This implies that

$$x_0 \ e_1 \ x_1 \ \ldots \ x_{k-1} \ e_k \ v \ vu \ u$$

is a longer walk than W, and it does not repeat any edges because $uv \notin E(W)$, contradicting the definition of W $\lightning$. □

This claim proves that W is closed, but we still need to prove that it traverses every edge to claim that it is indeed an Eulerian tour. By way of contradiction, assume that there exists at least one edge $e \notin E(W)$ not used by W. Consider an vertex x_i of W; by connectivity of G, there must be a path P beween x_i and each of the

endpoints of e. Let $x_i u$ be the first edge of P, for some $u \notin E(W)$. This implies that

$$u \ ux_i \ x_i \ e_{i+1} \ x_{i+1} \quad \ldots \quad x_{k-1} \ e_k \ x_k \ e_1 \ x_1 \ \ldots \ x_{i-1} \ e_i \ x_i$$

is a longer walk than W that does not repeat any edge, since $ux_i \notin E(W)$, again contradicting the definition of W ⨳.

$\square$

Note that this theorem proves that it is not possible to describe an Eulerian tour over the bridges and landmasses of Königsberg, because every vertex of the multigraph has odd degree.

1.2.4 Hamiltonian cycles

Eulerian tours are closed walks that traverse every edge of the graph exactly once. But what if we are interested in traversing each *vertex* exactly once instead?

Definition 1.16: Hamiltonian paths and cycles

A **Hamiltonian path** over a graph G is a subgraph P_n of G. A **Hamiltonian cycle** over a graph G is a subgraph C_n of G.

Hamiltonian paths and cycles are named after W. R. Hamilton. Note that the notation P_n (or C_n) implies that the length of the path (or cycle) is n, hence this definition matches our requirements.

As for the case of Eulerian tours, when some conditions are met, Hamiltonian cycles are guaranteed to exist, as discussed in the following theorem.

Theorem 1.6: Hamiltonian cycles

A graph G such that $\delta \geq \frac{n}{2}$ contains a Hamiltonian cycle.

Proof. First, we will prove that the condition of the statement implies that G is connected.

Claim: G is connected.

Proof of the Claim. By way of contradiction, suppose that G is not connected; therefore G has at least two connected components. Let H be the smallest connected component of G; then, clearly $|V(H)| \leq \frac{n}{2}$. However, note that $|\mathcal{N}(x)| \geq \delta \geq \frac{n}{2}$ and since $\{x\} \cup \mathcal{N}(x) \subseteq V(H)$, we get that $V(H)$ must have at least $\frac{n}{2} + 1$ nodes ⨳. $\square$

Let P be the longest path of G, and let $x_0, \ldots, x_k$ be its vertices.

Claim: There exists an index ℓ such that $x_0 \ e_1 \ x_1 \ \ldots \ x_k \ e_\ell \ x_\ell \ e_{\ell+1} \ x_{\ell+1} \ e_0 \ x_0$ is a cycle.

Proof of the Claim. By the same argument used in the proof of Proposition 1.2, we

know that
$$\mathcal{N}(x_0), \mathcal{N}(x_k) \subseteq \{x_0, \ldots, x_k\}$$

Let I_0 and I_k be the following two sets

$$I_0 := \{i \mid i \in [1, k], x_i \in \mathcal{N}(x_0)\} \implies |I_0| = |\mathcal{N}(x_0)|$$

$$I_k := \{i \mid i \in [1, k], x_{i-1} \in \mathcal{N}(x_k)\} \implies |I_k| = |\mathcal{N}(x_k)|$$

Since $\delta \geq \frac{n}{2}$, we have that $|I_0|, |I_k| \geq \frac{n}{2}$. However, note that $k \leq n - 1$ — since we started counting at 0 — hence by the pigeonhole principle there must be at least one index $\ell \in I_0 \cap I_k$, meaning that $x_0 \sim x_\ell$ and $x_k \sim x_{\ell-1}$, defining a cycle as described in the statement of the claim. $\qquad \square$

This means that we found a cycle C in the graph that uses all the vertices of P, namely $x_0, \ldots, x_k$. By way of contradiction, assume that $k < n - 1$, i.e. C has less than n nodes, meaning that C is a non-Hamiltonian cycle. In particular, if $k < n - 1$, we have that $|V(G)| - |V(C)| \neq \varnothing$, thus let $y \in V(G) - V(C)$. By the previous claim, we know that G is connected, there must be an edge $xy \in E(G)$ such that $x \in V(C)$. However, this would imply that $P \cup \{xy\}$ is a path of longer path than P ↯. $\qquad \square$

Note that the statement of this theorem cannot be improved, even by 1; for instance consider the following graph
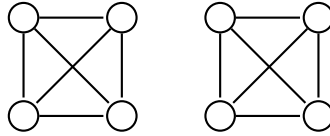
composed of two disconnected K_4. Here, we have that

$$\delta = 4 - 1 = 3 \geq \frac{2 \cdot 4}{2} - 1 = 4 - 1 = 3 = \frac{n}{2} - 1$$

Therefore, $\delta \geq \frac{n}{2} - 1$ is not sufficient to guarantee connectivity.

1.3 Exercises

Problem 1.1

Let $G = (V, E)$ be a graph of n vertices, where $n \geq 2$. Show that there must be two vertices $x, y \in V$ such that $\deg(x) = \deg(y)$.

Solution. By definition, the range of the possible degrees for any node of G is $[0, n - 1]$. By way of contradiction, assume that for any two vertices $x, y \in V$ it holds that $\deg(x) \neq \deg(y)$; hence, since the graph has n nodes, it must be that each node is assigned a different degree, and that we use all the possible degrees in $[0, n - 1]$. In particular, this implies that there are two vertices $u, v \in V$ such that $\deg(u) = 0$ and $\deg(v) = n - 1$, but this is a contradiction because if the degree of v is $n - 1$, it must be adjacent to all the other nodes of V, including u, and $\deg(u) = 0$ ↯.

2

Matchings

Given a graph $G = (V, E)$, a **matching** of G is a set of edges $M \subseteq E$ such that

$$\forall e, e' \in M \quad e \cap e' = \varnothing$$

Figure 2.1: A *matching* of the previous graph.

As shown in figure, a matching is nothing more than a set of edges that must not share endpoints with each other — for this reason, in literature it is often referred to as **independent edge set**.

Given a matching M in a graph G, we say that a vertex $v \in V(G)$ is **free** w.r.t. M if there are no edges $e \in M$ such that $e \cap v \neq \varnothing$. An edge $xy \in E(G)$ is said to be *disjoint* from M if both x and y are free w.r.t. M.

In graph theory we are often interested in the matching that has the largest possible cardinality of a graph. For this purpose, we often distinguish the two following concepts, namely *maximal* and *maximum* matching.

Definition 2.2: Maximal matching

A **maximal matching** is a matching that cannot be extended any further.

For instance, the matching shown in Figure 2.1 is actually a **maximal matching**, because

no other edge in E can be added to che current set of edges M of the matching without breaking the matching condition.

> **Definition 2.3: Maximum matching**
>
> A **maximum matching** is a matching that has the largest cardinality.

Clearly, the previous example does not repreent a **maximum matching**, because the following set of edges

is still a valid matching for the graph, but has a larger cardinality than the previous set.

Given a matching M in a graph G, what conditions must be met to increase its cardinality? Trivially, if there exists an edge e disjoint from M, clearly $M \cup \{e\}$ is a larger matching than M — implying that M was not *maximal* in G. However, this is not the only situation in which the cardinality of M can be extended. In fact, consider the following definitions.

2.1 Augmenting paths

> **Definition 2.4: Alternating path**
>
> Given a graph G, and a matching M in G, an M-**alternating path** is a path of G that starts at a free node w.r.t. M, and is composed of edges that alternate between M and $E(G) - M$.

Figure 2.2: For instance, if M is the set of *red* edges — which forms a matching of the graph — then 6 $\{6, 2\}$ 2 $\{2, 7\}$ 7 $\{7, 3\}$ 3 $\{3, 8\}$ 8 is an M-alternating path.

> **Definition 2.5: Augmenting path**
>
> Given a graph G, and a matching M in G, an M-**augmenting path** is an M-alternating path that ends at a free vertex w.r.t. M.

For example, if we consider the path

$$6 \; \{6, 2\} \; 2 \; \{2, 7\} \; 7 \; \{7, 3\} \; 3 \; \{3, 8\} \; 8 \; \{8, 4\} \; 4$$

this is actually an M-augmenting path of the previous graph. Augmenting paths are very useful because they can be used to *expand* the cardinality of an initial matching. In fact, in the previous graph we can actually define a *larger* matching by **swapping** the edges of this augmenting path, as shown below

2.1.1 Berge's theorem

This suggests that the presence of augmenting paths in a graph is a *sufficient* condition for a matching *not* to be maximum, but we can actually prove that it is also *necessary*, as stated in the following theorem, proved by Berge [Ber57] in 1957.

> **Theorem 2.1: Berge's theorem**
>
> Given a graph G, M is a maximum matching of G if and only if there are no M-augmenting paths.

Proof.

Direct implication. By contrapositive, consider a graph G and a matching M such that there is an M-augmenting path P in G. Moreover, by way of contradiction assume that M is maximum; however $M\Delta E(P)$ is a larger matching than M — here, Δ is the symmetric difference, therefore the operation $M\Delta E(P)$ has the same effect of *swapping* the edges of P between the ones in M and in $E(P) - M$.

Converse implication. By contrapositive, consider a graph G and a matching M of G that is not maximum, i.e. there exists a matching M^* of G such that $|M| < |M^*|$. Consider the subgraph of G that has the vertices of $V(G)$ and the edges described by $M\Delta M^*$; the symmetric difference of these two sets will yield the set of edges that are either in M or in M^*, but not in $M \cap M^*$, therefore this subgraph is *not* a multigraph. Moreover, since M and M^* are both matchings, we have that

(1) the degrees of the vertices of this subgraph can be either 0, 1 or 2

(2) in each component of the subgraph the edges *must* alternate between M and M^*

By the observation (1), we have that each component of the subgraph can be either

- a isolated vertex

- a cycle

- a path

and by observation (2), we have that all the cycle components must have *even* length, which implies that they have the same number of edges of M and M^*. On the other hand, path components may have either even or odd length; in particular, even-length paths must have the same number of edges of M and M^* — as for cycle components — while odd-length paths have a different number of edges of M and M^*. However, since $|M| < |M^*|$, there must be at least one path component of this subgraph such that its edges of M are less than the edges of M^*, and this is clearly an M-augmenting path.

$\square$

Given a graph G, and a matching M of G, what is the maximum possible value for $|M|$? To answer this question, we need to introduce the following combinatorial structure.

2.1.2 Kőnig's theorem

> **Definition 2.6: Vertex cover**
>
> Given a graph G, a **vertex cover** for G is a set of vertices $C \subseteq V(G)$ such that every edge in G is incident to at least one vertex in C. Using symbols
>
> $$\forall (u, v) \in E(G) \quad u \in C \lor v \in C$$
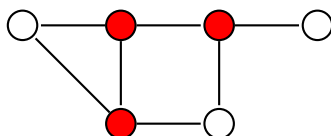
Figure 2.3: An example of a vertex cover.

As shown in figure, a vertex cover is simply a set of vertices that must *cover* all the edges of the graph. For vertex covers, we are interested in the *minimum* possible cardinality — the concepts of *minimal* and *minimum* are defined analogously.

As introduced before, through vertex covers we can bound the size of any matching of a graph.

> **Theorem 2.2: Matchings bound vertex covers**
>
> Given a graph G, a matching M, and a vertex cover S of G, it holds that $|M| \leq |S|$.

