Script Graph Theory

Based on a lecture by Thomas Gatter and Prof. Peter Stadler and a special lecture by Annachiara Korchmaros

Nicolas Handke

Winter semester 2023/2024

This script has been put together by NICOLAS HANDKE
during winter semester 2023/2024.

The layout is inspired by FELIX LINKER.

Contents

I Base Lecture **4**

1 Foundation **5**

2 Algebraic Graph Theory **16**

3 Cycles in Graphs **32**
3.1 Linear Independence & Cycle bases 36
3.2 Fundamental & Strictly Fundamental Cycle Bases 43
3.3 Minimal Cycle Bases & Canonical Greedy 45

4 Planar Graphs **57**
4.1 Cycle Bases in Planar Graph 63

5 Graph Coloring **70**
5.1 Coloring in Planar Graphs 72

6 CoGraphs **79**

7 Graph Isomorphisms **84**

8 Random Graphs **95**

Bibliography **98**

Part I.

Base Lecture

1. Foundation

Definition 1.1. A graph $\mathcal{G}$ is a tuple (V, E) consisting of

- a set of vertices V and
- a set of edges $E \subseteq V \times V$.

The *size* of $\mathcal{G}$ is equivalent to the number of vetices

$$|\mathcal{G}| := |V|.$$

Edges are denoted by the two incident vertices.

- If $\mathcal{G}$ is undirected the order of the vertices does not matter and we therefore write

$$e = \{x, y\} \in E.$$

- If $\mathcal{G}$ is a directed graph, an edge *from x to y* is written as ordered tuple

$$e = (x, y) \in E.$$

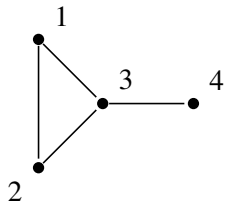
If no further specification regarding $\mathcal{G}$ being directed or undirected is required or the properties are clear from the context, we can use the simplified notation $e = xy$.

If clear from the context we can write $v \in \mathcal{G}$ to abbreviate $v \in V$ for a graph $\mathcal{G} = (V, E)$.

Example 1.1. The undirected example graph $\mathcal{G}_e$ given by

$$\mathcal{G}_e := (\{1,2,3,4\}, \{\{1,2\},\{1,3\},\{2,3\},\{3,4\}\})$$

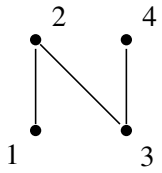
can be visually represented as:



Example 1.2. For directed graphs like

$$\mathcal{G}_1 := (\{1,2,3,4\}, \{(1,2),(2,3),(3,4)\})$$

the order of the vertices for each edge matters. In visualizations this is typically indicated via a directional arrow.



Example 1.3. A specific graph is the *empty* given by

$$\mathcal{G}_2 = (\emptyset, \emptyset)$$

which sometimes abbreviated to $\mathcal{G}_2 = \emptyset$.

Remark. *If not declared otherwise, we will assume simple graphs, i.e. graphs with no self-loops (or 'reflexive' edges) and no multi-edges.*
If not directly stated or otherwise clear from the context, we will also assume that a graph is undirected.

Example 1.4. The following visualizes the distinguishing of

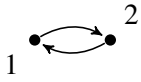
- multi-graphs with self-loops,

  undirected  and directed 

- multi-graphs with multi-edges,

  undirected  and directed 

- and *simple*, directed graphs with multiple edges

  because in the latter case both edges can be clearly distinguished through the vertices alone as $(1, 2) \neq (2, 1)$.

Definition 1.2 (Subgraph)**.** Let $\mathcal{G} = (V, E)$ and $\mathcal{G}' = (V', E')$ s.t.

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

then $\mathcal{G}'$ is called a subgraph of $\mathcal{G}$.

Furthermore, $\mathcal{G}'$ is an induced subgraph of $\mathcal{G}$ if

$$E' = E \cap (V' \times V')$$

E' inherits all present edges between vertices in V'

and we write $\mathcal{G}' = \mathcal{G}[V']$ or $\mathcal{G}' = \mathcal{G}\big|_{V'}$.
(Speaking: *$\mathcal{G}'$ is the subgraph of $\mathcal{G}$ induced by V'*)

For a graph $\mathcal{G} = (V, E)$ and subgraph $\mathcal{G}' = (V', E')$ of $\mathcal{G}$ we can denote their (vertex-based) difference as

$$\mathcal{G} - \mathcal{G}' := \mathcal{G}[V \setminus V'].$$

Definition 1.3 (Adjacency)**.** Let $\mathcal{G} = (V, E)$ be a graph.
Two vertices $x, y \in V$ are adjacent, or neighbors if

$$\exists\, xy \in E.$$

there exists an edge between x and y

Definition 1.3.1 (Predecessor & Successor)**.** Let $\mathcal{G} = (V, E)$ be a *directed* graph.
Regarding an edge $(x, y) \in E$ from x to y we denote

- x being a predecessor to y and
- y being a successor to x.

Definition 1.4 (Neighborhood)**.** Let $\mathcal{G} = (V, E)$ be an undirected graph.
The set of neighbors of a vertex $v \in V$ can be denoted by

$$\mathrm{N}(v) := \{y \in V \mid \{v, y\} \in E\}.$$

Definition 1.4.1 (Predecessors & Successors)**.** Let $\mathcal{G} = (V, E)$ be a directed graph.
For a vertex $v \in V$ we denote

- the set of predecessors

  $$\mathrm{N}_-(v) \text{ or } \mathrm{N}_{\text{in}} := \{y \in V \mid (y, v) \in E\}$$

  and
- the set of successors

  $$\mathrm{N}_+(v) \text{ or } \mathrm{N}_{\text{out}} := \{y \in V \mid (v, y) \in E\}.$$

Remark. *In a directed graph, a vertex's neighborhood most commonly refers to its set of successors, however it doesn't have to and the meaning should always be drawn from the context.*

Definition 1.5 (Degree)**.** Let $\mathcal{G} = (V, E)$ be an undirected graph.
The degree of a vertex $v \in V$ is

$$d(v) \text{ or } \deg(v) := |N(v)|.$$

the number of its neighbors

Definition 1.5.1 (In-Degree, Out-Degree & Full-Degree)**.** Let $\mathcal{G} = (V, E)$ be a directed graph.
For a vertex $v \in V$ we denote

- its in-degree as

$$\deg_{\text{in}}(v) := |\,\mathrm{N}_{\text{in}}(v)|,$$

- its out-degree as

$$\deg_{\text{out}}(v) := |\,\mathrm{N}_{\text{out}}(v)|,$$

  and

- its full-degree as

$$\deg_{\text{full}}(v) \text{ or } \deg(v) := |\,\mathrm{N}_{\text{in}}(v)| + |\,\mathrm{N}_{\text{out}}\,|.$$

Definition 1.6. Let $\mathcal{G} = (V, E)$ be an undirected graph.
We denote

- the minimal degree in $\mathcal{G}$

$$\delta(\mathcal{G}) := \min_{v \in V} \deg(v),$$

- the maximal degree in $\mathcal{G}$

$$\Delta(\mathcal{G}) := \max_{v \in V} \deg(v),$$

  and

- the average degree in $\mathcal{G}$

$$\deg(\mathcal{G}) := \frac{1}{|V|} \sum_{v \in V} \deg(v).$$

Intuitively the following holds:

$$\delta(\mathcal{G}) \leq \deg(\mathcal{G}) \leq \Delta(\mathcal{G})$$

Proposition 1.7. Let $\mathcal{G} = (V, E)$ be an undirected graph.
The number of vertices of odd degree is always even.

Proof. Let $\mathcal{G} = (V, E)$ be an undirected graph.
As each edge is defined by exactly two vertices, the number of edges is equivalent to

$$|E| = \frac{1}{2} \sum_{v \in V} \deg(v).$$

Due to $|E| \in \mathbb{N}_0$, the sum of all vertices must be even and, therefore, the number of vertices of odd degree must also be even. □

Definition 1.8 (Path)**.** Let $\mathcal{G} = (V, E)$ be a graph.
A path is a non-empty graph $P = (V_P, E_P)$ with

- a set of unique vertices $V_P = \{x_0, x_1, \ldots, x_k\} \subseteq V$ and
- a set of (unique) edges $E_P = \{x_0x_1, x_1x_2, \ldots, x_{k-1}x_k\}$.

P is a path of length $k = |E_P|$.

Typically we write paths as their natural sequence of vertices

$$P = \underbrace{x_0}_{=e_0} \overbrace{x_1 x_2}^{=e_1} \ldots x_k.$$

If the inner vertices are of no specific relevance, we can abbreviate it as a path from x_0 to x_k.

Definition 1.9 (Cycle)**.** Let $\mathcal{G} = (V, E)$ be an undirected graph and

$$P = (V_P, E_P) = x_0 \dots x_{k-1}$$

a path in $\mathcal{G}$ with $k \geq 3$. If $\exists\, x_{k-1}x_0 \in E$, then

$$C := (V_P, E_P \cup \{x_{k-1}x_0\}) \text{ or shorter } C := P + x_{k-1}x_0$$

the extension of P by the edge $x_{k-1}x_0$

is a cycle of length k in $\mathcal{G}$.

Definition 1.10 (Walk)**.** Let $\mathcal{G} = (V, E)$ be a graph.
A walk of length k in $\mathcal{G}$ is a non-empty sequence

$$w = v_0e_0v_1e_1 \dots e_{k-1}v_k \text{ s.t. } \forall\, 0 \leq i < k :\ e_i = v_iv_{i+1} \in V.$$

If $v_0 = v_k$ we call w a closed walk.

Definition 1.11 (Trail)**.** A trail is a walk where all edges are distinct.

Remark. *A path is a trail in which all vertices are distinct.*

Proposition 1.12. Let $\mathcal{G} = (V, E)$ be a graph.
Every walk between two vertices contains a path between these vertices.

Proof. Let $\mathcal{G} = (V, E)$ be a graph and $x, y \in V$.
W.l.o.g. we can assume there exists a walk between x and y and $x \neq y$.
Proof by induction over the length of w:

- *Induction Predicate* $P(n)$:
  If there exists a walk w of length n between x and y, then w contains a path between x and y.
- *Basic step*:
  For a length of 1, there only exists a single edge uv in any walk w. Therefore, $P(1)$ is true.

- *Induction requirement (IR)*:
  Let $P(k)$ be true for some $k \geq 1$.
- *Induction hypothesis (IH)*:
  $P(k) \Rightarrow P(k+1)$
- *Induction step (IS):*
  W.l.o.g. let $w = v_0 e_0 v_1 \dots e_k v_{k+1}$ be a walk of length $k+1$ between $v_0 = x$ and $v_{k+1} = y$.
  a) Assume w has no vertex appearing more than once. Therefore, w defines a path.
  b) Assume there exist indices $0 \leq i < j \leq k+1$ s.t. $v_i = v_j$. We can therefore define a shorter walk

$$w' := x \dots e_{i-1} \underbrace{v_i}_{=v_j} e_j \dots y$$

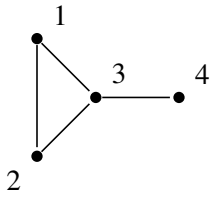
  which is by construction still a valid walk between x and y.
  The walk w' has length $k + 1 - (j - i) \leq k$. By (IR) we know that $P(k)$ is true and w' therefore contains a path between x and y. Due to the construction of w', it follows that w contains a path between x and y.

  $\Rightarrow P(k+1)$ is true.

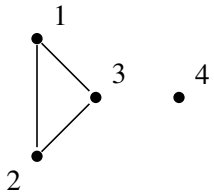
$\Rightarrow$ Proposition 1.12 holds. □

Definition 1.13 (Connectedness)**.** A graph $\mathcal{G}$ is called connected if any two of its vertices are linked by a path in $\mathcal{G}$.

Example 1.5. Our example graph $\mathcal{G}_e$ is connected:



while the variation $\mathcal{G}'_e$ is not a connected graph:



Proposition 1.14. The vertices of an undirected, connected graph $\mathcal{G}$ with $n = |\mathcal{G}|$ can always be enumerated, say $v_1, \ldots, v_n$, s.t. for every $1 \leq i \leq n$ the subgraph $\mathcal{G}_i := \mathcal{G}[\{v_1, \ldots, v_i\}]$ is also connected.

Proof. Let $\mathcal{G}$ be an undirected, connected graph.
Proof by induction over i:

- *Induction Predicate* $P(n)$:
  For every $1 \leq n \leq |\mathcal{G}|$ there exists an enumeration of the vertices of $\mathcal{G}$ s.t. the subgraph $\mathcal{G}_i := \mathcal{G}[\{v_1, \ldots, v_i\}]$ is connected.

- *Basic step*:
  For $i = 1$, every possible subgraph $\mathcal{G}_1$ has only a single vertex and is connected by definition. Therefore, $P(1)$ is true.