Proof. By definition, any vertex cover S of $G = (V, E)$ is also a vertex cover for $G^B = (V, B)$, for any set of edges $B \subseteq E$, and in particular this is true for $G^M = (V, M)$.

Now consider G^M, and a vertex cover C on it: by construction we have that $\Delta \leq 1$, therefore any vertex in C will cover at most 1 edge of M. This implies that if $|C| = k$, then C will cover at most k edges of G^M.

Lastly, since G^M has $|M|$ edges by definition, any vertex cover defined on G^M has to contain at least $|M|$ vertices. This implies that no vertex cover S of G smaller than $|M|$ can exist, because S will have to cover at least the edges in M. $\square$

> **Corollary 2.1**
>
> Given a graph G, a maximum matching M^* and a minimum vertex cover S^*, it holds that $|M^*| \leq |S^*|$.

Moreover, if the graph is bipartite this theorem is actually *stronger*, as proved by Konig [Kon] in 1931.

> **Theorem 2.3: Kőnig's theorem**
>
> Given a bipartite graph G, a maximum matching M^* and a minimum vertex cover S^*, it holds that $|M^*| = |S^*|$.

Proof. Consider a graph G, a maximum matching M^* and a minimum vertex cover S^* of G; by the previous corollary, it follows that to prove the statement it suffices to show that there exists a vertex cover S such that $|S| = |M^*|$, because

$$|M^*| \leq |S^*| \leq |S| = |M^*| \implies |M^*| = |S^*|$$

Hence, we are going to construct the following vertex cover. Let G be bipartitioned through (A, B); then, for each edge $ab \in M^*$ such that $a \in A$ and $b \in B$, we place $b \in S$ if and only if there exists an M^*-alternating path that starts in A and ends at b, otherwise we place $a \in S$.

Figure 2.4: For instance, given this graph bipartitioned into (A, B) — where A is the uppermost row of vertices — and the *red* matching, we would construct the *green* vertex cover.

Note that, by definition, it holds that $|S| = |M^*|$.

Claim: S is a vertex cover for G.

Proof of the Claim. Consider an edge $ab \in E(G)$ such that $a \in A$ and $b \in B$; note

that, by definition of S all the edges in M^* are already covered, hence we may assume that $ab \notin M^*$. We have two cases.

- a is free, i.e. $\nexists ab' \in M^*$ for $b' \in B$. Note that b cannot be free, otherwise $M^* \cup \{ab\}$ would still be a matching of G but greater than M^*. Hence, there must be an edge $a'b \in M^*$ for some $a' \in A$. Therefore, since a is free, the edge ab is a trivial M^*-alternating path ending at $b \in B$, meaning that $b \in S$ by definition, implying that ab is covered by S.

- a is matched, i.e. $\exists ab' \in M^*$ for $b' \in B$. Therefore, by definition of S, either a or b' lies in S. In particular, if $a \in S$, then ab is trivially covered by S, hence suppose that $b' \in S$.

  Observe that, by definition of S this implies that there must be an M^*-alternating path P that starts in A and ends at b'.

  - If $ab, ab' \notin E(P)$, then $P \cup \{b'a\} \cup ab$ is an M^*-augmenting path, which would contradict the fact that M^* is maximum by [Theorem 2.1](#) $\lightning$

  - If $ab \notin E(P)$ but $ab' \in E(P)$, then P could not have been an M^*-alternating path starting at a vertex in A $\lightning$.

  - If $ab \in E(P)$ but $ab' \notin E(P)$, then P must have had the following form

    $$\ldots\; b''\; \{b'', a\}\; a\; \{a, b\}\; b\; \{b, a'\}\; a'\; \{a', b'\}\; b'$$

    where $a' \in A$, $b'' \in B$. However, since $ab' \in M^*$ and $ab \notin M^*$, and the edges of P must alternate w.r.t. M^*, ab'' must lie inside M^*, contradicting $ab' \in M^*$ by definition of matching $\lightning$.

  This implies that $ab', ab \in E(P)$, meaning that P encounters b "before" b'. However, this implies that $P - \{ab'\}$ is an M^*-alternating path that starts in A and ends at b, thus $b \in S$ by definition, concluding that ab is still covered by S.

  $\square$

Hence, we have constructed a vertex cover S such that $|S| = |M^*|$, meaning that the statement holds because of the previous observation. $\square$

2.1.3 Finding maximum matchings

Consider a matching M of a graph G, and an M-augmenting path P; the idea of *swapping* the edges of P between M and $E(G) - M$ is very useful when G is **bipartite**. In fact, we can actually describe a procedure which is able to return a *maximum matching* of a bipartite graph, by swapping the edges of the augmenting paths present in G. However, for this algorithm to work, we first need a procedure capable of finding augmenting paths in bipartite graphs, which is defined down below:

1. Assume that the considered graph G is bipartite through (A, B), and consider a matching M of G

2. Starting from a node $a \in A$ free w.r.t. M, compute a *modified* BFS such that the edges of its tree alternate between $E(G) - M$ and M

3. If the tree of the BFS contains a free leaf $b \in B$, then the path $v \to b$ is M-augmenting

For instance, given the following bipartite graph G, and a matching M of G — outlined in *red*

the *modified* BFS rooted in a would produce the following tree
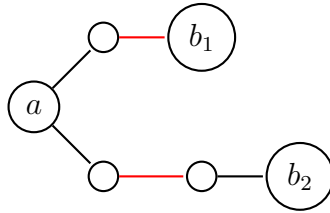
and we observe that the path $a \to b_1$ is M-alternating, while the path $a \to b_2$ is M-augmenting. The next proposition guarantees that if there are M-augmenting paths that start in a, our *modified* BFS will find at least one of them.

> **Proposition 2.1**
>
> Given a bipartite graph G, bipartitioned into (A, B), and a matching M of G, if there exists an M-augmenting path in G that starts in a vertex $a \in A$ free w.r.t. M, then there exists an M-augmenting path in the tree T of the *modified* BFS.

Proof. Let P be an M-augmenting path that starts in a and minimizes the edges in $E(P) - E(T)$ and, by way of contradiction, assume that $E(P) - E(T) \neq \varnothing$, i.e. P is not completely contained in T. Therefore, let xy be the first edge in $E(P) - E(T)$ encountered while traversing P, starting at a, and without loss of generality assume that $x \in V(P)$.

Claim: $x \in A$.

Proof of the Claim. By way of contradiction, assume that $x \in B$.

- Assume that x is not a leaf of T. Since the BFS starts at $a \in A$, and the edges of T alternate between M and $E(G) - M$, if $x \in B$ then the next edge xy' in T is an edge in M. Moreover — by the same reasoning — since P is M-augmenting, it must be that $xy \in M$, meaning that $xy, xy' \in M$ contradicting the definition of matching ↯

- Now, assume that x is a leaf of T. By the same reasoning, xy must be in M because P is M-augmenting, but $xy \notin E(T)$ would imply that the BFS stopped before adding the edge xy to T ↯.

$\square$

For instance, given the following setting

a possible path for P would be $a \to x \to y \to z$. However, the path $a \to y \to z$ is still an M-augmenting path that starts at a but has one fewer edge not in T w.r.t. P, contradicting the definition of $P \nmid$.

Note that, in the general case we would consider the path $P' := a \ T \ y \cup y \ P$, however this is not guaranteed to be a path. The complete proof leverages the fact that G is bipartite in order to prove that P' is indeed a path, but it is very technical and outside the scope of these notes. $\square$

Finally, now that we have a procedure which is guaranteed to find an augmenting path in a given bipartite graph, to return a maximum matching it suffices to run the following algorithm.

Algorithm 2.1: Maximum matching (bipartite graphs)

Given a bipartite graph G, the algorithm returns a maximum matching for G.

1: **function** MAXIMUMMATCHINGBIPGRAPHS(G)
2:     $M := \varnothing$
3:     **do**
4:         $P :=$ FINDAUGMENTINGPATH(G) $\triangleright$ the previous procedure
5:         Swap the edges between M and $E(G) - M$ in P
6:     **while** $P \neq$ None
7:     **return** M
8: **end function**

In particular, the output of this algorithm is guaranteed to be a *maximum* matching thanks to Theorem 2.1, since the algorithm terminates when there are no more augmenting paths left in the graph G.

2.2   Perfect matching

By definition, a matching is *not* forced to cover all the vertices of a graph. However, if this happens the matching is called **perfect matching**.

> **Definition 2.7: Perfect matching**
>
> Given a graph G, a **perfect matching** of G is a matching that covers all the vertices of G, i.e. M is a perfect matching if and only if
>
> $$\forall v \in V(G) \quad \exists e \in M \quad v \cap e \neq \varnothing$$
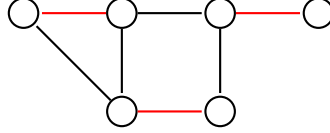
Figure 2.5: An example of a perfect matching.

Perfect matchings are an interesting topic of study when related to *bipartite graphs*. We observe that if a bipartite graph G, bipartitioned into (A, B), admits a perfect matching M, it must be that $|A| = |B|$. This is because every matched edge must connect one vertex from A to one vertex from B, and by definition M matched *each* vertex exactly once. However, the converse is not true.
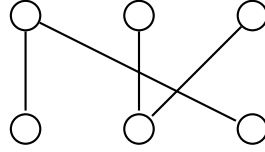
Figure 2.6: For instance, this bipartite graph, bipartitioned into (A, B) such that A is the uppermost row of vertices, even if $|A| = |B|$ this graph does not admit a perfect matching.

2.2.1 Hall's theorem

Because of this characterization on bipartite graphs, in addition to the concept of *perfect matching*, if $|A| \neq |B|$ we can define a *weaker* version of "perfect".

> **Definition 2.8: A-perfect matching**
>
> Given a bipartite graph G, bipartitioned through (A, B) such that $|A| \leq |B|$, we say that a matching M is an A-**perfect matching** if it covers all the vertices in A.

Note that we can always assume that $|A| \leq |B|$, without loss of generality. The following theorem, known as the Hall's marriage theorem — proved by Hall [Hal35] in 1935 — shows the conditions that guarantee that an A-perfect matching exists.

> **Theorem 2.4: Hall's marriage theorem**
>
> Given a bipartite graph G, bipartitioned into (A, B), then G admits an A-perfect matching if and only if
> $$\forall S \subseteq A \quad |S| \leq |\mathcal{N}(S)|$$

Proof. The direct implication is trivially true by definition of matching. We will prove the converse implication by contrapositive. Therefore, suppose that the bipartite graph G does not admit any A-perfect matching, and consider a minimum vertex cover V^* of G. Then, by Theorem 2.3, since G is bipartite we know that for any maximum matching M^* of G it holds that $|M^*| = |V^*|$. However, since we are assuming that there are no A-perfect matchings of G, any maximum matching must have size strictly less than $|A|$, meaning that $|V^*| = |M^*| < |A|$.

Now, consider the set $S := A - V^*$; since G is bipartite, it must hold that
$$\mathcal{N}(S) \subseteq V^* \cap B \implies |\mathcal{N}(S)| \leq |V^* \cap B| = |V^*| - |A \cap V^*|$$

Moreover, observe that
$$V^* \cap A = A - (A - V^*) = A - S \implies |V^* \cap A| = |A| - |S|$$

therefore, we have that
$$|\mathcal{N}(S)| \leq |V^*| - |V^* \cap A| = |V^*| - |A| + |S|$$

Lastly, since $|V^*| < |A| \iff |V^*| - |A| < 0$, we conclude that
$$|\mathcal{N}(S)| \leq |V^*| - |A| + |S| < 0 + |S| = |S| \implies |\mathcal{N}(S)| < |S|$$

which proves that there is at least one set S for which the *Hall condition* does not hold. □

2.2.2 Tutte's theorem

In the general case, what obstructs the existence of a perfect matching in a graph? Consider a graph G; clearly, if n is odd, the graph does not admit perfect matchings, since each edge matches exactly two nodes, meaning that there will always be a free vertex in the graph.

Figure 2.7: For instance, no matching of C_5 can be perfect.

For the same reasoning, if G has a connected component with an odd number of vertices, G does not admit perfect matchings. Hence, if G admits a perfect matching, there cannot

be any connected component with an odd number of vertices. But is the converse true as well? Consider the following graph

This graph is connected, and has 10 vertices, but it does not admit perfect matchings, meaning that the converse does not hold. Can we find a condition that guarantees that a given graph always admits a perfect matching?

Given a graph G, let $\mathcal{O}(G)$ be the number of components with an odd number of vertices. The next theorem, proved by Lovász and Plummer [LP09] in 1947 shows that the following property

$$\forall S \subseteq V(G) \quad \mathcal{O}(G[V(G) - S]) \leq |S|$$

which we will refer to as **Tutte condition** — is a *necessary* and *sufficient* condition to guarantee that a graph admits a perfect matching. In fact, in the example of C_5 we observe that the set that violates the Tutte condition is $S = \varnothing$, and in the second example is $S = \{x\}$.

Lemma 2.1

Given a graph G, and a supergraph G' of G, if G satisfies the Tutte condition, then G' satisfies the Tutte condition as well.

Proof. By contrapositive, suppose that G' fails the Tutte condition, meaning that there exists a set $S \subseteq V(G') = V(G)$ such that $\mathcal{O}(G[V(G') - S]) > |S|$. Since G' is obtained by adding edges to G, the number of odd components in G' can either remain the same, or decrease if two different components having an odd number of vertices became connected in G'. Therefore, we have that

$$|S| < \mathcal{O}(G[V(G') - S]) \leq \mathcal{O}(G[V(G) - S])$$

meaning that G fails the Tutte condition as well. □

We are now ready to prove Tutte's theorem.

Theorem 2.5: Tutte's theorem

A graph G admits a perfect matching if and only if for any $S \subseteq V(G)$ it holds that $\mathcal{O}(G[V(G) - S]) \leq |S|$ — meaning that it satisfies the Tutte condition.

Proof.

Direct implication. For the direct implication, consider a graph G that admits a perfect matching M, and by way of contradiction assume that there exists a set $S \subseteq V(G)$ that violates the Tutte condition. For the previous observation, we know that each component of $G[V(G) - S]$ that has an even number of vertices will not have free nodes w.r.t. M, and each component of $G[V(G) - S]$ that has an odd number of vertices will have one free node w.r.t. M. In particular, these free vertices must be covered by M, since M is perfect, and the only way they can be covered by M is through the vertices of S. However, since $\mathcal{O}(G[V(G) - S]) > |S|$, there exists at least one free vertex in $G[V(G) - S]$ w.r.t. M that cannot be matched to any of the vertices of S, meaning that M is not perfect $\lightning$.

Converse implication. We will prove the contrapositive of the converse implication. Therefore, suppose that G does not admit any perfect matching, and we will prove that there exists a set of vertices for which the Tutte condition does not hold. Let G' be the maximal supergraph of G that still does not admit perfect matchings. We say that a set X satisfies the *clique-adjacency* condition if every component of $G'[V(G') - X]$ is a clique, and every vertex $u \in X$ is adjacent to every vertex $v \notin X$.

Claim: If there is a subset $S' \subseteq V(G')$ that satisfies the clique-adjacency condition, then G does not satisfy the Tutte condition.

Proof of the Claim. Note that $S' \subseteq V(G') = V(G)$, hence if S' violates the Tutte condition on G, the claim is trivially true. Therefore, we will assume that S' satisfies both the clique-adjacency condition, and the Tutte condition on G.

Let S' be a subset that satisfies the clique-adjacency condition; hence, by definition every component of $G'[V(G') - S']$ is a clique. Thus, consider a clique, and let n be the number of its vertices.

- If n is even, we can always find a perfect matching restricted on it.

- If n is odd, we can always find a perfect matching restricted to $n - 1$ vertices, but since we are assuming that S' satisfies the Tutte condition on G, we know that $\mathcal{O}(G[V(G) - S']) \leq |S'|$. Therefore, by the clique-adjacency condition we know that we can always match the n-th vertex of the clique with a vertex of S'.

TODO ── □ da finire

Consider the set
$$S := \{v \in V(G') \mid \deg_{G'}(v) = n - 1\}$$

By way of contradiction, suppose that S violates the clique-adjacency condition. This happens if at least one of the following holds:

(1) there are two vertices $u \in S$ and $v \notin S$ such that $u \notin S$

(2) there is a component of $G'[V(G') - S]$ that is not a clique

However, by definition of S, each vertex $v \in S$ has $\deg_{G'} = n - 1$, hence (1) cannot be true, meaning that (2) must hold. Hence, consider a connected component of

$G'[V(G') - S]$ that is not a clique, i.e. there are at least two vertices x and y such that $x \nsim y$.

Let P be the shortest path from x to y in $G'[V(G') - S]$, and let a, b and c be the first three vertices of P — note that $a, b, c \notin S$. We observe that $a \nsim c$, otherwise P would not be the shortest path. Note that $b \notin S$, therefore $\deg_{G'}(b) < n - 1$, i.e. there exists a vertex d such that $b \nsim d$.
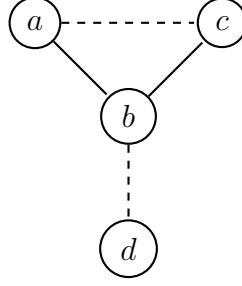

Figure 2.8: The "kite-shaped" figure we are considering — the dashed lines represent the *missing edges.*

By maximality of G', we know that $G' \cup \{ac\}$ must have a perfect matching M_1, and $G' \cup \{bd\}$ must have a perfect matching M_2. Moreover, $ac \in M_1$ and $bd \in M_2$, otherwise M_1 and M_2 would be perfect matchings on G'.

Claim: $(M_1 - \{ac\}) \cup (M_2 - \{bd\})$ contains a perfect matching for G'.

Proof of the Claim. Let $\overline{M} = (M_1 - \{ac\}) \Delta (M_2 - \{bd\})$. By the same reasoning applied for Theorem 2.1, since M_1 and M_2 are two matchings, every vertex of $\overline{M}$ has degree at most 2. Hence, consider a vertex z in the subgraph induced by $\overline{M}$

- if $\deg_{\overline{M}}(z)$ is 0 or 2, z is matched both by M_1 and M_2

- since M_1 is a perfect matching, each vertex of G' is matched by $M_1 - \{ac\}$, except for a and c

- likewise, since M_2 is a perfect matching, each vertex of G' is matched by $M_2 - \{bd\}$, except for b and d

which implies that the only vertices z that have $\deg_{\overline{M}}(z) = 1$ are precisely a, b, c and d.

Moreover, each component of the subgraph induced by $\overline{M}$ is either an isolated vertex, a cycle or a path, and the edges of the components must alternate between M_1 and M_2. Therefore, because of the degrees of a, b, c and d, it must be that these vertices are endpoints of two alternating paths P_1 and P_2. Moreover, since $ac \in M_1$ and $bd \in M_2$, one of these paths must be an M_1-alternating path, while other must be M_2 alternating — and without loss of generality let P_1 be the M_1-alternating. Note that the set

$$M' := (M_1 \cup M_2) - (E(P_1) \cup E(P_2))$$

is a perfect matching on $G[V(G) - (V(P_1) \cup V(P_2))]$. We have three cases for P_1 and P_2:

(a) P_1 is a path $a \to c$, and P_2 is a path $b \to d$; then

$$(M_1 \cap E(P_2)) \cup (M_2 \cap E(P_1)) \cup M'$$

is a perfect matching on G'

(b) P_1 is a path $a \to b$, and P_2 is a path $c \to d$; then

$$(M_1 \cap E(P_2)) \cup (M_2 \cap E(P_1)) \cup M' \cup \{bd\}$$

is a perfect matching on G'

(c) P_1 is a path $a \to d$, and P_2 is a path $b \to c$; then

$$(M_1 \cap E(P_2)) \cup (M_2 \cap E(P_1)) \cup M' \cup \{ac\}$$

is a perfect matching on G'

$\square$

Therefore, because of this claim G' always contains a perfect matching, contradicting the definition of G'. This implies that S must satisfy the clique-adjacency condition. Hence, by the first claim we have that G satisfies the Tutte condition

> first claim is wrong?

$\square$

2.3 Stable matching

Matchings in bipartite graphs are particularly useful as they can be applied to model various types of problems across different fields. One well-known example is the stable matching problem, which arises in scenarios like job assignments, college admissions, and matchmaking systems. In this problem, the goal is to find a *stable pairing* between two sets of entities — such as students and universities — where no two unmatched entities would prefer each other over their current assignments.

> **Definition 2.9: Stable matching**
>
> Given a graph G, bipartitioned through (A, B), and a matching M of G, consider a family of *preference* functions $\{w_v\}_{v \in V(G)}$ that for each vertex $v \in V(G)$, assign a value to the edges $vu \in E(G)$, for all $u \in \mathcal{N}(v)$
>
> $$\forall v \in V(G) \quad w_v : \mathcal{N}(v) \to \mathbb{R}$$
>
> We say that M is **stable** if for each $ab \in E(G)$ it does <u>not</u> happen that
>
> $$(a \text{ free } \vee (\exists ab' \in M \quad w_a(b) > w_a(b')))$$
>
> $$\wedge$$
>
> $$(b \text{ free } \vee (\exists a'b \in M \quad w_b(a) > w_b(a')))$$

For instance, consider the following bipartite graph G, and a matching M — outlined in *red*:

If the family of preference functions $\{w_v\}_{v \in V}$ is such that

$$w_a(b) > w_b(b') \land w_b(a) > w_b(a')$$

which implies that <u>both</u> a and b would prefer to *switch* their current matched vertex — then M is *not* stable. Note that the empty matching is *not* stable, by definition.

The following theorem, proved by Gale and Shapley [GS62] in 1962, proves that a stable matching can be always constructed in a bipartite graph, regardless of the preference functions.

Theorem 2.6: Gale-Shapley theorem

Given a bipartite graph G, and a family of preference functions $\{w_v\}_{v \in V(G)}$, there exists a stable matching of G.

Proof. Before proving the theorem, we need to introduce some definitions.

Consider a bipartite graph G, bipartitioned through (A, B), and consider a matching M. Given two vertices $a \in A$ and $b \in B$, we say that a is **acceptable to** b if

- b is free, or
- b is matched to a vertex a', and $w_b(a) > w_b(a')$

Moreover, we say that a vertex $a \in A$ is **happy** if either

- a is free, or
- $ab \in M$ and for each b' such that a is acceptable to b', it holds that $w_a(b) \geq w_a(b')$

Observe that in the empty matching every vertex is happy.

Claim: Consider a matching M such that each vertex is happy; if M is not stable, then there exists a vertex $a \in A$ such that a is free and acceptable to some $b \in B$.