- *Induction requirement (IR)*:
  Let $P(i)$ be true for some $i \geq 1$.

- *Induction hypothesis (IH)*:
  $P(i) \Rightarrow P(i+1)$

- *Induction step (IS):*
  From (IR) we know that for some $i < |\mathcal{G}|$ we already have an enumeration $v_1, \ldots, v_i$ satisfying $P(i)$, i.e. $\mathcal{G}_i := \mathcal{G}[\{v_1, \ldots, v_i\}]$ is connected.
  Pick any vertex $v \in \mathcal{G} - \mathcal{G}_i$. Due to the connectedness of $\mathcal{G} \Rightarrow \exists$ path $P = v \ldots v_1$.

Choose the newly enumerated vertex v_{i+1} as the last vertex of P in $\mathcal{G} - \mathcal{G}_i$.
Then v_{i+1} has a neighbor in the already connected subgraph $\mathcal{G}_i$.
Therefore $\mathcal{G}_{i+1} := \mathcal{G}[\{v_1, \dots, v_{i+1}\}]$ is connected.
$\Rightarrow P(k+1)$ is true.

$\Rightarrow$ Proposition 1.14 holds. □

Definition 1.15 (Euler-Tour)**.** Let $\mathcal{G}$ be an undirected graph.
An Euler-Tour is closed walk in $\mathcal{G}$ which traverses each edge exactly one.

If such a walk exists, the graph $\mathcal{G}$ is called eularian.

Theorem 1.16 (EULER (1736))**.** Let $\mathcal{G} = (V, E)$ be an undirected, connected graph.

$$\mathcal{G} \text{ is eularian } \Leftrightarrow \forall\, v \in V :\ \deg(v) \text{ is even.}$$

Proof. Let $\mathcal{G} = (V, E)$ be an undirected, connected graph.

($\Rightarrow$) Given an Euler Tour, any vertex appearing k times (or $k+1$ times for the start/end vertex) must have a degree of $2k$.

($\Leftarrow$) W.l.o.g. let all vertices in $\mathcal{G}$ have even degree.
Let $w = v_0 e_0 \dots e_{k-1} v_k$ be the longest possible walk in $\mathcal{G}$ that traverses no edge more than once.

To prove:

- w is a closed walk. Proof by contradiction:

  Assume: w is not a closed walk, i.e. $v_0 \neq v_k$.

  $\Rightarrow$ v_0 is incident to an odd number of edges (in w). ↯ Definition of $\mathcal{G}$
  $\Rightarrow$ since $\deg(v_0)$ is even $\exists\, y \in V :\ e = v_0 y \in E$ which is not in w.
  $\Rightarrow$ we can therefore construct a walk

  $$w' := y e v_0 e_0 \dots v_k$$

  which would be longer than w. ↯ Definition of w

- w traverses all edges of $\mathcal{G}$. Proof by contradiction:

Assume: $\exists\, e \in E$ which is not traversed by w.

W.l.o.g. we can assume that e is not only outside of w but incident to it (i.e. incident to one of its traversed vertices)
Let v_i be a vertex in w incident to e s.t. $\exists\, v \in V : e = uv_i$.
As w is a closed walk we can construct a walk

$$w' := uev_ie_i \dots e_{k-1} \underbrace{v_k}_{=v_0} e_0v_1 \dots e_{i-1}v_i$$

which would be longer than w. ↯ Definition of w

$\Rightarrow$ w is an Euler Tour.

□

2. Algebraic Graph Theory

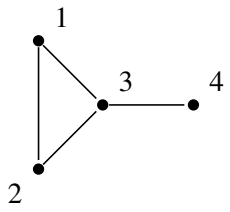
Definition 2.1 (Adjacency Matrix)**.** Let $\mathcal{G} = (V, E)$ be an undirected graph.
The adjacency matrix $A \in \{0,1\}^{|V|\times|V|}$ is defined by

$$\forall\, x, y \in V: \qquad A_{xy} = A_{yx} := \begin{cases} 1 & \text{if } xy \in E, \\ 0 & \text{otherwise.} \end{cases}$$

The entries in the adjacency matrix A for vertices x and y are often also denoted with a_{xy}.

Example 2.1. We again consider our undirected example graph

$$\mathcal{G}_e := (\{1,2,3,4\}, \{\{1,2\},\{1,3\},\{2,3\},\{3,4\}\}).$$



Using the numerical ordering of the vertices we get the graphs adjacency matrix as

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

Definition 2.2 (Degree Matrix)**.** Let $\mathcal{G} = (V, E)$ be an undirected graph.
The degree matrix $D \in \mathbb{N}_0^{|V|\times|V|}$ is defined by

$$\forall\, x, y \in V: \qquad D_{xy} = D_{yx} := \begin{cases} \deg x & \text{if } x = y, \\ 0 & \text{otherwise.} \end{cases}$$

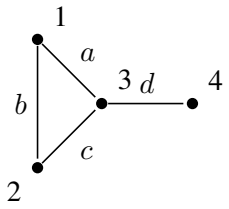
Example 2.2. For our example graph $\mathcal{G}_e$, its degree matrix is

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

Definition 2.3 (Incidence Matrix)**.** Let $\mathcal{G} = (V, E)$ be an undirected graph.
The incidence matrix $H \in \{0, 1\}^{|V|\times|E|}$ is defined by

$$\forall\, x \in V, e \in E: \qquad H_{xe} := \begin{cases} 1 & \text{if } x \in e, \\ 0 & \text{otherwise.} \end{cases}$$

Example 2.3. For our example $\mathcal{G}_e$ with additional edge enumeration via



we get the incidence matrix as

$$H = \begin{pmatrix} H_{1a} & H_{1b} & H_{1c} & H_{1d} \\ H_{2a} & H_{2b} & H_{2c} & H_{2d} \\ H_{3a} & H_{3b} & H_{3c} & H_{3d} \\ H_{4a} & H_{4b} & H_{4c} & H_{4d} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Observation 2.3.1. Let $\mathcal{G} = (V, E)$ be an undirected graph and H the corresponding incidence matrix.

$$\forall\, e \in E : \sum_{v \in V} H_{xe} = 2$$

Observation 2.3.2. Let $\mathcal{G} = (V, E)$ be an undirected graph and H and A the corresponding incidence and adjacency matrix, respectively.

$$\forall\, x \in V : \sum_{e \in E} H_{xe} = \sum_{y \in V} A_{xy} = \sum_{y \in V} A_{yx} = \deg(x)$$

Using these observations, it is interesting to look at the product of the incidence matrix and its transposed version.

Let $\mathcal{G} = (V, E)$ be an undirected graph and A, D, and H the matrices per above notation.

$$\left(HH^T\right)_{xy} = \sum_{e \in E} H_{xe} \cdot H^T_{ey} = \sum_{e \in E} H_{xe} \cdot H_{ye}$$

Let $x, y \in V$.

- if $x \neq y$

$$\left(HH^T\right)_{xy} = \begin{cases} 1 & \text{if } x \text{ and } y \text{ are adjacent,} \\ 0 & \text{otherwise} \end{cases}$$

  $\Rightarrow \left(HH^T\right)_{xy} = A_{xy}$

- if $x = y$

$$\left(HH^T\right)_{xx} = \sum_{e \in E} H_{xe} \cdot H_{xe} = \sum_{e \in E} H^2_{xe} \overset{(*)}{=} \sum_{e \in E} H_{xe} = \deg(x)$$

  $(*)$ due to $H_{xe} \in \{0, 1\}$
  $\Rightarrow \left(HH^T\right)_{xy} = D_{xy}$

Therefore the following holds:

$$HH^T = A + D$$

As a conversion method from undirected to directed graphs we define the concept of an *orientation*:

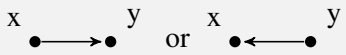
Definition 2.4 (Orientation)**.** Let $\mathcal{G} = (V, E)$ be an undirected graph.
An orientation of $\mathcal{G}$ is a directed graph $\mathcal{G}' = (V, E')$ with

$$\{x, y\} \in E \;\Leftrightarrow\; (x, y) \in E' \text{ or } (y, x) \in E'$$

As a visualization, each undirected edge



gets converted to either



Remark. *Intuitively, for each directed graph $\mathcal{G}$ there also exists a (unique) directed graph $\mathcal{G}'$ with $\mathcal{G}$ being an orientation of $\mathcal{G}'$.*

We can now define the the adjacency and incidence matrix for directed graphs as well.

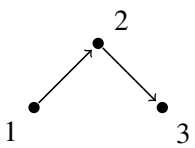
Definition 2.5 (Adjacency Matrix)**.** Let $\mathcal{G} = (V, E)$ be a *directed* graph.
The adjacency matrix $A \in \{0, 1\}^{|V| \times |V|}$ is defined by

$$\forall\, x, y \in V: \qquad A_{xy} := \begin{cases} 1 & \text{if } (y, x) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Note: The entry A_{xy} corresponds to an edge *from y to x*.

Example 2.4. We consider the directed example graph

$$\mathcal{G}_d := (\{1, 2, 3\}, \{(1, 2), (2, 3)\}).$$



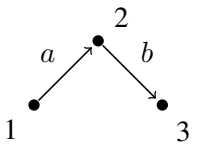
Using the numerical ordering of the vertices we get the graphs adjacency matrix as

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Definition 2.6 (Incidence Matrix)**.** Let $\mathcal{G} = (V, E)$ be a *directed* graph.
The incidence matrix $H \in \{-1, 0, 1\}^{|V| \times |E|}$ is defined by

$$\forall\, x \in V, e \in E: \qquad H_{xe} := \begin{cases} 1 & \text{if } \exists\, y \in V:\ (y, x) \in e, \\ -1 & \text{if } \exists\, y \in V:\ (x, y) \in e, \\ 0 & \text{otherwise.} \end{cases}$$

Example 2.5. For our directed example $\mathcal{G}_d$ with additional edge enumeration via



we get the incidence matrix as

$$H = \begin{pmatrix} H_{1a} & H_{1b} \\ H_{2a} & H_{2b} \\ H_{3a} & H_{3b} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Notation. For most cases, we will simply use the respective symbols for the adjacency and incidence matrix, namely A and H, for directed and undirected graphs.
However, if not clear from the context, e.g., when talking about an undirected graph and its orientation, we will use the symbol

$$A^{\leftrightarrow} \text{ for undirected and } A^{\rightarrow} \text{ for directed graphs.}$$

The notation for the incidence matrix is equivalent.

Using these definitions, it is again interesting to look at the product of the incidence matrix and its transposed version for *directed* graphs (or orientations of undirected graphs).

Let $\mathcal{G} = (V, E)$ be an undirected graph and $\mathcal{G}' = (V, E')$ an orientation of $\mathcal{G}$.
Additionally let $A := A^{\leftrightarrow}$, $D := D^{\leftrightarrow} = D^{\rightarrow}$, and $H := H^{\rightarrow}$ the matrices per above notation.

$$\left(HH^T\right)_{xy} = \sum_{e \in E} H_{xe} \cdot H^T_{ey} = \sum_{e \in E} H_{xe} \cdot H$$

Let $x, y \in V$.

- if $x = y$

$$\left(HH^T\right)_{xx} = \deg_{\text{full}}(x)$$

$\Rightarrow \left(HH^T\right)_{xy} = D_{xy}$

- if $x \neq y$

$$\left(HH^T\right)_{xy} = \begin{cases} 1 & \text{if } x \text{ and } y \text{ are adjacent,} \\ 0 & \text{otherwise} \end{cases}$$

$\Rightarrow \left(HH^T\right)_{xy} = -A_{xy}$

Hence, we can define the following result:

Definition 2.7 (Laplace Matrix)**.** Let $\mathcal{G} = (V, E)$ be an undirected graph and $\mathcal{G}' = (V, E')$ an orientation of $\mathcal{G}$.
The Laplace matrix (or *Laplacian*) of the graph is defined by

$$L := H^{\rightarrow} H^{\rightarrow T} = D^{\leftrightarrow} - A^{\leftrightarrow}.$$

Excursion: Spectral graph theory

Definition 2.8 (Eigenvalues & Eigenvectors)**.** Let $A \in \mathbb{R}^{nxn}$ be a quadratic matrix.
The spectrum of A is the set of all values $\lambda \in \mathbb{C}$ satisfying

$$A \cdot x = \lambda \cdot x \text{ for some vector } 0 \neq x \in \mathbb{R}^n.$$

These values λ are called *Eigenvalues* (abrv. E.V.) and the corresponding vectors x are called *Eigenvectors* (abrv. E.Vectors).

If A is a symmetric matrix, then additionally the following holds:

(1) All Eigenvalues are real ($\lambda \in \mathbb{R}$).