Proof of the Claim. By instability of M, there must be an edge $ab \notin M$ such that a is free, or it prefers b to its current partner, and vice versa. By way of contradiction, suppose that a is matched to some $b' \in B$, hence by instability of M through ab we know that $w_a(b) > w_a(b')$

- If b is free, then by definition a is acceptable to b; however, by happiness of a it must be that $w_a(b') \geq w_a(b)$ ↯

- If b is matched by some edge $a'b \in M$, by instability of M through ab we know that $w_b(a) > w_b(a')$, hence a is acceptable to b, and by happiness of a it must be that $w_a(b') \geq w_a(b)$ ↯

This implies that a must be free; therefore, we have that

- If b is free as well, then by definition a is acceptable to b

- If b is matched by some edge $a'b \in M$, by instability of M through ab we know that $w_b(a) > w_b(a')$, hence a is still acceptable to b

$\square$

Now, consider another matching M' of G; we say that M is **better than** M' if

- $\forall a'b \in M' \quad \exists ab \in M \quad w_b(a) \geq w_b(a')$, meaning that every vertex $b \in B$ prefers its match in M at least as much as its match in M', and

- $\exists a'b \in M', ab \in M \quad w_b(a) > w_b(a')$, meaning that there is at least one vertex b that strictly perfers its match in M over its match in M'

In other words, M is better than M' if no match in M is worse than in M', and at least one match is strictly better.

Let M_k be an unstable matching such that every vertex is happy; therefore, by the previous claim we know that there exists a vertex $a \in A$ such that a is free and and acceptable to some $b \in B$. We will construct a matching M_{k+1} as follows:

$$b^* \in \underset{\substack{b \in \mathcal{N}(a): \\ a \text{ acceptable to } b}}{\arg\max} w_a(b) \implies M_{k+1} := (M_k \cup \{ab^*\}) - \{a'b^* \in M_k \mid a' \in A\}$$

meaning that M_{k+1} is obtained from M_k by adding the edge ab^*, where b^* maximizes $w_a(b^*)$, and removing the edge $a'b^*$ from M_k, if present — the last set is either $\{a'b^*\}$ or $\varnothing$.

Claim: M_{k+1} is better than M_k, and if M_k is ensures that every vertex is happy, M_{k+1} does as well.

Proof of the Claim. First, we prove that M_{k+1} is better than M_k. The first condition that M_{k+1} has to satisfy in order to be better than M_k is true simply because we removed the edge $a'b^*$ if it was present in M_k, and the rest of the matching has not been altered. Moreover, if b^* was free then the second condition is vacuously true, otherwise if the edge $a'b^*$ was present in M_k, in M_{k+1} there we added the edge ab^* and we know that $w_{b^*}(a) > w_{b^*}(a')$ because a is acceptable to b^* by definition.

Now, assume that M_k is such that every vertex is happy. Since b^* is the neighbor of a that maximizes $w_a(b^*)$, a is happy by definition. Now, if b^* was free in M_k, then we only added ab^* to M_{k+1}, hence every vertex is happy w.r.t. M_{k+1}. Otherwise, if b^* was not free in M_k, i.e. there was an edge $a'b^* \in M_k$, by definition M_{k+1} will not contain $a'b^*$, meaning that a' is free w.r.t. M_{k+1}, hence a' is happy by definition. $\square$

Now, consider the empty matching $M_0 := \varnothing$; we already observed that M_0 is not stable, but is such that each vertex is happy, therefore by the previous claim we can extend M_0

into M_1 through some vertex $a \in A$ that satisfied the condition of the first claim. In particular, since we proved that this process preserves the happiness of the vertices, we can repeat this process as long as the current matching M_i is still unstable.

Observe that every matching in the sequence is better than the previous ones, therefore such sequence cannot cycle. Moreover, since there is a finite number of possible matching of G, the sequence will eventually reach a matching where every vertex of A is matched — if $|A| \neq |B|$ we can assume that $|A| < |B|$ without loss of generality; call this matching $\hat{M}$. Hence, by contrapositive of the first claim, we know that $\hat{M}$ is either unstable, or has an unhappy vertex. However, by construction of our sequence, every vertex of $\hat{M}$ is happy, therefore it must be that $\hat{M}$ is stable. $\qquad \square$

2.4 Exercises

> **Problem 2.1**
>
> Let G be a k-regular bipartite graph, bipartitioned through (A, B). Prove that
>
> 1. $|A| = |B|$
>
> 2. G has a perfect matching

Proof. Since G is k-regular, it holds that the number of edges that have an endpoint in A is precisely $k|A|$, and the number of edges that have an endpoint in B is exactly $k|B|$; moreover, since G is bipartite, we have that

$$k|A| = k|B| \implies |A| = |B|$$

We will prove the second statement by using Theorem 2.4. By way of contradiction, assume that there is a set $S \subseteq V(G)$ such that $|S| > |\mathcal{N}(S)| \implies k|S| > k|\mathcal{N}(S)|$. Applying the same argument as before, the number of edges ab such that $a \in S, b \in \mathcal{N}(S)$ is exactly $k|S|$, hence by the pigeonhole principle there must be at least one vertex in $\mathcal{N}(S)$ that has more than k vertices, contradicting the k-regularity of G. $\qquad \square$

3

Graph packing

> **Definition 3.1:** *A-B* **paths**
>
> Given a graph G, and two sets $A, B \subseteq V(G)$, an *A-B* **path** is a path that has one end in A, one end in B, and no internal vertices in $A \cup B$.
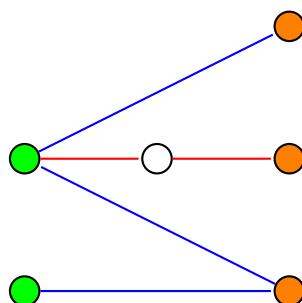

Figure 3.1: For instance, if A is the set of *green* vertices, and B is the set of *orange* vertices, then the *red* is an *A-B* path but the *blue* one is not.

Note that if $A \cap B \neq \varnothing$, any vertex in $A \cap B$ is a trivial *A-B* path.

> **Definition 3.2:** *A-B* **hitting set**
>
> Given a graph G, and two sets $A, B \subseteq V(G)$, an *A-B* **hitting** set is a set $X \subseteq V(G)$ such that every *A-B* path has a vertex in X.

Figure 3.2: For example, if A is the set of *green* vertices, and B is the set of *orange* vertices, then the *blue* is an A-B hitting set.

Note that if X is an A-B hitting set, then $A \cap B \subseteq X$, otherwise any single vertex in $A \cap B$ would represent an A-B trivial path that does not pass through the hitting set X. Moreover, consider an A-B hitting set X of a graph G; by definition, all the A-B paths must have at least one vertex in X, but since A-B paths do not have internal vertices in $A \cup B$, there cannot be $|X| + 1$ disjoint A-B paths.

Definition 3.3: A-B separation

Given a graph G, and two sets $A, B \subseteq V(G)$, two sets $X, Y \subseteq V(G)$ describe a **separation** (X, Y) of G if and only if

- $A \subseteq X$

- $B \subseteq Y$

- $X \cup Y = V(G)$

- no edge has one end in $X - Y$ and the other in $Y - X$

We say that a separation (X, Y) has **order** $|X \cap Y|$.

Given a graph G, two sets $A, B \subseteq V(G)$, and an A-B separation of G, we observe that $G[V(G) - (X \cap Y)]$ is disconnected.

Proposition 3.1

Given a graph G, and two sets $A, B \subseteq V(G)$, there exists an A-B hitting set of size k in G if and only if there exists an A-B separation of order k.

Proof. The converse implication is trivially true, because if (X, Y) is an A-B separation of G, then $X \cap Y$ itself is an A-B hitting set of G by definition; therefore, we just need to prove the direct implication.

Let Z be an A-B hitting set of G such that $|Z| = k$, and consider the connected components $C_1, \ldots, C_\ell$ of $G[V(G) - Z]$. Note that there cannot be any component C_i that has a vertex $a \in V(G) \cap (A - Z)$ and a vertex $b \in V(C) \cap (B - Z)$, because otherwise by connectivity of C_i there would be a path $a \to b$ — which is an A-B path — that avoids

the hitting set Z. Hence, we can define the two following sets

$$X := Z \cup \bigcup_{\substack{C \text{ component} \\ \text{of } G[V(G)-Z] \\ \text{such that } A \cap C \neq \varnothing}} V(C)$$

$$Y := Z \cup \bigcup_{\substack{C \text{ component} \\ \text{of } G[V(G)-Z] \\ \text{such that } A \cap C = \varnothing}} V(C)$$

We observe that, by definition

- $A \subseteq X$

- Y contains all the vertices in $B - Z$, therefore if $B \cap Z = \varnothing$ then trivially $B \subseteq Y$, and if $B \cap Z \neq \varnothing$ we observe that $Z \subseteq Y$ by definition, hence B is always completely covered by Y

- $X \cap Y = Z$

and for the previous observation there cannot be edges with one end in $X - Y$ and the other in $Y - X$. Hence, we conclude that (X, Y) is a separation of G, and $X \cap Y = Z \implies |X \cap Y| = |Z| = k$. $\qquad \square$

The following theorem establishes a deep connection between *connectivity* and *separation* in graphs, and was proved by Menger [Men27] in 1927.

> **Theorem 3.1: Menger's theorem**
>
> Given a graph G, and two sets $A, B \subseteq V(G)$, the number of disjoint A-B paths is equal to the minimum order of an A-B separation of G.

Proof. Consider an A-B separation (X, Y) of minimum order of G, i.e. a separation that minimizes $|X \cap Y|$; observe that each A-B path must traverse at least one vertex of $X \cap Y$, therefore the number of disjoint A-B paths that can exist is at most $|X \cap Y|$. This proves the direct implication of the theorem.

To prove the converse implication, we would need to prove that the minimum order of an A-B separation of G is no more than the number of disjoint A-B paths. Instead, we will prove that, for a fixed value k, in G there are k disjoint A-B paths, or there is an A-B separation of order at most $k-1$, which is logically equivalent. In particular, we will proceed by induction on m, i.e. the number of edges of G.

> *Base case.* If $m = 0$, the only possible A-B paths in G are trivial paths, i.e. isolated vertices in $A \cap B$. Hence, if $|A \cap B| \geq k$ the base case holds; instead, if $|A \cap B| \leq k - 1$, we observe that $A \cap B$ itself is an A-B hitting set of G of size at most $k - 1$, which means that there exists an A-B separation of order at most $k - 1$ by Proposition 3.1.
>
> *Inductive hypothesis.* Assume that the statement holds on a graph that has $m - 1$ edges.

Inductive step. We will prove that the statement holds for a graph G that has m edges as well. In particular, fix an edge $xy \in E(G)$, and note that the graph $G-\{xy\}$ has $m-1$ edges, meaning that the inductive hypothesis holds for $G-\{xy\}$. However, if $G-\{xy\}$ has k disjoint A-B paths, then G also has k disjoint A-B paths, therefore we may assume that the inductive hypothesis holds for $G-\{xy\}$ precisely because the latter has an A-B separation of order at most $k-1$.

Hence, let (X,Y) be an A-B separation of $G - \{xy\}$ of order at most $k-1$; we observe that if $x, y \in X$, or $x, y \in Y$, then (X,Y) is also a separation for G, which would conclude the proof. Thus, suppose that $x \in X - Y$ and $y \in Y - X$; then $(X, Y \cup \{x\})$ is an A-B separation of G that either has still order at most $k-1$, or it becomes a separation of order k with $k-1$ A-B paths. Since the first case concludes the proof, we may assume the second one to hold.

Consider the induced subgraphs $G[X]$ and $G[Y]$ ———————————————— da finire

$\square$

3.1 k-connectivity

> **Definition 3.4: k-connectivity**
>
> A graph G is said to be k-**connected** if $|V(G)| \geq k+1$, and for each $X \subseteq V(G)$ such that $|X| \leq k-1$ it holds that $G[V(G) - X]$ is connected.