(2) If we have two Eigenvalues $\lambda_1 \neq \lambda_2$ with

$$A \cdot x = \lambda_1 \cdot x \text{ and } A \cdot y = \lambda_2 \cdot y$$

for some vectors $x, y \in \mathbb{R}^n$, then

$$\langle x, y \rangle = \sum_i x_i \cdot y_i = 0.$$

Different Eigenvectors have to be orthogonal to each other.

Remark. *Eigenvalues for a matrix A can be calculated by*

$$\det(A - \lambda \mathrm{I}) = 0$$

which is a polynomial of degree n in λ, called the characteristic polynomial of A.

The only interesting thing for us right now, is that as a result there are at most n different Eigenvalues, and exactly n Eigenvalues when counting multiplicity.
The spectrum of A can therefore be noted as

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$$

Another interesting concept regarding matrices is the quadratic form.

Definition 2.9 (Quadratic Form)**.** Let $A \in \mathbb{R}^{nxn}$ be a quadratic matrix and $\vec{x} \in \mathbb{R}^n$ a vector.
The quadratic form is given by:

$$\langle \vec{x}, A\vec{x} \rangle = \sum_{i,j} \vec{x}_i \cdot A_{ij} \cdot \vec{x}_j = \sum_i \vec{x}_i \cdot (A\vec{x})_i$$

If $\vec{x}$ is an Eigenvector of A, then there exists an E.V. λ s.t. $A\vec{x} = \lambda\vec{x}$ and as a result

$$\langle \vec{x}, A\vec{x} \rangle = \sum_i \vec{x}_i \cdot \lambda\vec{x} = \lambda \cdot \sum_i x_i^2$$

End of Excursion

Proposition 2.10. Every Laplacian of a graph has the vector

$$\mathbf{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

as one of its Eigenvectors.

Proof. Let $L \in (\mathbb{N}_0 \cup \{-1\})^{n \times n}$ be the Laplacian of a graph $\mathcal{G} = (V, E)$.

$$\begin{aligned} \forall\, x \in V: \quad (L \cdot \mathbf{1})_x &= \sum_{y \in V} L_{xy} \cdot \underbrace{(\mathbf{1})_y}_{=1} \\ &= \sum_{y \in V} L_{xy} \\ &= \deg_{\text{full}}(x) + \sum_{x \neq y \in V} L_{xy} \\ &= \deg_{\text{full}}(x) - \deg_{\text{full}}(x) = 0 \end{aligned}$$

$\Rightarrow\ L \cdot \mathbf{1} = 0 \cdot \mathbf{1}.$ □

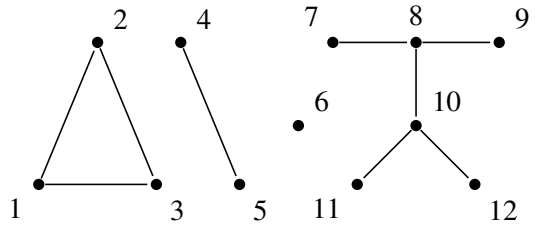
Corollary 2.11. Every Eigenvector $\vec{x}$ with Eigenvalue $\lambda \neq 0$ must satisfy

$$\langle \vec{x}, \mathbf{1} \rangle = 0 = \sum_i x_i \cdot 1 = \sum_i x_i.$$

$\vec{x}$ must be orthogonal to the constant vector $\mathbf{1}$.

Definition 2.12 (Connected Components)**.** Let $\mathcal{G} = (V, E)$ be a graph.
The connected components of $\mathcal{G}$ are the (vertex based) maximal subgraphs of $\mathcal{G}$, s.t. each subgraph is connected.

Example 2.6. Consider the graph $\mathcal{G}$ with



then its connected components would be

- $\mathcal{G}_1 = \mathcal{G}[\{1, 2, 3\}]$,
- $\mathcal{G}_2 = \mathcal{G}[\{4, 5\}]$,
- $\mathcal{G}_3 = \mathcal{G}[\{6\}]$, and
- $\mathcal{G}_4 = \mathcal{G}[\{7, 8, 9, 10, 11, 12\}]$.

Notation (Block-diagonal Form)**.** For any graph $\mathcal{G}$ with connected components, its symmetric matrices (exemplified with the adjacency matrix) can be written (through enumeration of the vertices) as the block-diagonal form using the matrices of its components. For the graph $\mathcal{G}$ from Example 2.6 we get:

$$A_{\mathcal{G}} = \left(\begin{array}{c|c|c|c} A_{\mathcal{G}_1} & 0 & 0 & 0 \\ \hline 0 & A_{\mathcal{G}_2} & 0 & 0 \\ \hline 0 & 0 & A_{\mathcal{G}_3} & 0 \\ \hline 0 & 0 & 0 & A_{\mathcal{G}_4} \end{array}\right)$$

Observation 2.12.1. Let $L_1, \ldots, L_4$ be Laplace-matrices an L the Laplace-matrix resulting through

$$L := \left(\begin{array}{c|c|c|c} L_1 & 0 & 0 & 0 \\ \hline 0 & L_2 & 0 & 0 \\ \hline 0 & 0 & L_3 & 0 \\ \hline 0 & 0 & 0 & L_4 \end{array}\right).$$

Additionally, let $\vec{x}$ be an Eigenvector of L_2 s.t. $\exists\, \lambda : L_2 \cdot \vec{x} = \lambda \cdot \vec{x}$.

$$\left(\begin{array}{c|c|c|c} L_1 & 0 & 0 & 0 \\ \hline 0 & L_2 & 0 & 0 \\ \hline 0 & 0 & L_3 & 0 \\ \hline 0 & 0 & 0 & L_4 \end{array}\right) \cdot \underbrace{\begin{pmatrix} \mathbf{0} \\ \vec{x} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}}_{=\vec{x}'} = \begin{pmatrix} \mathbf{0} \\ L_2 \cdot \vec{x} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} = \lambda \cdot \begin{pmatrix} \mathbf{0} \\ \vec{x} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}$$

Therefore, $\vec{x}'$ is an Eigenvector of L and as a result is λ also an Eigenvalue of L.

Corollary 2.13. Let $\mathcal{G}$ be a graph an L its Laplacian.
If $\mathcal{G}$ has k connected components then L has k orthogonal Eigenvectors with the same Eigenvalue of 0.

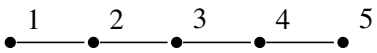
Example 2.7. For two Laplacian L_1, L_2 we get

$$\left(\begin{array}{c|c} L_1 & 0 \\ \hline 0 & L_2 \end{array}\right) \cdot \begin{pmatrix} \mathbf{1} \\ \mathbf{0} \end{pmatrix} = \underbrace{0}_{=\lambda_1} \cdot \begin{pmatrix} \mathbf{1} \\ \mathbf{0} \end{pmatrix} = \underbrace{0}_{=\lambda_2} \cdot \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \end{pmatrix} = \left(\begin{array}{c|c} L_1 & 0 \\ \hline 0 & L_2 \end{array}\right) \cdot \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \end{pmatrix}.$$

We will now take a closer look at quantifiable values for graph connectivity.

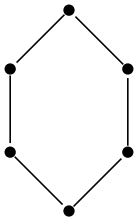
Definition 2.14 (Vertex Cut-Set)**.** Let $\mathcal{G} = (V, E)$ be a graph.
A vertex cut-set K of $\mathcal{G}$ is a set of vertices s.t. removing K leaves $\mathcal{G}[V \setminus K]$ disconnected.

Example 2.8. For the example graph $\mathcal{G}_1$



possible cut-sets are $\{2\}, \{3\}, \{4\}, \{2, 4\}$ with the smallest being of size 1.

The example graph $\mathcal{G}_2$ has a smallest possible cut-set of size 2:



Definition 2.15 (Vertex-Connectivity)**.** Let $\mathcal{G} = (V, E)$ be a graph.
$\mathcal{G}$ is k-vertex-connected if one needs to remove at least k vertices to make $\mathcal{G}$ disconnected or to leave only a single vertex.
We denote the vertex-connectivity as

$$v(\mathcal{G}) = \min\{|S| \mid S \subseteq V \text{ is vertex cut-set of } \mathcal{G}\} \cup \{|V| - 1\}.$$

Similarly, we can define edge cut-sets and edge-connectivity.

Definition 2.16 (Edge Cut-Set)**.** Let $\mathcal{G} = (V, E)$ be a graph.
An edge cut-set K of $\mathcal{G}$ is a set of edges s.t. removing K leaves $\mathcal{G}' = (V, E \setminus K)$ disconnected.

Definition 2.17 (Edge-Connectivity)**.** Let $\mathcal{G} = (V, E)$ be a graph.
$\mathcal{G}$ is k-edge-connected if one needs to remove at least k edges to make $\mathcal{G}$ disconnected.
We denote the edge-connectivity as

$$e(\mathcal{G}) = \min\{|S| \mid S \subseteq E \text{ is edge cut-set of } \mathcal{G}\}.$$

Corollary 2.18. The vertex-connectivity (and also the edge-connectivity) of a disconnected graph is 0.

Proposition 2.19. For every (non-trivial) graph $\mathcal{G}$ the following holds:

$$v(\mathcal{G}) \leq e(\mathcal{G})$$

vertex-connectivity is lower or equal to edge-connectivity

Proof, adapted from Diestel (1997). Let $\mathcal{G} = (V, E)$ be a (non-trivial) graph and let $F \subseteq E$ be any edge cut-set s.t. $\mathcal{G}' = (V, E \setminus F)$ is disconnected, i.e. $e(\mathcal{G}) = |F|$.
We show that $v(\mathcal{G}) \leq |F|$.

- Suppose first that $\mathcal{G}$ has a vertex $v \in V$ that is not incident with an edge in F. Let $C \subseteq V$ be the connected component of $\mathcal{G}'$ containing v.
  Then the vertices of C that are incident with an edge in F separate v from $\mathcal{G}[V \setminus C]$, i.e. $C \setminus v$ is a vertex cut-set of $\mathcal{G}$.
  Since, by minimality of F, no edge in F has both ends in C, there are at most $|F|$ such vertices giving $v(\mathcal{G}) \leq C \setminus v \leq |F|$.

- Suppose now that every vertex of $\mathcal{G}$ is incident with an edge in F.
  Let $v \in V$ be any vertex and let $C \subseteq V$ be the component of $\mathcal{G}'$ containing v.
  Then the neighbors w of v with $vw \notin F$ lie in $\mathcal{G}[C]$ and are incident with distinct edges in F, giving $\deg_{\mathcal{G}}(v) \leq |F|$.
  As $N_{\mathcal{G}}(v)$ separate v from all other vertices in $\mathcal{G}$, this yields $v(\mathcal{G}) \leq N_{\mathcal{G}}(v) \leq |F|$, unless there are no other vertices, i.e. unless $N_{\mathcal{G}}(v) \cup v = V$.
  However, as v was an arbitrary vertex, we may assume that $\mathcal{G}$ is complete, giving $v(\mathcal{G}) = e(\mathcal{G}) = |V| - 1$.

□

Definition 2.20 (Algebraic-Connectivity)**.** Let $\mathcal{G}$ be a graph and L its Laplacian.
The 2nd-smallest Eigenvalue λ_2 of L is called the algebraic-connectivity of $\mathcal{G}$

$$a(\mathcal{G}) := \lambda_2.$$

Lemma 2.21. Let $\mathcal{G}$ be a graph and L its Laplacian.
All Eigenvalues of L are non-negative.

Proof. Let L be a graph Laplacian and $\mathcal{G} = (V, E)$ a corresponding directed graph with its incidence matrix H.
Let λ be an E.V. with $\vec{x}$ as its E.Vector.
From Definition 2.9 we know

$$\langle \vec{x}, L\vec{x} \rangle = \lambda \cdot \underbrace{|\vec{x}|^2}_{\geq 0}$$

and therefore

$$\begin{aligned}\langle \vec{x}, L\vec{x} \rangle = \langle \vec{x}, HH^T\vec{x} \rangle = \langle \vec{x} \cdot H \cdot H^T \cdot \vec{x} \rangle \\ &= \sum_{e \in E} \left(\underbrace{\left(\sum_i \vec{x}_i \cdot H_{ie} \right)}_{b_e :=} \cdot \underbrace{\left(\sum_i \vec{x}_i \cdot H_{ie} \right)}_{= b_e} \right) \\ &= \sum_{e \in E} b_e^2 \geq 0\end{aligned}$$

□

Lemma 2.22. Let $\mathcal{G}$ be a graph.

$$a(\mathcal{G}) = 0 \Leftrightarrow \mathcal{G} \text{ is not connected.}$$

Lemma 2.22 stated here without proof.

Theorem 2.23. For all graphs $\mathcal{G}$ holds:

$$a(\mathcal{G}) \leq v(\mathcal{G})$$

Before we prove Theorem 2.23, we first introduce some smaller results.