In other words, a graph is k-connected if it has at least $k+1$ vertices, and it remains connected whenever fewer than k vertices are removed.
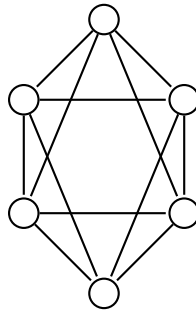

Figure 3.3: A 4-connected graph.

Clearly, any connected graph is 1-connected; moreover, any K_{t+1} clique is t-connected.

> **Proposition 3.2**
>
> If G is k-connected, then $\delta \geq k$.

Proof. By way of contradiction, assume that $\delta < k$, i.e. there is a vertex $x \in V(G)$ such

that $\deg(x) < k \implies |\mathcal{N}(x)| < k$. However, this violates the k-connectivity of G, because $G[V(G) - \mathcal{N}(x)]$ is disconnected even though $|\mathcal{N}(x)| \leq k - 1 \nmid$. $\qquad \square$

> **Proposition 3.3**
>
> If G is k-connected, then for any $xy \in E(G)$ it holds that $G - \{xy\}$ is $(k-1)$-connected.

Proof. By way of contradiction, suppose that $G - \{xy\}$ is not $(k-1)$-connected, i.e. there exists a set $X \subseteq V(G)$ such that $|X| \leq k - 2$ and $(G - \{xy\})[V(G) - X]$ is disconnected. By k-connectivity of G, since $|X| \leq k - 2$, we know that $G[V(G) - X]$ is connected, therefore if $(G - \{xy\})[V(G) - X]$ is disconnected it must imply that by removing xy we disconnect $G[V(G) - X]$. Now, since $|X| \leq k - 2$, we know that $|X \cup \{y\}| \leq k - 1$, but $G[V(G) - (X \cup \{y\})]$ does not contain the edge xy, hence it is disconnected, contradicting the k-connectivity of $G \nmid$. $\qquad \square$

> **Definition 3.5: Internal disjointness**
>
> Given a graph G, two paths P_1 and P_2 of G are said to be **internally disjoint** if every vertex of $V(P_1) \cap V(P_2)$ is an endpoint of P_1 and P_2.
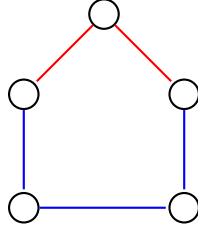

Figure 3.4: For instance, the *red* and *blue* paths are internally disjoint.

> **Proposition 3.4**
>
> Given a k-connected graph, and two sets $A, B \subseteq V(G)$, if $|A|, |B| \geq k$ then there exist k disjoint A-B paths in G.

Proof. By way of contradiction, if G does not admit k disjoint A-B paths, therefore by Theorem 3.1 the minimum order of an A-B separation is at most $k - 1$; hence, let (X, Y) be a minimum order A-B separation of G, i.e. $|X \cap Y| \leq k - 1$. Observe that $|A| \geq k$ and $A \subseteq X$, therefore $|X| \geq k$ which means that $X - Y \neq \varnothing$, and analogously $Y - X \neq \varnothing$. However, by definition of A-B separation we observe that $G[V(G) - (X \cap Y)]$ is disconnected, and we removed less than k vertices, contradicting the k-connectivity of $G \nmid$. $\qquad \square$

Proposition 3.5

Given a k-connected graph, for any pair of vertices $x, y \in V(G)$ such that $x \neq y$ there exist k $\{x\}$-$\{y\}$ internally disjoint paths.

Proof. Fix a pair of distinct vertices $x, y \in V(G)$; we have two cases.

- The two vertices are not adjacent, i.e. $x \nsim y$. By Proposition 3.2, we know that $|\mathcal{N}(x)|, |\mathcal{N}(y)| \geq k$, hence by the previous proposition we there exist k disjoint $\mathcal{N}(x)$-$\mathcal{N}(y)$ paths, which will form k *internally disjoint* $\{x\}$-$\{y\}$ paths along with the edges to x and y.

- The two vertices are not adjacent, i.e. $x \nsim y$. By k-connectivity of G, we know that $G - \{xy\}$ is $(k-1)$-connected by Proposition 3.3, therefore by the same argument of the previous case we have $k-1$ internally disjoint $\{x\}$-$\{y\}$ path, therefore G will have k internally disjoint $\{x\}$-$\{y\}$ paths along with the edge xy.

$\square$

Corollary 3.1

Given a k-connected graph G, a vertex $x \in V(G)$, and a subset $Y \subseteq V(G)$ such that $|Y| \geq k$, there exist k $\{x\}$-Y paths $P_1, \ldots, P_k$ such that $P_i \cap P_j = \{x\}$ — for all $i, j \in [k]$ such that $i \neq j$.

Proof. Add a vertex z to the vertices of G, and for each $y \in Y$ add an edge yz; we observe that $G[V(G) \cup \{z\}]$ is still k-connected, because $|Y| \geq k$. Now, since $x \neq z$, by the previous proposition there exist k $\{x\}$-$\{z\}$ internally disjoint paths, and these paths must traverse Y, therefore each path P_i must have the form $x \to y_i \to z$ for some $y_i \in Y$. Hence, if we consider the subpaths $x \to y_i$ for each i, in G these are precisely k $\{x\}$-Y paths intersects only in x. $\square$

Theorem 3.2

Given a k-connected graph G such that $k \geq 2$, and k vertices $x_1, \ldots, x_k \in V(G)$, there exist a cycle in G that contains all the vertices $x_1, \ldots, x_k$.

Proof. By way of contradiction, let C be the cycle that contains as many vertices x_i as possible; if $x_1, \ldots, x_k \in V(C)$, the theorem is trivially verified, hence we may assume that at least one vertex x_i is not in $V(C)$, and without loss of generality assume that $x_k \notin V(C)$. By Proposition 3.2 we know that $\delta \geq k$, and by Proposition 1.3 we know that $|V(C)| \geq k + 1$. Hence, by the previous corollary we know that exist k $\{x_k\}$-$V(C)$ paths that only intersect in x_k.

Now, since there are at most $k - 1$ vertices x_i in $V(C)$, and there are k $\{x_k\}$-$V(C)$ paths, by the pigeonhole principle there must be a subpath Q of C of the form $x_i \to x_j$ — for

$i, j \in [k-1]$ distinct — that has no internal vertex in $\{x_1, \ldots, x_{k-1}\}$, and such that there are two paths P_1 and P_2 that x_k as one endpoint and the other lies in Q.
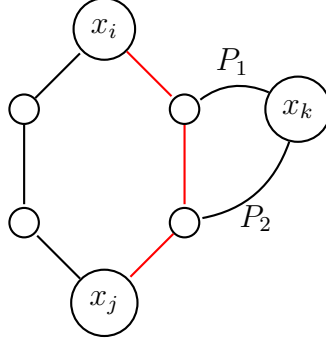

Figure 3.5: In the figure, the *red* segments compose the subpath Q.

Therefore, we can reroute Q by passing through P_1 and P_2 in order to find a cycle $C \cup \{x_k\}$ that has a larger number of vertices x_i, contradicting the definition of C ↯. $\qquad \square$

Note that this theorem *does not* guarantee the order in which the vertices are present in the cycle. In fact, asking to preserve the order of the vertices $x_1, \ldots, x_k$ given is a *much harder* question, and the only result we currently know is the following.

Theorem 3.3

Given a $10k$-connected graph G, and k vertices $x_1, \ldots, x_k \in V(G)$, there exist a cycle in G that contains all the vertices $x_1, \ldots, x_k$ in the same order.

3.2 Feedback vertex set

Consider the following combinatorial structure.

Definition 3.6: Feedback vertex set

Given a graph G, a subset $X \subseteq V(G)$ is said to be a **feedback vertex set** (FVS) if $G[V(G) - X]$ is acyclic; equivalently, X is an FVS if it intersects all cycles of G.
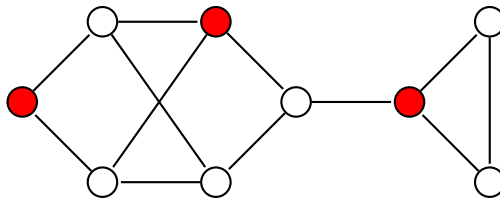

Figure 3.6: For instance, the *red* set of vertices is an FVS of this graph.

In many applications, it is important to find a **minimum** FVS; however, this problem is known to be NP-Complete — as proved by Karp [Kar72] in 1972. Therefore, we often focus on estimating the size of this set by computing lower and upper bounds. In particular, in this section we will present a result proved by Erdös and Pósa [EP65] in 1965. But before introducing their result about FVSs, we must discuss some combinatorial structures first.

3.2.1 Topological minors

Consider a graph G, and an edge $xy \in E(G)$; to **subdivide** xy means to *remove* the edge xy from $E(G)$ and *replacing* it with a 2-edge path $x\,z\,y$ for some new vertex z. For instance, if G contains the edge xy
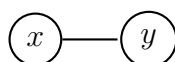
to *subdivide* the edge xy means to replace xy in G with the following 2-edge path
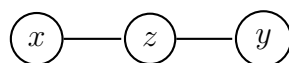
> **Definition 3.7: Subdivision**
>
> Given a graph G, a **subdivision** of G is a graph obtained from G by repeatedly subdividing its edges.
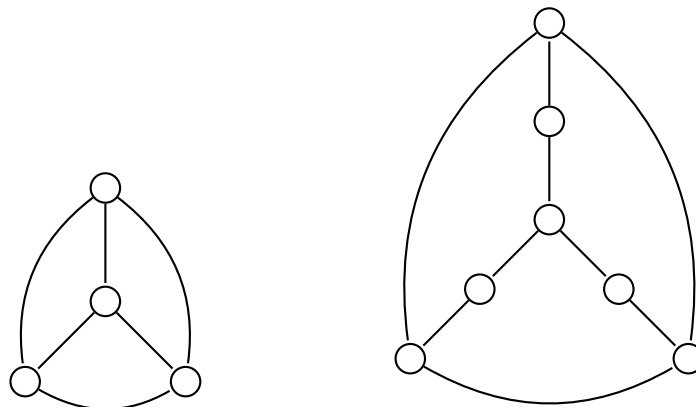

Figure 3.7: On the left: K_4. On the right: a subdivision of K_4.

Trivially, any graph is a subdivision of itself.

> **Definition 3.8: Topological minor**
>
> Given a graph G, and a graph H, we say that G contains H as **topological minor** if G has a subgraph which is a subdivision of H. Similarly, we say that H is a *topological minor* of G.
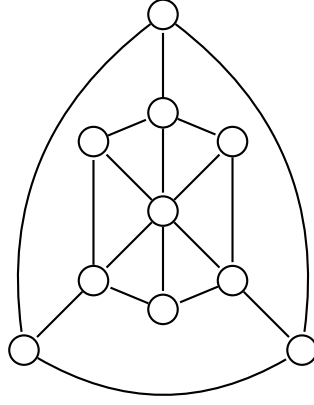
Figure 3.8: For instance, this graph has K_4 as topological minor, because it contains a subdivision of K_4 as subgraph.

Theorem 3.4

If G is 3-connected, then G has K_4 as topological minor.

Proof. If G is 3-connected, then by Proposition 3.2 we know that $\delta(G) \geq 3$, and by Proposition 1.3 this means that in G there is a cycle C. Moreover, by 3-connectivity of G, we know that G cannot be a cycle itself, therefore there must be at least one vertex $v \notin V(C)$. Now, by 3-connectivity agin we can apply Corollary 3.1, obtaining 3 $\{v\}$-$V(C')$ paths P_1, P_2 and P_3 that only intersect in v, and finally $G[V(C) \cup V(P_1) \cup V(P_2) \cup V(P_3)]$ is a subdivision of K_4. □

3.2.2 Erdős-Pósa theorem

Proposition 3.6

If G is 3-regular multigraph, then G has a cycle of length at most $2\lceil \log n \rceil$.