Proposition 2.24. Let $\mathcal{G}_1 = (V, E_1), \mathcal{G}_2 = (V, E_2)$ be two edge-disjoint graphs on the same vertex set.
There is a well-defined union graph

$$\mathcal{G}_1 \cup \mathcal{G}_2 := (V, E_1 \cup E_2)$$

with

$$a(\mathcal{G}_1 \cup \mathcal{G}_2) \geq a(\mathcal{G}_1) + a(\mathcal{G}_2).$$

Proof. Let $\mathcal{G}_1 = (V, E_1), \mathcal{G}_2 = (V, E_2)$ be two edge-disjoint graphs on the same vertex set. Furthermore, let L_i be the graph Laplacian of $\mathcal{G}_i$ for $i \in \{1, 2\}$.
We take the following result as already proven:

For the second-smallest Eigenvalue λ_2 of a graph Laplacian L, the following holds:

$$\lambda_2 = \min_{\vec{x} \in W} \langle \vec{x}, L\vec{x} \rangle$$

with

$$W := \{\vec{x} \mid 1 = |\vec{x}|^2 \ \wedge \ \langle \vec{x}, \mathbf{1} \rangle = 0\}.$$

The construction of W thereby ensures that if $\vec{x}$ is an Eigenvector of L we get

$$\langle \vec{x}, L\vec{x} \rangle = \lambda \cdot |\vec{x}|^2 = \lambda.$$

We can then conclude

$$\begin{aligned} a(\mathcal{G}_1 \cup \mathcal{G}_2) &= \min_{x \in W} \{\langle \vec{x}, \left(L(\mathcal{G}_1) \overset{(*)}{+} L(\mathcal{G}_2) \right) \vec{x} \rangle\} \\ &= \min_{x \in W} \{\langle \vec{x}, L_1\vec{x} \rangle + \langle \vec{x}, L_2\vec{x} \rangle\} \\ &\geq \min_{x \in W} \{\langle \vec{x}, L_1\vec{x} \rangle\} + \min_{x \in W} \{\langle \vec{x}, L_2\vec{x} \rangle\} \\ &= a(\mathcal{G}_1) + a(\mathcal{G}_2) \end{aligned}$$

$(*)$ valid, as $\mathcal{G}_1$ and $\mathcal{G}_2$ are edge-disjoint □

Corollary 2.25. If $\mathcal{G} = (V, E), \mathcal{G}' = (V, E')$ are two graphs on the same vertex set and $E' \subseteq E$ then

$$a(\mathcal{G}') \leq a(\mathcal{G}).$$

Now onto the actual proof of Theorem 2.23:

Proof. Let $\mathcal{G} = (V, E)$ be a graph with $|V| = n$.
Let $v \in V$ be an arbitrary vertex.

Let $\mathcal{G}_1 = (V_1, E_1) = \mathcal{G}[V \setminus \{v\}]$ be the graph $\mathcal{G}$ after removing vertex v and, additionally, let $\mathcal{G}'$ be the graph retrieved after connecting v to all vertices in $\mathcal{G}_1$

$$\mathcal{G}' := (V, E_1 \cup \{vx \mid x \in V_1\}).$$

From Corollary 2.25 we get

$$a(\mathcal{G}) \leq a(\mathcal{G}').$$

Let A, D be the undirected adjacency matrix and degree matrix of $\mathcal{G}_1$ respectively.
Futhermore, let L_1 be the Laplacian of $\mathcal{G}_1$ and L the Laplacian of $\mathcal{G}'$.

$$L = \left(\begin{array}{cccc:c} & & & & -1 \\ & & & & -1 \\ & 1+D-A & & & -1 \\ & & & & -1 \\ & & & & \vdots \\ \hdashline -1 & -1 & -1 & -1 & \dots & n-1 \end{array}\right) = \left(\begin{array}{c:c} L_1 + \mathrm{I} & -\mathbf{1} \\ \hdashline -\mathbf{1}^T & n-1 \end{array}\right)$$

Let $\vec{x}$ be an E.Vector of L_1 with E.V. $\lambda^{(1)}$:

$$L_1 \cdot \vec{x} = \lambda^{(1)} \cdot \vec{x}.$$

We claim that $\begin{pmatrix} \vec{x} \\ 0 \end{pmatrix}$ is an E.Vector of L with E.V. $\lambda^{(1)} + 1$.

Proof.

$$\begin{aligned} L \cdot \begin{pmatrix} \vec{x} \\ 0 \end{pmatrix} &= \left(\begin{array}{c:c} L_1 + \mathrm{I} & -\mathbf{1} \\ \hdashline -\mathbf{1}^T & n-1 \end{array}\right) \cdot \begin{pmatrix} \vec{x} \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} (L_1 + \mathrm{I}) \cdot \vec{x} + (-\mathbf{1}) \cdot 0 \\ -\mathbf{1}^T \cdot \vec{x} + (n-1) \cdot 0 \end{pmatrix} \\ &= \begin{pmatrix} L_1 \cdot \vec{x} + \vec{x} \\ 0 \end{pmatrix} \\ &= (\lambda^{(1)} + 1) \cdot \begin{pmatrix} \vec{x} \\ 0 \end{pmatrix} \end{aligned}$$

This results in $a(\mathcal{G}_1) \leq a(\mathcal{G}') - 1$.
$\Rightarrow\ a(\mathcal{G}) - 1 \leq a(\mathcal{G}_1)$.

This concludes that if you remove k vertices from a graph, its algebraic-connectivity goes down by at most k. However, if a graph is disconnected, its algebraic connectivity is 0.
$\Rightarrow\ a(\mathcal{G}) \leq v(\mathcal{G})$. □

Remark. *For completion, the connection between the edge-connectivity and the algebraic-connectivity is*

$$e(\mathcal{G}) \cdot \left(1 - \cos\left(\frac{\pi}{|V|}\right)\right) \leq a(\mathcal{G})$$

for a graph $\mathcal{G} = (V, E)$.

Definition 2.26 (Fiedler-Vector)**.** Let $\mathcal{G}$ be a graph with its Laplacian L.
The vector $\vec{x}$ with

$$L \cdot \vec{x} = a(\mathcal{G}) \cdot \vec{x}$$

is called the Fiedler-Vector.

3. Cycles in Graphs

We expand the definition of cycles (1.9) by

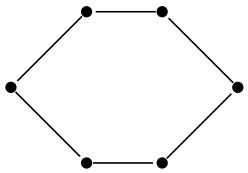
Definition 3.1. Cycles can be distinguished into

(1) elementary (or simple) cycles,
- closed path (original definition)
- each contained node has a degree of 2

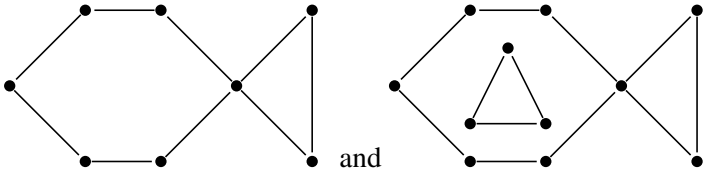
(2) and generalized cycles.
- edge-disjoint union of elementary cycles
- each contained node has an even degree

Example 3.1. Using Definition 3.1 we can distinct

- elementary cycles,

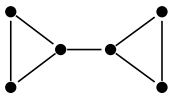


- generalized cycles,



and

- and walks, that are neither elementary or generalized cycles.



Remark. *Every Euler-Tour is a generalized cycle and can therefore be decomposed into edge-disjoint elementary cycles.*

To conceptualize the symmetric difference of edge sets as a way of joining cycles, we first introduce the Semiring $S = (\{0,1\}, \oplus, \odot, 0, 1)$ with the binary operators defined through

$\oplus$	0	1
0	0	1
1	1	0

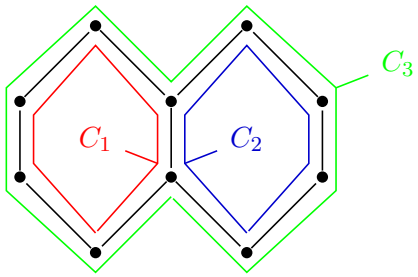
and

$\odot$	0	1
0	0	0
1	0	1

The additive operation is thereby identical to "XOR".

Defining (generalized) cycles C_1 and C_2 with their respective edge sets E_1 and E_2 we can create a new (generalized) cycle C_3, spanning the vertices of both cycles, via

$$E_1 \oplus E_2 = E_3 = (E_1 \cup E_2) \setminus (E_1 \cap E_2).$$



Corollary 3.2. Every sum ($\oplus$) of two generalized cycles is again a generalized cycle.

This allows for the increase of a cycles span (and the later concept of *cycle bases*), through XOR-union of smaller cycles' edge sets.

Observation 3.2.1. Let $\mathcal{G} = (V, E)$ be an undirected graph and H its incidence matrix. As H is also defined over the set $\{0,1\}$, in combination with the characteristic vector $\chi(C)$ of a cycle C, being defined through

$$\forall\, e \in E:\ \chi(C)_e = \begin{cases} 1 & if\, e \in C, \\ 0 & \text{otherwise,} \end{cases}$$

we get (under the foundation of the Semiring S)

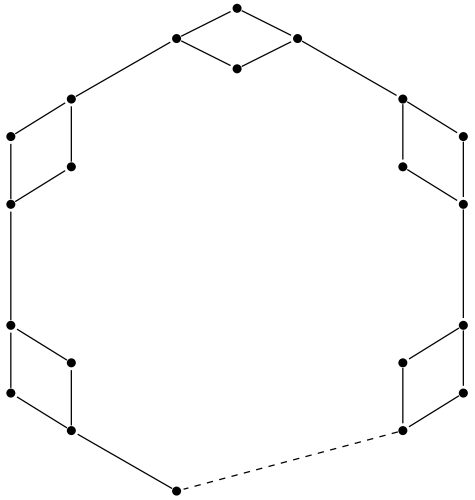
$$\forall\, x \in V:\ \bigoplus_{e \in E} H_{xe} \cdot \chi(C)_e = 0.$$

Therefore, we can conclude that every subset $C \subseteq E$ is a cycle if and only if

$$H \cdot \chi(C) = \mathbf{0}.$$

Observation. Let $\mathcal{G} = (V, E)$ be a graph with $m = |E|$.
As each unique cycle C as a corresponding unique representation through its unique, characteristic vector $\chi(C) \in \{0,1\}^m$, the total number of generalized cycles in $\mathcal{G}$ is bound to be lower or equal to $2^m = |\{0,1\}^m|$.

Example 3.2 (Neclace-graph)**.** We consider the undirected example graph $\mathcal{G}_N = (V, E)$ with k "diamonds" and $n = |V| = 4k$:



$\Rightarrow$ $\mathcal{G}_N$ has a maximum of $2^k + k$ cycles.

Definition 3.3 (Generating Set)**.** Let $\mathcal{G} = (V, E)$ be a graph.
A *set of cycles* G is a generating set if for every cycle D in $\mathcal{G}$ holds

$$\exists\, G' \subseteq G : \quad \bigoplus_{C \in G'} = D.$$

D can be constructed through the sum ($\oplus$) over a subset of G

(see Example 3.3 for visualization)

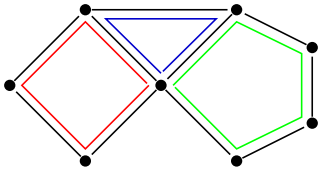
Instead of a subset, we can construct a cycle D through a *selection* vector

$$(\lambda_1, \lambda_2, \ldots, \lambda_{|G|}) \in \{0, 1\}^{|G|}$$

such that

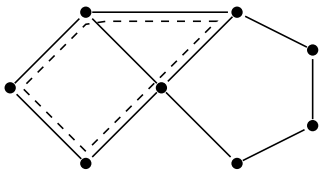
$$D = \bigoplus_{C \in G} \lambda_C \cdot C.$$

Example 3.3. Consider the following graph $\mathcal{G}$ with a generating set G of three elementary cycles C_1, C_2, and C_3:

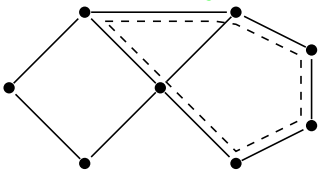


All other (non-trivial) cycles can be generated as follows:

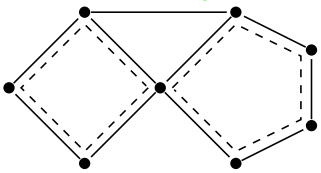
$C_1 \oplus C_2$



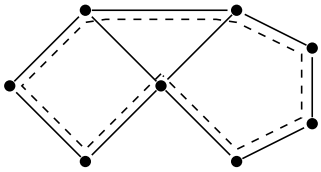
$C_1 \oplus C_3$



$C_2 \oplus C_3$



$C_1 \oplus C_2 \oplus C_3$



3.1. Linear Independence & Cycle bases

Definition 3.4 (Linear Independence)**.** Let $\mathcal{J}$ be a set of cycles.
$\mathcal{J}$ is linear independent (l.i.), if no cycle $D \in \mathcal{J}$ can be written as the sum ($\oplus$) of some other cycles in $\mathcal{J} \setminus \{D\}$:

$$\forall\, (\lambda_1, \ldots, \lambda_{|\mathcal{J}|}) \in \{0,1\}^{|\mathcal{J}|} : \quad \left(\bigoplus_{C \in \mathcal{J}} \lambda_C \cdot C = 0 \right) \Rightarrow (\forall\, C \in \mathcal{J} :\ \lambda_C = 0)$$

Remark. *The expression in Definition 3.4 is a sufficient description of linear independence.*
Consider the set of cycles G with $D \in G$. If G is not linear independent, i.e. there exists $G' \subseteq G \setminus \{D\}$ s.t.

$$\bigoplus_{C \in G'} \underbrace{1}_{\lambda_C} \cdot C = D$$

then we get

$$\left(\bigoplus_{C \in G'} \underbrace{1}_{\lambda_C} \cdot C \right) \oplus \underbrace{1}_{\lambda_D} \cdot D = D \oplus D = 0$$

violating the denoted expression with $1 = \lambda_C \neq 0$ for each $C \in G' \cup \{D\}$.

Definition 3.5 (Cycle Basis)**.** A cycle basis is a minimal generating set of cycles if and only if it is a maximal linear independent set of cycles.

Before discussing a method of constructing a cycle basis, we first need to (briefly) introduce *spanning trees*.

Definition 3.6 (Trees & Forests)**.** An *acyclic* graph $\mathcal{F}$, i.e. one not containing any cycles is called a forest.

If $\mathcal{F}$ is also a connected graph, then it is called a tree.

Corollary 3.7. A connected graph with n vertices is a tree $\Leftrightarrow$ it has $n-1$ edges.

Definition 3.8 (Spanning Trees)**.** Let $\mathcal{G} = (V, E)$ be a connected graph.
Let $T = (V', E')$ with $V' \subseteq V$ and $E' \subseteq E$ be a tree in $\mathcal{G}$.
We call T a spanning tree of $\mathcal{G}$ if $V' = V$.

Remark. *Every connected graph contains a (normal) spanning tree.*

Construction 3.9 (KIRCHHOFF-Basis)**.** Let $\mathcal{G} = (V, E)$ be a connected graph.
Construct a cycle basis $\mathcal{B}$ as follows:

1) Construct a *spanning tree* $T = (V, E')$ in $\mathcal{G}$.
2) For each edge $e \in E \setminus E'$ consider $T_\cup = (V, E' \cup \{e\})$.
   Claim: $T_\cup$ contains a unique cycle.

   > Let $x, y \in V$ s.t. $e = xy$. Then x and y are connected by a *unique* path P_{xy} in T (due to connectedness of trees).
   > $\Rightarrow$ $P_{xy} \cup e$ is a unique cycle

   We denote this cycle as $\mathrm{cyc}(T, e) := P_{xy} \cup e$.
3) $\mathcal{B} = \{\mathrm{cyc}(T, e) \mid e \in E \setminus E'\}$

$\mathcal{B}$ is l.i. because for every cycle $\mathrm{cyc}(T, e)$ the edge e is by construction not present in any other cycle $\mathrm{cyc}(T, e')$ with $e \neq e'$.

Let $n := |V|$ and $m := |E|$. The cardinality of $\mathcal{B}$ is

$$|\mathcal{B}| = |E \setminus E'| = |E| - |E'| = m - (|V| - 1) = \underbrace{m}_{\leq n^2} - n + 1$$

This is called the *cyclomatic/ Euler-/ 1st Betti number* $\mu(\mathcal{G}) := |\mathcal{B}|$.

Observation 3.9.1. Let $\mathcal{G} = (V, E)$ be a graph with k connected components

$$\mathcal{G}_1 = (V_1, E_1), \mathcal{G}_2 = (V_2, E_2), \ldots, \mathcal{G}_k = (V_k, E_k).$$

Due to the definition of connected components we know that the vertex sets V_i and edge sets E_i are pairwise disjoint, full partitions of V and E respectively. Thus

$$\mu(\mathcal{G}) = \sum_{i=1}^{k} \mu(\mathcal{G}_i) = \sum_{i=1}^{k} (|E_i| - |V_i| + 1) = |E| + |V| + k$$

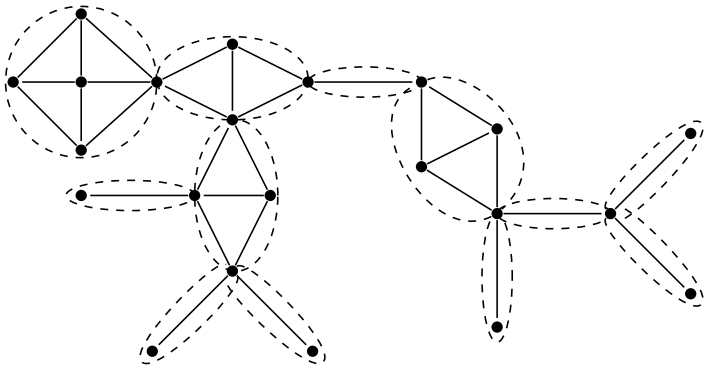
Definition 3.10 (2-Connectedness)**.** Let $\mathcal{G}$ be a graph.
$\mathcal{G}$ is 2-connected if and only if every two vertices are contained in a cycle.

Definition 3.11 (Block)**.** A block in a graph is a maximal 2-connected subgraph or a singular edge, that is not contained in any cycle.

The decomposition of the edge set of a graph into blocks is unique.

Observation. All elementary/simple cycles are contained in a single block.

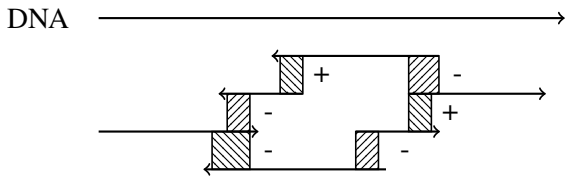
Example 3.4. The following example graph contains 2-connected subgraphs. Dashed outlines mark the boundaries of each respective block.



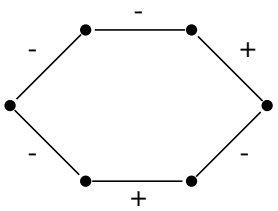
Observation. Every cycle basis consisting of elementary cycles is the union of cycle bases for the respective block of the graph.

Example. One use case for cycle bases in the bioinformatics is for *genome assembly*.

A sequencing result from a double-stranded DNA region (boxes represent overlaps)

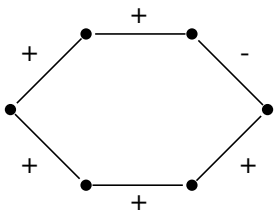


can be converted in a corresponding *signed overlap graph*



with:
- vertex $\hat{=}$ read
- edge $\hat{=}$ overlap
- sign $\hat{=}$ relative reading direction

However, for an alignment to be correct it the changes in read direction must cancel out over the entire alignment, i.e. an overlap graph such as



is not possible and would indicate a problem in the dataset (false-positive overlap).
As these 'impossible' alignments are represented through cycles with odd parity, genome assemby algorithms could use cycle parity as an easy-to-compute filtering step.
This could be done using the Kirchhoff construction in linear time and check the basis cycles.

In praxis, genome assembler construct a min-/max-weight spanning tree where the weight function represents some quality measure of the overlaps.

In addition to the Kirchhoffbasis, we now introduce another approach for constructing bases: ear cycle bases.

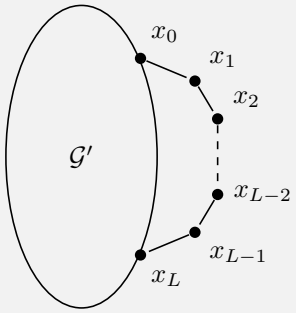
Construction 3.12 (Ear Decomposition)**.** Let $\mathcal{G} = (V, E)$ be a (at least) 2-connected graph.
An ear of $\mathcal{G}$ is a path P from x_0 to x_L (thus of length L) in $\mathcal{G}$ s.t.

(i) $x_0 \neq x_L$,

(ii) $\forall\, 0 < k < L :\ \deg(x_k) = 2$, and

(iii) $\deg(x_0), \deg(x_L) > 2$.

Example: Ear from x_0 to x_L and remaining subgraph $\mathcal{G}'$:



Removal of an ear means:

a) Removal of every inner vertex $x_i\ \forall\, 0 < i < L$ of the ear

b) if $L = 1$: Remove edge x_0x_L

Therefore, removal of an ear results in the subgraph

$$\mathcal{G}' = (V \setminus \{x_i \mid 0 < i < L\}, E \setminus P)$$

An ear decomposition is a sequence of graphs $(\mathcal{G}_i)$ and respective removals of ears (O_i) as follows:

$$(\mathcal{G}_0, O_1, \mathcal{G}_1, O_2, \mathcal{G}_2, \ldots, O_N, \mathcal{G}_N)$$

with the conditions:

(i) $\mathcal{G} = \mathcal{G}_0$,

(ii) $\mathcal{G}_{i+1}$ is the result of removing ear O_{i+1} from $\mathcal{G}_i$,

(iii) every $\mathcal{G}_i$ is 2-connected, and

(iv) $\mathcal{G}_N$ is an elementary cycle.

For the correctness of above construction, we state the following theorem without proof:

Theorem 3.13. Every 2-connected graph has an ear decomposition.

Definition 3.14 (Ear cycle basis)**.** Let $\mathcal{G}$ be a 2-connected graph with an ear decomposition

$$(\mathcal{G}, O_1, \mathcal{G}_1, O_2, \mathcal{G}_2, \ldots, O_N, \mathcal{G}_N)$$

and additionally let C_{N+1} denote the remaining elementary cycle in $\mathcal{G}_N$.
For every (ear) path O_i between nodes $x_0^{(i)}$ and $x_L^{(i)}$, let P_i be a path in $\mathcal{G}_i$ between nodes $x_L^{(i)}$ and $x_0^{(i)}$ (for all $1 \leq i \leq N$). As $\mathcal{G}_i$ is by definition 2-connected, we know that one such path P_i exists.

We define an ear cycle basis of $\mathcal{G}$ as

$$\mathcal{B}_0 = (\underbrace{P_1 \cup O_1}_{C_1}, \underbrace{P_2 \cup O_2}_{C_2}, \ldots, \underbrace{P_N \cup O_N}_{C_N}, C_{N+1}).$$

It leaves to show that $\mathcal{B}_0$ is really a cycle basis. For this matter, we have to prove that

(i) $\mathcal{B}_0$ is l.i. and

(ii) $|\mathcal{B}_0| = \mu(\mathcal{G})$.

Proof. Let $\mathcal{G} = (V, E)$ be a 2-connected graph with an ear cycle basis

$$\mathcal{B}_0 = (\underbrace{P_1 \cup O_1}_{C_1}, \underbrace{P_2 \cup O_2}_{C_2}, \ldots, \underbrace{P_N \cup O_N}_{C_N}, C_{N+1}).$$

- We first prove (i):
  By construction O_1 contains edges that are not in $\mathcal{G}_1$ and thereby not in any other cycle.
  $\Rightarrow$ $C_1 = P_1 \cup O_1$ cannot be constructed by a combination of any other cycles $C_2, \ldots, C_{N+1}$.
  Thus, C_1 is l.i. from all other (following) cycles.
  $\Rightarrow$ by setting $\mathcal{G} = \mathcal{G}_i$ for any $1 \leq i < N$, the same follows by induction.
  $\Longrightarrow$ $\mathcal{B}_0$ is l.i.