Proof. Clearly, if G admits loops or parallel edges, the statement is trivially true, therefore we may assume that G is a simple graph.

Fix a vertex $v \in V(G)$, and grow a BFS tree T from x as long no cycles of length at most $2\lceil \log n \rceil$ are encountered. Since G is 3-regular, T will be a tree in which the root x has 3 children, and all the other nodes that are not leaves will have 2 children. If the BFS stopped before visiting all the vertices of G, the statement holds, so we may assume that T covers all $V(G)$.

By 3-regularity of G, we have that

- $\delta \geq 3$, hence by Proposition 1.3 we know that G must contain a cycle
- the only possible cycle that can be formed in G is by connecting two leaves of T through an edge

and lastly, since $V(T) = V(G) = n$, the height of T is $\lceil \log n \rceil - 1$, which means that such a cycle must have length

$$\lceil \log n \rceil - 1 + \lceil \log n \rceil - 1 + 1 = 2 \lceil \log n \rceil - 1$$

$\square$

Along with the *subdivision* operation that we introduced previously, it naturally follows to defined the *inverse* operation of the subdivision, i.e. the **suppression**. Given a graph G, and two edges $xz, zy \in E(G)$ such that $\deg(z) = 2$, to *suppress* z means to *remove* the vertex z from $V(G)$ along with the edges xz and zy, and *replacing* them with an edge xy in $E(G)$. For instance, if G contains the edges xz and zy
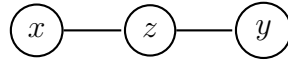
to *suppress* the vertex z means to replace xz and zy in G with the following edge xy
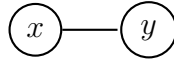
Now, consider the following definition.

> **Definition 3.9: Cut**
>
> Given an undirected graph $G = (V, E)$, and two subsets $S, T \subseteq V$, the **cut** induced by S and T on G is defined as follows
>
> $$\mathrm{cut}(S, T) = \{e \in E : |S \cap e| = |T \cap e| = 1\}$$
>
> In the case in which $T = \overline{S} = V - S$ we will simply write $\mathrm{cut}(S) = \mathrm{cut}(S, \overline{S})$.

In other words, a *cut* induced by a two sets S and T of vertices is the set of edges that have one endpoint in S and the other one in T.
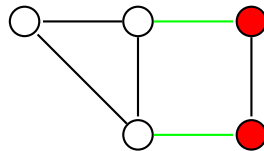

Figure 3.9: For instance, if the set S is described by the *red* vertices, then $\mathrm{cut}(S)$ is the set of green edges.

We will use the definition of cut in order to prove the following lemma.

Lemma 3.1

Given a 3-regular graph H such that $|V(H)| \geq c'k \log(k+1)$ for some constant c', then H contains k disjoint cycles.

Proof. We proceed by induction on k.

Base case. For $k = 1$, if H is 3-regular then $\delta(H) = 3$, therefore by Proposition 1.3 it must contain a cycle.

Inductive hypothesis. Assume that the statement holds for $k - 1$.

Inductive step. Let C be a cycle of H, and consider the graph $H - C$ — note that we are *not* considering the induced subgraph $H[V(H) - V(C)]$, but we are just removing the edges of C from H.

Claim: $H - C$ has a 3-regular graph H' as topological minor, such that $|V(H')| \geq |V(H)| - 2|V(C)|$.

Proof of the Claim. Starting with $V := V(C)$, as long as $H[V(H) - V]$ contains vertices of degree at most 1, move them into V; let M be the number of moved vertices. Then, since we are adding vertices to V, we have that $|\text{cut}(V)| \leq |\text{cut}(C)| - M$. We observe that, by construction $\delta(H[V(H) - V) \geq 2$, and the number of vertices that have degree 2 in $H[V(H) - V]$ is at most $|\text{cut}(V)| \leq |\text{cut}(C)| - M$.

Lastly, let H' be the graph obtained from $H[V(H) - V]$ by suppressing all the vertices of degree 2; then we have that H' contains at least the vertices of H, without the vertices of C, the M vertices that we moved during the procedure, and the vertices of degree 2 that we suppressed. In other words, we have that

$$|V(H')| \geq |V(H)| - |V(C)| - M - (|\text{cut}(C)| - M) \geq |V(H)| - 2|V(C)|$$

Note that the last inquality comes from the observation that $|\text{cut}(C)| \leq |V(C)|$, which follows by 3-regularity of H

Finally, since H' is 3-regular by construction, and $H[V(H) - V]$ is a subdivision of H' that is contained in $H - C$, this concludes the claim. $\square$

Now, since $|V(H)| \geq c'k \log(k+1)$, and by the previous proposition we know that

$$|V(C)| \leq 2\log(|V(H)|) + 2 \iff -2|V(C)| \geq -4\log(|V(H)|) - 4$$

where the added term 2 takes care of the rounding error from the ceiling operation — we obtain the following

$$\begin{aligned}
|V(H')| &\geq |V(H)| - 2|V(C)| \\
&\geq |V(H)| - 4\log(\log V(H)) - 4 \\
&\geq c'k\log(k+1) - 4\log(c'k\log(k+1)) - 4 \\
&\geq c'k\log(k+1) - 4[\log c' + \log k + \log(\log(k+1))] - 4
\end{aligned}$$

which is at least $c'(k-1)\log k$ for sufficiently big values of c'. In particular, since $|V(H')| \geq c'(k-1)\log k$, and we know that H' is 3-regular, we can apply the inductive hypothesis on H', which means that H' contains $k-1$ disjoint cycles.

Finally, if H' contains k disjoint cycles, and $H - C$ has H' as topological minor, i.e. $H - C$ contains a subdivision H'' of H', by definition of the subdivision operation H'' will contain $k-1$ disjoint cycles as well. Therefore, since $H - C$ contains H'', H contains k disjoint cycles along with the cycle C.

$\square$

We are now ready to prove the theorem that we introduced at the beginning of this section.

Theorem 3.5: Erdős-Pósa theorem

There is a constant c such that for any graph G, and any $k \in \mathbb{N}$

- either G has k vertex-disjoint cycles or

- there is an FVS X of G such that $|X| \leq ck \log k$

Proof. If G contains k disjoint cycles, the theorem is trivially true, hence we may assume that G does not have k disjoint cycles. Fix H to be the largest subgraph of G such that $\forall v \in V(H)$ $2 \leq \deg_H(v) \leq 3$ — note that H always exists because we may assume that G contains at least 1 cycle, since if G is acyclic the theorem is trivially true (and we would pick the cycle as H itself). Moreover, fix $\overline{H}$ to be the subgraph of G without all the "cycle components" of G — i.e. components of G that are cycles — and let $U := \{v \in V(\overline{H}) \mid \deg_{\overline{H}}(v) = 3\}$.

By TODO we know that $|U| < c'k \log(k+1)$, for some constant c. Let W be a set of vertices composed of one vertex for each cycle component — the ones not present in $\overline{H}$. `discrepanze lemma`

Claim: Every cycle of G is in H.

Proof of the Claim. If there was a cycle of G not present in H, this cycle would have contained vertices of degree at least 2, contradicting the maximality of H. $\square$

Claim: There are no paths P in $G[V(G) - (U \cup W)]$ of length at least 1 such that

- both ends of P are in $H[V(H) - (U \cup W)]$

- no internal vertex of P is in H

- no edge of P is in H

Proof of the Claim. If such a path P exists, then $H \cup P$ is a bigger graph than H that has every vertex of degree 2 or 3, contradicting the maximality of H. $\square$

Claim: Every cycle of $G - (U \cup W)$ intersects H in exactly one vertex.

Proof of the Claim. TODO $\square$ `da finire`

TODO □ da finire

3.3 Exercises

Problem 3.1

Given a k-connected graph G such that $V(G) \geq 2k$, show that G contains a cycle of length at least 2k.

Solution. TODO da fare

<div style="text-align: right; font-size: 4em; color: gray;">4</div>

Extremal graph theory

Extremal graph theory is a field of mathematics that investigates how *global properties* of a graph influence the **presence or absence** of specific substructures. Broadly speaking, it seeks to understand how large or dense a graph can be while avoiding certain configurations, or conversely, what constraints guarantee the appearance of particular patterns.

The central theme of the discipline lies in *exploring thresholds*: determining the minimal or maximal values of graph parameters — such as the number of edges, degrees, or vertices — beyond which a given substructure must necessarily exist. These substructures can range from simple motifs like *paths* and *cycles* to more complex or irregular patterns.

This area plays a crucial role in understanding the fundamental behavior of graphs under constraints, offering insights that are both theoretically significant and applicable in various areas such as network design, data analysis, and algorithm optimization.

4.1 Existence of substructures

4.1.1 Cliques

We will start our discussion of extremal graph theory with the question of **determining the existence** of cliques within a given graph. In general, finding the largest clique in a graph is an NP-Hard problem. This leads to a natural question: under what conditions can we guarantee the existence of a K_n subgraph?

Pál Turán addressed this by providing an exact upper bound on the number of edges a graph can have without containing a clique of a given size. But before discussing his results, let us first introduce some important definitions.

Definition 4.1: Edge maximum graph without K_r subgraphs

A graph G is said to be **edge maximum without K_r subgraph** — denoted as $\mathrm{EM}K_r$, for short — if for any G' such that $|V(G')| = |V(G)|$ and $|E(G')| > |E(G)|$, G' has K_r as subgraph.

In other words, G is edge maximum without K_r if by adding an edge to G we obtain a graph G' that contains K_r as subgraph.
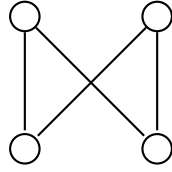
Figure 4.1: For instance, in this graph clearly there are no cliques as subgraph, and if we add one edge we construct K_3 as subgraph.

Definition 4.2: Independent set

Given a graph G, a subset $X \subseteq V(G)$ is said to be an **independent set** if for all u and v in X, it holds that $u \nsim v$.
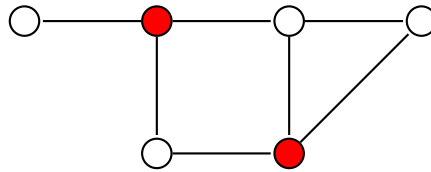
Figure 4.2: For instance, the *red* set of vertices is an independent set.

Note that we already encountered independent sets when discussing bipartite graphs: in fact, if G is bipartite through the bipartition (A, B), we observe that both A and B must be independent sets.

Now, consider the following type of graphs.

Definition 4.3: r-partite graph

A graph is sait to be r-**partite** if it can be partitioned into $X_1, \ldots, X_r$ sets of vertices such that $X_1, \ldots, X_r$ are independent sets.

Moreover, the graph is said to be **complete r-partite** if forall $i, j \in [r]$ such that $i \neq j$, if $x \in X_i$ and $y \in X_j$ then $x \sim y$.
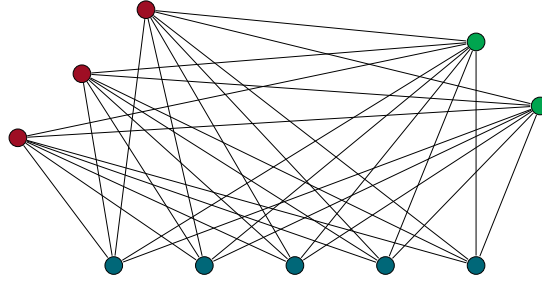
Figure 4.3: A complete 3-partite graph.

The following proposition shows that such graphs cannot contain cliques as subgraphs.

Proposition 4.1

An $(r-1)$-partite graph cannot contain K_r as subgraph.

Proof. Let G by an $(r-1)$-partite graph; by definition, G is composed of $r-1$ independent sets, therefore any subgraph of G that has r vertices $x_1, \ldots, x_r$ must have 2 vertices lying in the same independent set by the pigeonhole principle, meaning that any subgraph of vertices $x_1, \ldots, x_r$ will have two non-adjacent vertices. $\square$

From this proposition, it follows that the number of edges of a *complete* $(r-1)$-partite graph provides a *lower bound* for the maximum number of edges that a graph can have before inducing K_r as subgraph. Therefore, consider the following definition.

Definition 4.4: Turán graph

The Turán graph $T(n, r)$ is the complete r-partite graph that has n nodes, partitioned into $X_1, \ldots, X_r$, such that for all $i, j \in [r]$, if $i \neq j$ then $||X_i| - |X_j|| \leq 1$ — i.e. the cardinality of the partitions are as close to equal as possible.
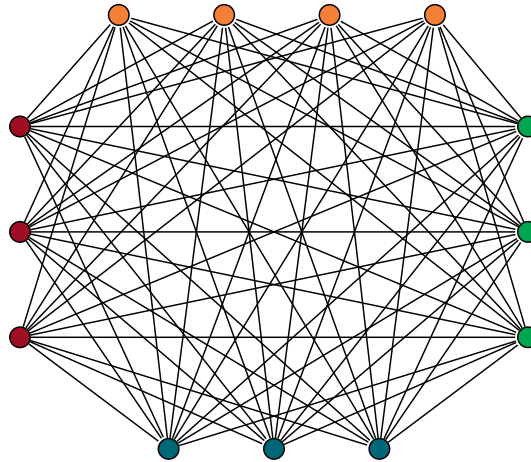
Figure 4.4: The Turán graph $T(13, 4)$.

We will refer to the number of edges of $T(n, r)$ as follows

$$t(n, r) := |E(T(n, r))| \sim \binom{n}{2} \frac{r-1}{r}$$

Note that, by definition, if $n \leq r$ then $T(n, r)$ can be partitioned into n independent sets that contain exactly 1 vertex each, and $r - n$ empty sets, meaning that $T(n, r) = K_n$.

> **Proposition 4.2**
>
> If G is EMK_r, then $t(n, r-1) \leq |E(G)|$.

Proof. We observe that $T(n, r-1)$ is $(r-1)$-partite in particular, therefore it does not contain K_r as subgraph by Proposition 4.1. However, since G has n vertices as well, and G is EMK_r, it must be that $t(n, r-1) \leq |E(G)|$. $\qquad\square$

Turán graphs are important not only because of the *lower bound* we discussed for Proposition 4.1, but also because Turán was able to prove in 1941 [Tur41] that any EMK_r graph *must be* exactly $T(n, r-1)$.

> **Theorem 4.1: Turán's theorem**
>
> If G is EMK_r, then $G = T(n, r-1)$.

Proof. We proceed by strong induction on n.

Base case. For the base case, we consider $n < r$; in fact, G is EMK_r, and in particular it is edge maximum, therefore for $n < r$ we have that G is K_n. Hence, $G = T(n, r-1)$ — some of the $r-1$ partitions may be empty.

Strong inductive hypothesis. Assume that the statament is true for an EMK_r graph of $r \leq k \leq n-1$ vertices.

Inductive step. We are going to prove the statement for $n \geq r$. Note that G is EMK_r, therefore by adding an edge between two non-adjacent vertices we obtain K_r as subgraph of G; this must imply that K_{r-1} must be a subgraph of G. Hence, let $Y := \{y_1, \ldots, y_{r-1}\}$ be the vertices such that $G[Y] = K_{r-1}$.

Now, since G is EMK_r, then $G[V(G) - Y]$ does not contain K_r, therefore it must be that $|E(G[V(G) - Y])|$ is at most the number of edges of a graph H that is EMK_r of $n - (r-1)$ vertices. However, by strong inductive hypothesis we know that $H = T(n - (r-1), r-1)$, therefore

$$|E(G[V(G) - Y])| \leq |E(H)| = t(n - (r-1), r-1)$$

Claim: If $n \geq r$, then $t(n, r) = t(n-r, r) + (n-r)(r-1) + \binom{r}{2}$.

Proof of the Claim. Let $X_1, \ldots, X_r$ be a partition of $T(n, r)$. Since $n \geq r$, we

know that the independent sets are not empty, thus fix $x_i \in X_i$ for each X_i of the partition, and let $Z := \{x_1, \ldots, x_r\}$. Now, we observe the following.

- $G[Z] = K_r$, hence there are $\binom{r}{2}$ edges in $G[Z]$.

- $G[V(G) - Z] = T(n - r, r)$ since we only removed one vertex per independent set, thus there are $t(n - r, r)$ edges in $G[V(G) - Z]$.

- Fix an independent set X_i, and consider a vertex $y \in X_i$; this vertex has $r - 1$ neighbors in Z, i.e. $|N(y) \cap Z| = r - 1$, because $T(n, r)$ is a complete r-partite graph, and the r-th node of Z is in X_i itself. Therefore, because there are $n - r$ vertices outside G, there are $(r - 1)(n - r)$ edges between $G[Z]$ and $G[V(G) - Z]$.

The claim follows from the above statements. $\square$

Now, similar to what we did in the claim, we observe the following.

- $G[Y] = K_{r-1}$, hence there are $\binom{r-1}{2}$ edges in $G[Y]$.

- $|E(G[V(G) - Y])| \leq t(n - (r - 1), r - 1)$ for the previous observation.

- Any vertex $z \notin Y$ must have at most $r - 2$ neighbors in Y — if z had $r - 1$ neighbors, $G[Y \cup \{z\}]$ would be a K_r, contradicting the definition of G. Hence, there are $(r - 2)(n - (r - 1))$ edges between $G[Y]$ and $G[V(G) - Y]$.

Therefore, we get that

$$|E(G)| \leq \binom{r - 1}{2} + (r - 2)(n - (r - 1)) + t(n - (r - 1), r - 1) = t(n, r - 1)$$

where the last inequality follows from the previous claim. Moreover, for Proposition 4.2 this upper bound is tight. However, if this bound is tight, it must be that the vertices $z \notin Y$ have *exactly* $r - 2$ neighbors in Y; hence, for each $i \in [r - 1]$ let $X_i := \{z \in V(G) - Y \mid z \nsim y_i\} \cup \{y_i\}$. In particular, since z has exactly $r - 2$ neighbors in Y, it is adjacent to every vertex in Y except for y_i, meaning that for any vertex $z \in V(G)$ it holds that z is in *exactly* one set between $X_1, \ldots, X_{r-1}$. Therefore, these sets describe a *partition* of G.

Claim: For all $i \in [r - 1]$, it holds that X_i is an independent set.

Proof of the Claim. Fix $i \in [r - 1]$, and by way of contradiction suppose that X_i is not an independent set, i.e. $\exists z, z' \in X_i \quad z \sim z'$. However, $z \in X_i \implies z \notin Y$, and by the previous observation we know that z has $r - 2$ neighbors in Y, meaning that z is adjacent to every vertex in $Y - \{y_i\}$ — and the same reasoning can be applied on z'. Therefore, we have that $G[(Y - y_i) \cup \{z, z'\}]$ is K_r because $z \sim z'$, contradicting the definition of $G \nleq$. $\square$

Therefore, G is $(r - 1)$-partite through the partition $X_1, \ldots, X_{r-1}$.

Claim: If G is EMK_r and $(r - 1)$-partite, then $G = T(n, r - 1)$.

Proof of the Claim. Since no edges can be added to G without creating K_r as

subgraph of G, it must be that G is complete $(r-1)$-partite. Let $X_1, \ldots, X_r$ be the partition that defines G, and by way of contradiction suppose that $G \neq T(n, r-1)$. By definition of Turán graph, this implies that for distinct $i, j \in [r]$ it holds that $||X_i| - |X_j|| \geq 2$, and without loss of generality suppose that $|X_i| < |X_j|$. Fix a vertex $y \in X_j$, and construct the graph G' starting from G as follows: remove all the edges $\{y, x_i\}$ for $x_i \in X_i$, move y into X_i, and add all the edges $\{y, x_j\}$ for $x_j \in X_j - \{y\}$. We observe that G' is still $(r-1)$-partite by construction, and by Proposition 4.1 we know that G' cannot contain K_r as subgraph. However, we observe that

$$|E(G')| = |E(G)| - |X_i| + |X_j| - 1 > |E(G)|$$

contradicting the fact that G was $\mathrm{EM}K_r \not\downarrow$. $\qquad \square$

Finally, the statement follows from this last claim.

$\square$

As a corollary of this theorem, we have the answer to our original question: if a graph G has n vertices and more than $t(n, r-1)$ edges, it clearly cannot be $T(n, r-1)$, therefore by the theorem it must contain K_r as subgraph, hence the bound is

$$|E(G)| \geq t(n, r-1) + 1$$

4.1.2 k-connected subgraphs

TODO `intro`

Mader [Mad72]

> **Theorem 4.2: Mader's theorem**
>
> If G is such that $|V(G)| \geq 2k$ and $|E(G)| \geq (2k-1)(n-k)$, then G contains a k-connected subgraph.

Proof. We proceed by strong induction on n.

Base case. Since $n \geq 2k$, the base case of the induction is $n = 2k$, thus we have that

$$|E(G)| \geq (2k-1)k = \frac{(2k-1)2k}{2} = \binom{2k}{2}$$

which must imply that $G = K_{2k}$, which is $(2k-1)$-connected by definition.

Strong inductive hypothesis. Assume that the statement holds for any graph of h vertices such that $n - 1 \geq h > 2k$ vertices.

Inductive step. We are going to prove the statement for a graph G of n vertices such that $n > 2k$, i.e. $n \geq 2k + 1$.

Suppose that $\delta < 2k$, i.e. $\delta \leq 2k - 1$, implying that there exists a vertex v such that $\deg(v) \leq 2k - 1$, and consider $G[V(G) - \{v\}]$. Since this subgraph has $n - 1$

vertices, and we are assuming that $|E(G)| \geq (2k-1)(n-k)$ then we have that

$$
\begin{aligned}
|E(G[V(G) - \{v\}])| &\geq |E(G)| - (2k-1) \\
&\geq (2k-1)(n-k) - (2k-1) \\
&\geq (2k-1)(n-k-1) \\
&\geq (2k-1)((n-1)-k)
\end{aligned}
$$

Moreover, since $n \geq 2k+1 \iff n-1 \geq 2k$ and $G[V(G) - \{v\}]$ has $n-1$ vertices, the latter has at least $2k$ vertices, hence we can apply the inductive hypothesis on it and obtaining a k-connected subgraph to prove the statement trivially. Therefore, we may assume that $\delta \geq 2k$.

To recap, we are now assuming that

(a) G does not contain a k-connected graph — otherwise the statement is trivially true

(b) $|E(G)| \geq (2k-1)(n-k)$

(c) $\delta \geq 2k$

(d) $n \geq 2k+1$

and in particular, assumption (a) implies that we may assume there is an A-B separation (X, Y) of G having order $|X \cap Y| \leq k-1$, where $X - Y, Y - X \neq \varnothing$ — for some subsets $A, B \subseteq V(G)$.

Claim: It holds that $|E(G[X])| \geq (2k-1)(|X|-k)$ or $|E(G[Y])| \geq (2k-1)(|Y|-k)$.