- To prove (ii) we have to show that

  $$|\mathcal{B}_0| = N + 1 = \mu(\mathcal{G}) = |E| - |V| + 1$$

  holds. We recall that by construction the removal of ear O_1 with length L from $\mathcal{G}$ creates $\mathcal{G}_1$ as follows:

  $$\mathcal{G}_1 = (\underbrace{V \setminus \{x_i \mid 0 < i < L\}}_{V_1 :=}, \underbrace{E \setminus O_1}_{E_1 :=})$$

  We therefore know

  $$|V_1| = |V| - (L - 1) = |V| - L + 1 \text{ and } |E_1| = |E| - L$$

  which concludes to

  $$\begin{aligned} \mu(\mathcal{G}_1) = |E_1| - |V_1| + 1 &= |E| - L - (|V| - L + 1) + 1 \\ &= |E| - L - |V| + L - 1 + 1 \\ &= (|E| - |V| + 1) - 1 \\ &= \mu(\mathcal{G}) - 1 \end{aligned}$$

  $\Rightarrow$ by induction, we get that removal of 1 ear decreases $\mu(\mathcal{G})$ by 1.
  $\Rightarrow$ with N removals we thus have $\mu(\mathcal{G}_N) = \mu(\mathcal{G}) - N$.
  As $\mathcal{G}_N$ consists of a singular, elementary cycle C_{N+1} we also know that $\mu(\mathcal{G}_N) = 1$.

  $$\begin{aligned} \mu(\mathcal{G}_N) &= \mu(\mathcal{G}) - N \\ 1 &= \mu(\mathcal{G}) - N && \mid + N \\ N + 1 &= \mu(\mathcal{G}) \end{aligned}$$

Therefore, $\mathcal{B}_0$ is indeed a cycle basis. □

3.2. Fundamental & Strictly Fundamental Cycle Bases

Definition 3.15 (Strictly Fundamental)**.** Let $\mathcal{G}$ be a graph and $\mathcal{B} = (C_1, \ldots, C_{\mu(\mathcal{G})})$ a cycle basis of $\mathcal{G}$.
$\mathcal{B}$ is strictly fundamental if for every order of its cycles $C'_1, \ldots, C'_{\mu(\mathcal{G})}$

$$C'_k \setminus \bigcup_{i=1}^{k-1} C'_i \neq \emptyset$$

holds for all $1 \leq k \leq \mu(\mathcal{G})$.

Theorem 3.16. If $\mathcal{B}$ is a Kirchhoffbasis then $\mathcal{B}$ is strictly fundamental.

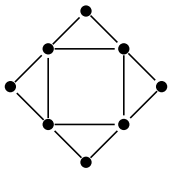
Proof. Let $\mathcal{G} = (V, E)$ be a connected graph, $T = (V, E')$ a spanning tree in $\mathcal{G}$ and

$$\mathcal{B} = \{\mathrm{cyc}(T, e) \mid e \in E \setminus E'\}$$

a Kirchhoffbasis of $\mathcal{G}$ as by construction.

We know that each cycle $\mathrm{cyc}(T, e)$ in $\mathcal{B}$ contains the corresponding edge e which (by construction) cannot be part of any other cycle in $\mathcal{B}$.
Therefore, $\mathcal{B}$ is strictly fundamental. □

Example 3.5. Consider the following graph $\mathcal{G}$.



A possible ear cycle basis $\mathcal{B}$ with

$$\mathcal{B} = \{\triangleleft, \triangle, \triangleright, \triangledown, \square\} \text{ and also } \triangleleft \oplus \triangle \oplus \triangleright \oplus \triangledown = \mathcal{G}$$

results in

$$\square \setminus (\triangleleft \cup \triangle \cup \triangleright \cup \triangledown) = \emptyset.$$

Observation 3.16.1. As seen from Example 3.5, $\mathcal{B}$ is an ear cycle basis which is not strictly fundamental.
$\Rightarrow$ Ear cycle bases in general are not strictly fundamental.

Definition 3.17 (Weakly fundamental)**.** Let $\mathcal{G}$ be a graph and $\mathcal{B} = (C_1, \ldots, C_{\mu(\mathcal{G})})$ a cycle basis of $\mathcal{G}$.
$\mathcal{B}$ is weakly fundamental if *any* order of its cycles $C'_1, \ldots, C'_{\mu(\mathcal{G})}$ exists s.t.

$$C'_k \setminus \bigcup_{i=1}^{k-1} C'_i \neq \emptyset$$

holds for all $1 \leq k \leq \mu(\mathcal{G})$.

Sometimes this property is simply called *fundamental*.

Proposition 3.18. Every ear cycle basis is weakly fundamental.

Proof. Let $\mathcal{G}$ be a 2-connected graph and $\mathcal{B} = (C_1, \ldots, C_{\mu(\mathcal{G})})$ an ear cycle as by Definition 3.14.
When we reverse the order of the cycles

$$\begin{array}{ccccc} C_{\mu(\mathcal{G})}, & C_{\mu(\mathcal{G})}, & \ldots & C_1 \\ \| & \| & & \| \\ C'_1, & C'_2, & \ldots & C'_{\mu(\mathcal{G})} \end{array}$$

every C'_k has edges which are by definition not present in any C'_i for $1 \leq i < k$. □

3.3. Minimal Cycle Bases & Canonical Greedy

Definition 3.19. Let $\mathcal{G} = (V, E)$ be a graph, C a cycle in $\mathcal{G}$, $\mathcal{B}$ a cycle basis of $\mathcal{G}$, and

$$w : \ E \to \mathbb{R}_{\geq 0}$$

a positive weight function. The base case for the weight function is $e \mapsto 1$ for all $e \in E$.

- The cardinality of the edge set of C is defined as

$$|C| := \sum_{e \in C} w(e).$$

- The length of the basis $\mathcal{B}$ is denoted as

$$\text{len}(\mathcal{B}) := \sum_{C \in \mathcal{B}} |C|.$$

A cycle basis is called minimal, if for every other cycle basis $\mathcal{B}'$ of $\mathcal{G}$ holds:

$$\text{len}(\mathcal{B}) \leq \text{len}(\mathcal{B}')$$

Excursion: Matroids

Definition 3.20 (Matroid)**.** A matroid M is a tuple $(X, \mathcal{I})$ with

- a base set X (for us: set of all cycles),
- a set of linear independent sets $\mathcal{I} \subseteq 2^X$ (all l.i. sets of cycles), and
- properties
  - (i) $\emptyset \in \mathcal{I}$
  - (ii) $A \in \mathcal{I} \ \wedge \ B \subseteq A \Rightarrow B \in \mathcal{I}$ und
  - (iii) $A, B \in \mathcal{I}$ mit $|B| < |A| \Rightarrow \exists\, x \in A \setminus B : \ B \cup \{x\} \in \mathcal{I}$.
    (Multiple-Exchange-Property, ger: “Auswechseleigenschaft”)

End of Excursion

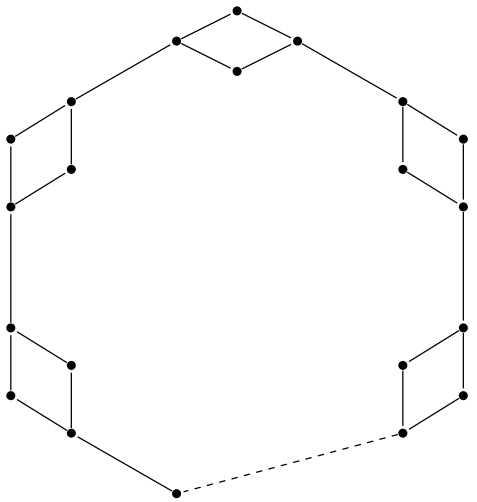
Theorem 3.21 (by KRUSKAL)**.** $M = (E, \mathcal{I})$ is a matroid if and only if the *canonical greedy algorithm* yields an optimal solution for any (positive) weight function.

Algorithm 3.1 (Canonical Greedy Algorithm)**.** Let $\mathcal{G} = (V, E)$ be a graph with k connected components and $M = (X, \mathcal{I})$ the matroid over all cycles in $\mathcal{G}$.

(i) order input X increasingly by cardinality of cycles (see Definition 3.19)

(ii) $B \leftarrow \emptyset$
```
while X ≠ ∅ do
    C ← dequeue(X)
    if (B ∪ {C}) ∈ I then
        B ← B ∪ {C}
    break if |B| = |E| − |V| + k
return B
```

Example 3.6. As a reminder: The "naclace-graph" $\mathcal{G}_N = (V, E)$ with k "diamonds" and $n = |V| = 4k$



has 2^n cycles with length $3n$.
$\Rightarrow$ # of cycles grows exponentially.

The canonical greedy algorithm (CGA) in theory guarantees an optimal solution, however, the required workload for indexing and ordering all cycles makes this approach not feasible for application. Instead we look at other distinctive properties of cycles.

Definition 3.22 (Relevant Cycles)**.** Let $\mathcal{G}$ be a graph.
A cycle C in $\mathcal{G}$ is called relevant if it is contained in at least one minimal cycle basis (MCB).

Definition 3.23 (Essential Cycles)**.** Let $\mathcal{G}$ be a graph.
A cycle C in $\mathcal{G}$ is called essential if it is contained in every MCB.

Proposition 3.24. Given cycles C, C_1, C_2 with $|C| > |C_1|, |C_2|$.

$$C = C_1 \oplus C_2 \;\Rightarrow\; C \text{ is not relevant}$$

Proof. Let $\mathcal{G}$ be a graph and C, C_1, C_2 cycles in $\mathcal{G}$ with $|C| > |C_1|, |C_2|$.
Due to their difference in cardinality, C_1 and C_2 are chosen before C in the CGA.
(C_1, C_2 could also be the sum of even shorter cycles, which still results in linear dependency)
$\Rightarrow$ C cannot be added due to l.i.-test failing. □

This claim can be further generalized:

Corollary 3.25. Let $\mathcal{G}$ be a graph.
A cycle C is not relevant (in $\mathcal{G}$) if it can be constructed from smaller cycles.

$$C = \bigoplus_{i=1}^{n} C_i \text{ with } C_i \in \mathcal{B}$$

for any cycle basis $\mathcal{B}$ (of $\mathcal{G}$).

This conclusion opens up the question about the properties of a cycle, that cannot be constructed as a sum ($\oplus$) of shorter cycles.

Observation 3.25.1. Let $\mathcal{G}$ be a graph and C a cycle in $\mathcal{G}$ that cannot be constructed from shorter cycles.
We modify the CGA (Algorith 3.1) as follows:

- order input X in step (i) s.t. C is the first cycle of length $|C|$

$\Rightarrow$ $\forall\, C'$ before C in $X : |C'| < |C|$
$\Rightarrow$ C is linear independent to the previously constructed $\mathcal{B}$ (i.e. linear independence test

succeeds) and is therefore added to $\mathcal{B}$.
$\Rightarrow$ C is in at least one MCB and thus relevant.

Theorem 3.26. Let $\mathcal{G}$ be a graph and C a cycle in $\mathcal{G}$.

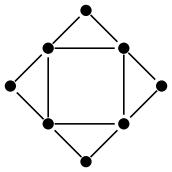
a) C is relevant if and only if C cannot be constructed as the sum ($\oplus$) of shorter cycles.

b) If C is relevant and also the only cycle of size $|C|$, then C is essential.

Proof. Trivial with Observation 3.25.1 and definition of the CGA. □

Example 3.7. Consider again the graph $\mathcal{G} = (V, E)$ with $|V| = 8$, $|E| = 12$, and

$$\mu(\mathcal{G}) = 12 - 8 + 1 = 5.$$

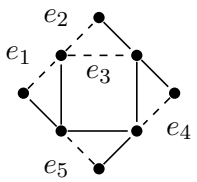


The noted ear cycle basis

$$\mathcal{B} = \{\triangleleft, \triangle, \triangleright, \triangledown, \square\}$$

is the only MCB of $\mathcal{G}$.

However, considering the spanning tree $T = (V, E')$ (full lines) with the separate edges $E \setminus E'$ (dashed lines)



the corresponding Kichhoffbasis

$$\mathcal{K} = (\text{cycle}(T, e_1), \text{cycle}(T, e_2), \text{cycle}(T, e_3), \text{cycle}(T, e_4), \text{cycle}(T, e_5))$$

is not a minimal cycle basis of $\mathcal{G}$.

Observation. Neither Kirchhoff nor ear decomposition produce *in general* a MCB.

Definition 3.27 (Chords & Chordfree)**.** Let $\mathcal{G} = (V, E)$ be a graph and C a cycle in $\mathcal{G}$. A chord in C in $\mathcal{G}$ is an edge $e = xy \in E \setminus C$ s.t. the endpoints x and y are incident to edges in C.

C is called chordfree (in $\mathcal{G}$) if no such chord exists in $\mathcal{G}$.

Proposition 3.28. Let $\mathcal{G}$ be a graph.
Necessary conditions for a cycle C in $\mathcal{G}$ to be *relevant* are:

(i) C is an elementary cycle, and

(ii) C is chordfree.

Proof. Let $\mathcal{G}$ be a graph and C a cycle in graph.
Condition (i) follows directly from Theorem 3.26.
We prove condition (ii) by proving its contra-positive:

$$C \text{ is not chordfree } \Rightarrow C \text{ is not relevant}$$

Let C be not chordfree, i.e. let $e = xy$ be a chord in C.
We can therefore define paths

$$P_1 := P_{xy} \text{ and } P_2 := P_{yx} \text{ s.t. } |P_1|, |P_2| \geq 2$$

which divide C into

$$C_1 = P_1 \cup \{e\} \text{ and } C_2 = P_2 \cup \{e\}$$

yielding

$$C = C_1 \oplus C_2 = P_1 \oplus P_2.$$

Due to

$$|C| = |P_1| + |P_2| \geq \underbrace{|P_1| + 1}_{=|C_1|}, \underbrace{|P_2| + 1}_{=|C_2|}$$

the cycles C_1 and C_2 are a linear combination of C from smaller cycles.
Therefore C is not relevant. □

Definition 3.29. Let $\mathcal{G} = (V, E)$ be a connected graph and $x, y \in V$ two arbitrary vertices.
The distance between x and y (in $\mathcal{G}$) is defined as

$$\operatorname{dist}_{\mathcal{G}}(x, y) := \min_{\text{path } P_{xy}} |P|$$

Length of the shortest path between x and y

If the respective graph $\mathcal{G}$ is clear from the context, the subscript may be dropped.

Definition 3.30 (Isometric)**.** Let $\mathcal{G} = (V, E)$ be a graph.
A subgraph $\mathcal{H} = (V', E')$ of $\mathcal{G}$ is isometric to $\mathcal{G}$ if

$$\forall\, x, y \in V' : \quad \operatorname{dist}_{\mathcal{H}}(x, y) = \operatorname{dist}_{\mathcal{G}}(x, y).$$

Proposition 3.31. Let $\mathcal{G} = (V, E)$ be a graph, C a cycle in $\mathcal{G}$ and $\mathcal{H} = (V', C)$ the subgraph of $\mathcal{G}$ induced by C.

$$C \text{ is relevant} \underset{\not\Leftarrow}{\Rightarrow} \mathcal{H} \text{ is isometric to } \mathcal{G}$$

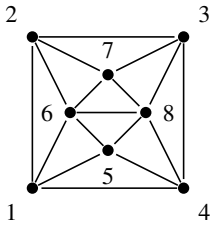
Proof. Let $\mathcal{G} = (V, E)$ be a graph, C a cycle in $\mathcal{G}$ and $\mathcal{H} = (V', C)$ the subgraph of $\mathcal{G}$ induced by C.

Proof of contra-positive:
Assume $\mathcal{H}$ is not isometric to $\mathcal{G}$
$\Rightarrow\ \exists\, x, y \in V' : \operatorname{dist}_{\mathcal{G}}(x, y) < |P_{xy}|, |P_{yx} = C \setminus P_{xy}|$
$\Rightarrow\ \exists$ path Q from x to y in $\mathcal{G}$ with $\operatorname{dist}_{\mathcal{G}}(x, y) = |Q| < |P_{xy}|, |P_{yx}|$
$\Rightarrow$ we can define cycles $C_1 := P_{xy} \oplus Q$ and $C_2 := P_{yx} \oplus Q$ with

$$C_1 \oplus C_2 = P_{xy} \oplus \underbrace{Q \oplus Q}_{\emptyset} \oplus P_{yx} = C$$

and by construction $|C_1|, |C_2| < |P_{xy}| + |P_{yx}| = |C|$.
$\Rightarrow$ The cycles C_1 and C_2 are a linear combination of C from smaller cycles.
Therefore C is not relevant. □

Example 3.8. As a counterexample to the second implication (from isometric to relevant) in Proposition 3.31, consider the example graph $\mathcal{G}$ given by:



The cycle

$$C = \{(1,2),(2,3),(3,4),(4,1)\}$$

is obviously not relevant as it can be constructed from the combination of all 10 elementary triangles.
However, C is isometric as for every pair of nodes x, y in C

(i) if $(x,y) \in C$, then the path $P_{xy} = \{(x,y)\}$ is already minimal with

$$|P_{xy}| = 1 = \mathrm{dist}(x,y)$$

(assuming default weight function), and

(ii) if $(x,y) \notin C$, i.e. for the pairs $(1,3)$ and $(2,4)$ the paths $P_{1,3} = \{(1,2),(2,3)\}$ and $P_{2,4} = \{(2,3),(3,4)\}$ respectively are minimal with $\mathrm{dist}(x,y) = 2$.

Definition (Edge-Short)**.** Let $\mathcal{G} = (V, E)$ be a graph.
A cycle C in $\mathcal{G}$ is edge short if it contains an edge $e = uv$ and a vertex $x \in V$ s.t.

$$C = \{e\} \uplus P_{ux} \uplus P_{vx}$$

where $|P_{ux}| = \text{dist}(x, u)$ and $|P_{vx}| = \text{dist}(x, v)$.

Without shift in definition (i.e. it provides an equivalent description/distiguisher for edge short cycles), the above statement can be reformulated to the following definition by introducing two cases.

Definition 3.32 (Edge-Short)**.** Let $\mathcal{G} = (V, E)$ be a graph.
A cycle C in $\mathcal{G}$ is edge short (e.s.) if $\forall\, e = uv \in C\ \exists$ vertex x in C s.t.

- if $|C|$ is odd, there exist paths P_{ux} and P_{vx} with

$$C = P_{ux} \uplus P_{vx} \uplus \{e\} \ \wedge\ \text{dist}(x, u) = |P_{ux}| = |P_{vx}| = \text{dist}(x, v),$$

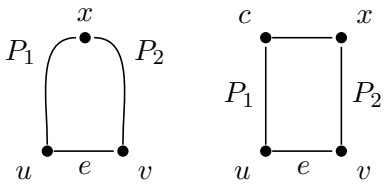
- if $|C|$ is even, there exist a vertex r with edge $rx \in C$ and paths P_{ur} and P_{vx} with

$$C = P_{rx} \uplus \{rx\} \uplus P_{vx} \uplus \{e\} \ \wedge\ \text{dist}(r, u) = |P_{ur}| = |P_{vx}| = \text{dist}(x, v).$$

This equivalent definition will later prevent multiple enumeration of identical cycles in the Horton-set; see Construction 3.34.

Observation. For a cycle C to be *edge short* it is necessary that for all edges $e = uv \in C$

- if C is of odd length, then there exists a vertex x in C s.t. $|P_1| = |P_2|$, and
- if C is of even length, then there exists an edge $e' = cx$ in path P_i s.t. $|P_1 \setminus \{cx\}| = |P_j|$ (for $i \neq j \in \{1, 2\}$).



odd even

Proposition 3.33 (HORTON, 1987)**.** Let $\mathcal{G}$ be a graph and C a cycle in $\mathcal{G}$.

$$C \text{ is relevant } \Rightarrow C \text{ is edge short}$$

Proof. (for odd-case, the even-case works analogue)
Let $\mathcal{G}$ be a graph and C a cycle in $\mathcal{G}$.
Proof of contra-positive:

Assume C is not e.s., hence there exists at least one edge $e = uv \in C$ with paths $P_{ux}, P_{vx} \subset C$ s.t.

(i) $\text{dist}(x, u) < |P_{ux}|$, or

(ii) $\text{dist}(x, v) < |P_{vx}|$.

W.l.o.g. we can assume that (i) is true. Therefore, there exists a path Q in $\mathcal{G}$ from u to x with $Q \cap C \neq \emptyset$ and

$$\text{dist}(x, u) = |Q| < |P_{ux}|$$

Q is the minimal path from u to x

resulting in a combination of $C = P_{ux} \oplus P_{vx} \oplus \{e\}$ from smaller cycles

$$\begin{aligned} C_u &= P_{ux} \oplus Q & |C| &> |C_u| \\ \text{and } C_v &= P_{vx} \oplus Q \oplus \{e\}. & |C| &\geq |C_v| \end{aligned}$$

$\Rightarrow$ C cannot be relevant. □

Construction 3.34 (Horton-Set of Cycles)**.** Let $\mathcal{G} = (V, E)$ be a graph.

Task: $\forall\, e = uv \in E, x \in V$: find edge short cycle of

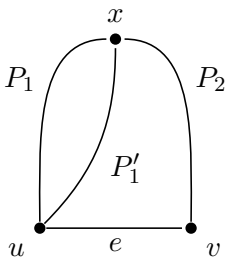
- odd length:

$$C = \{e\} \oplus P_{ux} \oplus P_{vx} \qquad \text{s.t. } |P_{ux}| = \text{dist}(x, u) = \text{dist}(x, v) = |P_{vx}|$$

- even length: find node $r \in V$ with $rx \in E$ and

$$C = \{e\} \oplus P_{ur} \oplus \{rx\} \oplus P_{vx} \quad \text{s.t. } |P_{ur}| = \text{dist}(r, u) = \text{dist}(x, v) = |P_{vx}|$$

This Construction 3.34 raises the question about *alternative shortest paths*:
On the example of an odd-length cycle C, let's assume that P_1' is an alternative shortest path to P_1, resulting in the following cycle constructions:



$$\begin{aligned} C^* &= P_1 \oplus P_1' & \qquad C' &= C \oplus C^* \\ C &= P_1 \oplus P_2 \oplus \{e\} & \qquad C &= C' \oplus C^* \\ C' &= P_1' \oplus P_2 \oplus \{e\} \end{aligned}$$

If P_1 and P_1' are of same length, then C and C' are interchangeable relevant but don't appear in the same MCB due to Greedy algorithm.
$\Rightarrow$ an arbitrary set of shortest paths between selected nodes is fine.

Another question is regarding the size of the Horton-set, and therefore the expected runtime:

- for odd length cycles the maximum number is bound by $O(|E| \cdot |V|)$,
- for even length cycles it is bound by $O(|E|^2)$, and
- for the greedy algorithm, especially the independence test we have essentially a *gaussian elimination* resulting in $O(|E|^2)$ per candidate

$\Rightarrow$ Overall runtime is in $O(|E|^4)$.

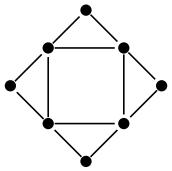
Definition 3.35 (ZyKov-Cycles)**.** Let $\mathcal{G} = (V, E)$ be a graph.
A ZyKov-Cycle is the *unique*, minimal cycle through an edge $e \in E$.

$$\forall\, e \in E:\ |\{C \mid e \in C\ \wedge\ |C| \text{ is minimal}\}| = 1$$

Corollary 3.36. Let $\mathcal{G}$ be a graph and C a cycle in $\mathcal{G}$.

$$C \text{ is a ZyKov-Cycle } \Rightarrow\ C \text{ is essential.}$$

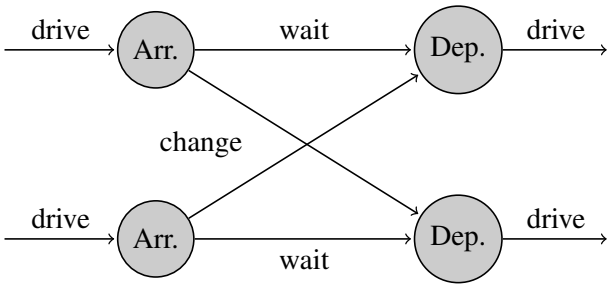
Example 3.9. The reverse direction of Corollary 3.36 is not true in general.
Consider again the example graph



where $\triangle$ is a ZyKov-Cycle while $\square$ is not. However, $\square$ is indeed an essential cycle.

Example (Periodic Event Scheduling Problem)**.** Use case for cycle bases.

Events $v_1, \ldots, v_n$ occur periodically with period T, however, are subject to constraints. The relations between events can be visualized as a graph $\mathcal{G} = (\mathcal{E}, \mathcal{A})$ with sets of events $\mathcal{E}$ and activities $\mathcal{A}$



with the corresponding time table TT and a total "headway" T.
Additionally, the constraints bound the travel between two events between a lower and upper limit

$$\forall\, e = (i,j) \in \mathcal{A}\ \exists\, l_{ij}, u_{ij} \in \mathbb{R}_{\geq 0} : \quad l_{ij} \leq TT_j - TT_i \leq u_{ij}.$$

Events may cross the cycle start, resulting in a periodic offset $P_{ij} \in \mathbb{Z}$ towards the headway time T:

$$l_{ij} \leq TT_j - TT_i + P_{ij} \cdot T \leq u_{ij}$$

The goal is to minimize the passenger travel time

$$\min \sum_{(i,j)\in\mathcal{A}} (TT_j - TT_i) \mod (T - l_{ij}).$$

One approach to this is with Mixed Integer linear programming (MIP):

We optimize stack variable ($\hat{=}$ unknowns) y_{ij} with

$$y_{ij} = TT_j - TT_i - l_{ij}$$

utilizing

- the spanning tree $\mathcal{T} = (\mathcal{E}, \mathcal{A}_\mathcal{T})$ and
- the fundamental cycle matrix Γ:

		edges			
		e_1	e_2	e_3	$\dots$
cycles	C_1	0	0	1	$\dots$
	C_2	1	1	0	$\dots$
	C_3	$\dots$			
	C_4	$\dots$			

$$\min \sum_{(i,j)\in\mathcal{A}} \omega_{ij} \cdot y_{ij}$$

s.t. $\Gamma(y_{ij} + \mathrm{I}) = T \cdot z_{ij}$ (periodic offset) with

$$\begin{aligned} 0 \le y_{ij} \le u_{ij} - l_{ij} &\qquad \forall\,(i,j) \in \mathcal{A} \\ y_{ij} \in \mathbb{R} &\qquad \forall\,(i,j) \in \mathcal{A} \\ z_{ij} \in \mathbb{Z} &\qquad \forall\,(i,j) \in \mathcal{A} \setminus \mathcal{A}_\mathcal{T}. \end{aligned}$$

4. Planar Graphs

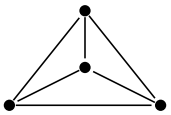
Definition (Planarity (informal))**.** A graph $\mathcal{G}$ is planar if there exists a 2D-drawing, s.t. no edges intersect, i.e. no two edges meet at a point outside of a vertex.