Proof of the Claim. By way of contradiction, suppose that $|E(G[X])| \leq (2k-1)(|X|-k) - 1$ and $|E(G[Y])| \leq (2k-1)(|Y|-k) - 1$. Thus, we have that

$$
\begin{aligned}
|E(G)| &\leq |E(G[X])| + |E(G[Y])| \\
&\leq (2k-1)(|X|-k) - 1 + (2k-1)(|Y|-k) - 1 \\
&= (2k-1)(|X| + |Y| - 2k) - 2 \\
&= (2k-1)(|X \cup Y| + |X \cap Y| - 2k) - 2 \\
&\leq (2k-1)(n + k - 1 - 2k) - 2 \\
&= (2k-1)(n - k - 1) - 2 \\
&< (2k-1)(n - k)
\end{aligned}
$$

which contradicts assumption (b) $\notlightning$. $\qquad\square$

This claim implies that at least one between $G[X]$ or $G[Y]$ satisfies the edge bound, hence if we can prove that they also satisfy the vertex bound we can apply the inductive hypothesis and conclude the proof.

Now consider assumption (c): if $\delta \geq 2k$ then for each vertex $v \in V(G)$ its degree must be at least $2k$, and in particular this is true for any $v \in X - Y$. However, (X, Y) is a separation of G, meaning that there are not edges in $\mathrm{cut}(X, Y)$; therefore, all the neighbors of v must lie inside X, meaning that $|X| \geq |\mathcal{N}(v) \cup v| \geq 2k + 1 > 2k$.

This last condition implies that the vertex bound of the inductive hypothesis is satisfied, so by applying the latter on one between $G[X]$ or $G[Y]$ the theorem holds.

$\square$

4.1.3 Cliques as minors

TODO intro

> **Theorem 4.3: Bollobás-Thomason theorem**
>
> There exists a constant c such that if G has at least ck^2n edges, then G has K_k as topological minor.
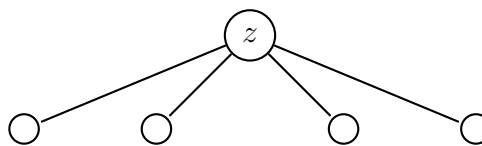
TODO $2 \leq c \leq 10$ interlude

Consider a graph G, and an edge $xy \in E(G)$; to **contract** xy means to *identify* the vertices x and y and delete any parralel edge that formed afterwards. For instance, if G contains the edge xy
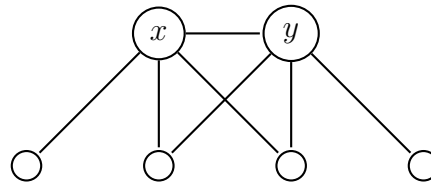
to *contract* the edge xy means to identify x and y in a new vertex z, and removing the 2 parallel edges that formed
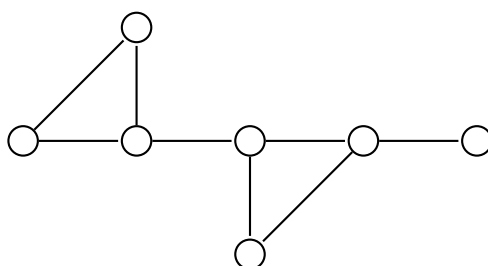

If we contracted the edge e in a graph G, we will denote the resulting graph as G/e.

> **Definition 4.5: Minor**
>
> Given a graph G, and a graph H, we say that G contains H as **minor** if H can be obtained from G by a series of *vertex deletions*, *edge deletions* and *edge contractions*. Similarly, we say that H is a *minor* of G.

For instance, consider the following graph G

then clearly the following graph H is a minor of G — the *dashed* edges and vertices are the removed ones, and the *gray* edge is the contracted edge
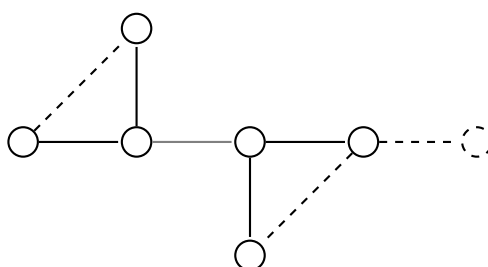
Consider a graph G that has H as topological minor, meaning that it contains a graph H' that is a subdivision of H; if we remove from G all the edges and vertices that are *not* part of H', we are left with a subdivision of H, and through the *edge contraction* operation we can actually reverse the *edge subdivision* in order to obtain H from H'. In fact, for any 2-edge path x y z, we can contract x y obtaining an edge y z, and no parallel edge can form because we previously removed all the other edges that are not in H', i.e. edges that were not constructed through subdivision. This means that if G contains H as topological minor, then G contains H as minor as well. However, the opposite is not true in general: for instance, consider the following graph G
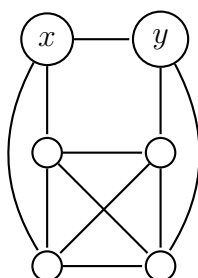
If we contract the edge xy, we obtain K_5, therefore G contains K_5 as minor, even if K_5 is not a topological minor of G: in fact, every subdivision of K_5 has 5 vertices of degree 4, but G has only 4 vertices of degree 4.

TODO

> ### Theorem 4.4
>
> If p is an integer such that $2 \leq p \leq 7$, then for any graph G such that
>
> $$|E(G)| \geq (p-2)n - \binom{p-1}{2} + 1$$
>
> it holds that G contains K_p as minor.

Proof. We will prove each value of p individually.

- For $p = 2$, we see that $|E(G)| \geq (2-2)n - \binom{2-1}{2} + 1 = 1$, therefore G contains at least 1 edge, and in fact K_2 is nothing more than a single edge.

- For $p = 3$, we have that $|E(G)| \geq (3-2)n - \binom{3-1}{2} + 1 = n - 1 + 1 = n$. Thus, by Proposition 1.5 this implies that G is not a tree (nor a forest), and in particular it is not acyclic, meaning that it contains at least one cycle C. Hence, if we contract all but 3 edges on C we obtain K_3.

- For $p = 4$, we have that $|E(G)| \geq (4-2)n - \binom{4-1}{2} + 1 = 2n - 3 + 1 = 2n - 2$. We proceed by induction on $n + m$.

  We observe that if $n = 2$ then $|E(G)| \geq 2 \cdot 2 - 2 = 2$ which is impossible with only 2 vertices, and similarly if $n = 3$ then $|E(G)| \geq 2 \cdot 3 - 2 = 4$ which is impossible with only 3 vertices. Hence, to satisfy the edge bound n must be at least 4. Therefore, for the base case assume that $n = 4$, which means that $|E(G)| \geq 2 \cdot 4 - 2 = 6$, and thus $G = K_4$.

  Now, for the inductive hypothesis, we need to be careful: since we are applying the induction on $n + m$, then

  - if the graph G we are considering has $m - 1$ edges, we are going to inductively assume that the property holds for G such that $|E(G)| \geq 2n - 2$

  - if the graph G we are considering has $n - 1$ vertices, we are going to inductively assume that the property holds for G such that $|E(G)| \geq 2(n-1) - 2$

  Now, because of the base case, consider a graph G of n vertices such that $n \geq 5$; moreover, suppose that $|E(G)| > 2n - 2$, and fix an edge $e \in E(G)$. Now, consider the graph $G - \{e\}$: clearly, this graph has still n vertices but since we removed an edge it holds that $|E(G - \{e\})| > 2n - 2 - 1 > 2n - 3 \iff |E(G - \{e\})| \geq 2n - 2$, therefore we can apply the inductive hypothesis as we described, and the theorem holds.

  Therefore, we may assume that $|E(G)| = 2n - 2$. Fix an edge $e \in E(G)$, and consider the contracted graph G/e: this graph has $n - 1$ vertices, hence if $|E(G/e)| \geq 2(n-1) - 2$ by the inductive hypothesis the theorem holds trivially. Thus, we may assume that $|E(G/e)| < 2(n-1) - 2 = 2n - 4 \iff |E(G/e)| \leq 2n - 5$, and since we are assuming that $|E(G)| = 2n - 2$, this means that by contracting e we removed at least 3 edges from G — one of which is e itself. Therefore, we know that in G we have the following structure
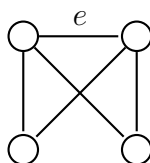
and in particular, the endpoints of e must have at least two common neighbors. This implies that e is contained in at least two distinct K_3's. By applying the same argument on every edge e of G, we get that $\delta \geq 3$ because the degree of the endpoints of e must be at least 3. Moreover, if $\delta \geq 4$, by Corollary 1.1 we have that $|E(G)| \geq \dfrac{4n}{2} = 2n$ which is greater than $2n - 2$, contradicting our assumption that $|E(G)| = 2n - 2$. This implies that $\delta = 3$. Finally, fix a vertex v of degree 3: TODO

da finire

- For $p = 5$ TODO

da finire

- Lastly, the cases for $p = 6$ and $p = 7$ were proved by TODO

da trovare i paper

Bibliography

[Ber57]   Claude Berge. "TWO THEOREMS IN GRAPH THEORY". In: *Proceedings of the National Academy of Sciences* 43.9 (Sept. 1957), 842–844. ISSN: 1091-6490. DOI: 10.1073/pnas.43.9.842. URL: http://dx.doi.org/10.1073/pnas.43.9.842.

[con25]   Wikipedia contributors. *Seven Bridges of Königsberg*. Jan. 2025. URL: https://en.wikipedia.org/wiki/Seven_Bridges_of_K%C3%B6nigsberg.

[EP65]   p. Erdös and L. Pósa. "On Independent Circuits Contained in a Graph". In: *Canadian Journal of Mathematics* 17 (1965), 347–352. DOI: 10.4153/CJM-1965-035-8.

[Eul41]   Leonhard Euler. "Solutio problematis ad geometriam situs pertinentis". In: *Commentarii academiae scientiarum Petropolitanae* (1741), pp. 128–140.

[GS62]   David Gale and Lloyd S Shapley. "College admissions and the stability of marriage". In: *The American mathematical monthly* 69.1 (1962), pp. 9–15.

[Hal35]   P. Hall. "On Representatives of Subsets". In: *Journal of the London Mathematical Society* s1-10.1 (Jan. 1935), 26–30. ISSN: 0024-6107. DOI: 10.1112/jlms/s1-10.37.26. URL: http://dx.doi.org/10.1112/jlms/s1-10.37.26.

[Kar72]   Richard M. Karp. "Reducibility among Combinatorial Problems". In: *Complexity of Computer Computations*. Springer US, 1972, 85–103. ISBN: 9781468420012. DOI: 10.1007/978-1-4684-2001-2_9. URL: http://dx.doi.org/10.1007/978-1-4684-2001-2_9.

[Kon]   Denés Konig. *Gráfok és mátrixok. Matematikai és Fizikai Lapok, 38: 116–119, 1931*.

[LP09]   L. Lovász and M.D. Plummer. *Matching Theory*. AMS Chelsea Pub., 2009. URL: https://books.google.it/books?id=OaoJBAAAQBAJ.

[Mad72]   Wolfgang Mader. "Existenz n-fach zusammenhängender Teilgraphen in Graphen genügend grosser Kantendichte". In: *Abhandlungen aus dem mathematischen Seminar der Universität Hamburg*. Vol. 37. Springer. 1972, pp. 86–97.

[Men27]   Karl Menger. "Zur allgemeinen kurventheorie". In: *Fundamenta Mathematicae* 10.1 (1927), pp. 96–115.

[Tur41]   Paul Turán. "On an external problem in graph theory". In: *Mat. Fiz. Lapok* 48 (1941), pp. 436–452.