A drawing of a graph upholding the above condition is called plane.

Example 4.1. Considering the example graph K_4 (complete graph on 4 vertices), the drawing of the graph as



is not plane, while the drawing



is indeed plane. Hence, K_4 is planar.

Definition 4.1 (Graph Embedding)**.** Let $\mathcal{G} = (V, E)$ be a graph with $|V| = n$ and X be a set of 2D-vectors

$$X = (\vec{x_1}, \ldots, \vec{x_n}) \in \left(\mathbb{R}^2\right)^n$$

An embedding of $\mathcal{G}$ in $\mathbb{R}^2$ (in 2D)

$$\mathcal{G}[X]$$

assigns each vertex $v_i \in V$ a unique vector $x_i \in X$ for $1 \leq i \leq n$.

The embedded graph $\mathcal{G}[X]$ is also called framework.

Definition 4.2 (Planarity)**.** Let $\mathcal{G} = (V, E)$ be an undirected graph and $\mathcal{G}[X]$ a framework.
For each edge $e = \{x, y\}$ (with respective vectors $\vec{x}$ and $\vec{y}$) we define the curve function

$$\phi_{xy} : [0, 1] \to \mathbb{R}^2$$

such that

(i) $0 \mapsto \vec{x}$ and $1 \mapsto \vec{y}$,

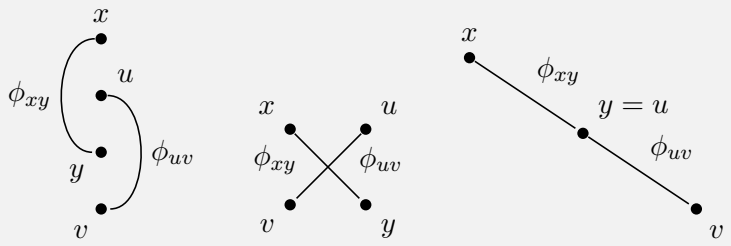
(ii) ϕ_{xy} is continuous and free of intersections, and

(iii) ϕ_{xy} is well-defined and injective

$$\forall\, t, t' \in [0, 1] : \quad \phi_{xy}(t) = \phi_{xy}(t') \;\Leftrightarrow\; t = t'.$$

Additionally, all pairs of curve functions must themselves be free of intersections, i.e. for all pairs of edges $\{x, y\}, \{u, v\} \in E$ (with respective vectors $\vec{x}$, $\vec{y}$, $\vec{u}$, and $\vec{v}$) holds

$$\phi_{xy}[0, 1] \cap \phi_{uv}[0, 1] = \{\vec{x}, \vec{y}\} \cap \vec{u}, \vec{v}.$$



ok not ok ok

If such a framework and curve functions exist, the graph is called planar.

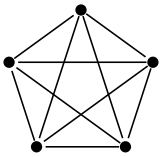
The logically next question would be: How do we know if a graph is planar?
A naive approach would list all possible drawings until a planar one is found, however, exhaustively creating all drawings of a graph is not feasible.

Observation 4.2.1. If a graph $\mathcal{G}$ is planar, then every subgraph of $\mathcal{G}$ is also planar.

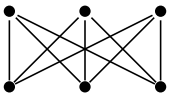
More importantly, the contra-positive states:

$$\exists \text{ non-planar subgraph } \mathcal{H} \text{ of } \mathcal{G} \;\Rightarrow\; \mathcal{G} \text{ is not planar.}$$

Example 4.2. Two famous non-planar graphs are the K_5 (complete graph on 5 vertices)



and the $K_{3,3}$ (complete bipartite graph on 6 vertices)



Definition 4.3 (Minor Operations)**.** Let $\mathcal{G} = (V, E)$ be a graph.

(i) removal of isolated vertex u ($\nexists\, v \in V : uv \in E$), resulting in the graph

$$\mathcal{G}' = (V \setminus \{u\}, E)$$

(ii) removal of an edge e (without removing the incident vertices), resulting in the graph

$$\mathcal{G}' = (V, E \setminus \{e\})$$

(iii) edge contraction, which identifies two vertices x, y (connected by an edge $e = xy$) to a new vertex $v_e \equiv (x = y)$



resulting in the graph $\mathcal{G}' = (V', E')$ with

$$\begin{aligned} V' :=& V \setminus \{x, y\} \cup \{v_e\} \\ E' :=& \{vw \in E \mid \{v, w\} \cap \{x, y\} = \emptyset\} \\ & \cup \{v_e w \mid wx \in E \setminus \{w\} \ \vee \ yw \in E \setminus \{e\}\} \end{aligned}$$

Definition 4.4 (Minor Graph)**.** Let $\mathcal{G}$ be a graph.
A graph $\mathcal{H}$ is called a minor of $\mathcal{G}$ if there exists a sequence of *minor operations* (i), (ii), and (iii), transforming $\mathcal{G}$ into $\mathcal{H}$.

Observation. Let $\mathcal{G}$ be a graph.
Every subgraph og $\mathcal{G}$ is also a minor of $\mathcal{G}$, but not every minor is also a subgraph.

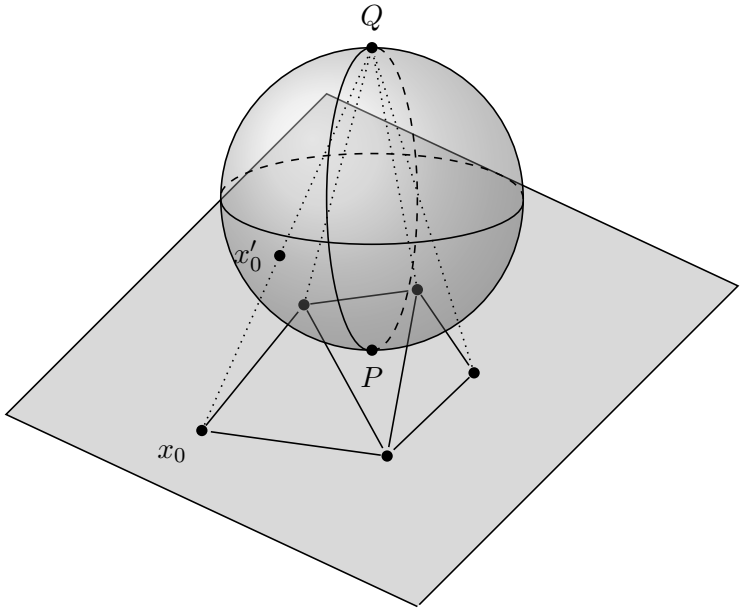
Theorem 4.5. Let $\mathcal{G}$ be a graph.

$$\mathcal{G} \text{ is planar } \Rightarrow \text{ every minor of } \mathcal{G} \text{ is planar.}$$

Hence, if $\mathcal{G}$ has a non-planar minor, then $\mathcal{G}$ is not planar.

Theorem 4.6 (KURATOWSKI, WAGNER)**.** A graph $\mathcal{G}$ is planar if and only if its minors include neither K_5 nor $K_{3,3}$ (see Example 4.2).

Observation 4.6.1. Let $\mathcal{G}$ be a graph.
If $\mathcal{G}$ is planar, then there exists an intersection free drawing in $\mathbb{R}^2$ and especially an intersection free drawing on the surface of a sphere.



The process starts by selecting a point $P \in \mathbb{R}^2$ s.t. P is not part of the graph drawing ("south pole" of the sphere) an a second point $Q \in \mathbb{R}^3$ above P along the normal vector of the plane ("north pole" of the sphere).
For each point x of the drawing, construct a straight line $\overline{xQ}$ and project the point x along

that line to the surface of the sphere:

$$\text{projection } x' = \overline{xQ} \cap \text{sphere} \setminus \{P\}$$

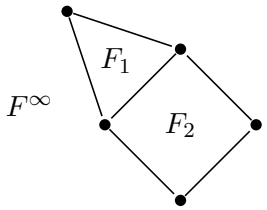
Definition 4.7 (Face)**.** Let $\mathcal{G}$ be a graph.
A face is an "enclosed area" in a drawing of $\mathcal{G}$.

They are further distinguished into

- *inner faces* $F_1, F_2, \ldots$ which are bound by edges on all sides, and
- the *outer face* F^∞ which 'contains' the graph and is bound by infinity.

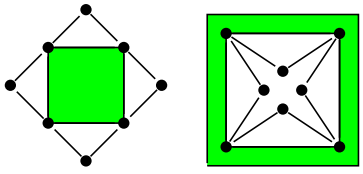
When denoting a *face of a planar graph* $\mathcal{G}$, we always imply that it is a face of a plane drawing of $\mathcal{G}$.

Example 4.3. The example graph $\mathcal{G}$ with the drawing



has F_1 and F_2 as inner faces and the (unbound) outer face F^∞.

Observation. Every face of a graph can be made the outer face by projecting the drawing onto the sphere, then rotating the sphere s.t. the point Q lies within the "new" outer face (the face spans over the "north pole" of the sphere), and then projecting back to 2D space.



For simplicity, the coloring of the actually infinite outer face on the rhs. is limited to a surrounding square.

Theorem 4.8 (Euler's Formula)**.** Let $\mathcal{G} = (V, E)$ be a connected plane graph (drawing) and $|F|$ be the number of faces.

$$|V| - |E| + |F| = 2$$

Proof. Let $\mathcal{G} = (V, E)$ be a connected graph with a plane drawing.
We prove Euler's Formula by induction over the number of edges $|E|$.

- *Induction Predicate* $P(n)$:
  If $|E| = n$, then $|V| - |E| + |F| = 2$.
- *Basic step*:
  For $|E| = 1$ we only have a single edge connecting two points. The only possible face thereby is the outer face F^{∞}:

  $$|V| - |E| + |F| = 2 - 1 - 1 = 2$$

  Therefore, $P(1)$ is true.
- *Induction requirement (IR)*:
  Let $P(k)$ be true for some $k \geq 1$.
- *Induction hypothesis (IH)*:
  $P(k) \Rightarrow P(k+1)$
- *Induction step (IS):*
  We know that for $|E| = k$ the formula holds. When increasing the number of edges on a connected graph, we only have two options:
  (i) We add a new vertex and connect it to an existing one. The new vertex therefore has to lie within an existing face and, therefore, adding the edge won't change the number of faces:

  $$(|V| + 1) - (|E| + 1) + |F| = 2$$

  (ii) We add an edge between two existing vertices, thus seperating the incident face into two parts (along the new edge):

  $$|V| - (|E| + 1) + (|F| + 1) = 2$$

  $\Rightarrow P(k+1)$ is true.

□

4.1. Cycle Bases in Planar Graph

Observation. Regarding a plane drawing of a graph $\mathcal{G} = (V, E)$, every inner (or bounded) face is delimited by an elementary cycle in $\mathcal{G}$.
Additionally, from Euler's formula we know

$$\begin{aligned} |F| &= |E| - |V| + 2 && | - 1 \\ |F| - 1 &= \underbrace{|E| - |V| - 1}_{\mu(\mathcal{G})} \end{aligned}$$

therefore, the cyclomatic number corresponds to the number of bound faces.

In fact, the cycles bounding each inner face can be ordered to form an ear cycle basis of $\mathcal{G}$.

Definition 4.9. Let $\mathcal{G} = (V, E)$ be a planar graph.
For any inner face F_i in $\mathcal{G}$ with $1 \leq i \leq \mu(\mathcal{G})$, in formulas we may use F_i as a denotation for its bounding cycle C_i in $\mathcal{G}$.
Regarding the outer (or exterior) face F^∞ we additionally denote

$$\operatorname{ext}(\mathcal{G}) := F^\infty \mathrel{\hat{=}} \bigoplus_{i=1}^{\mu(\mathcal{G})} F_i.$$

Observation. Let $\mathcal{G}$ be a planar graph.
Every edge is contained in *at most* two bounded faces (i.e. in the cycles bounding an inner face) and *exactly* two faces (including the exterior face):

$$\bigoplus_{i=1}^{\mu(\mathcal{G})} F_i \oplus \operatorname{ext}(\mathcal{G}) = \emptyset.$$

Definition 4.10 (Two-Bases)**.** Let $\mathcal{G} = (V, E)$ be a planar graph.
A cycle basis $\mathcal{B}$ of $\mathcal{G}$ is called a two-basis of $\mathcal{G}$ if every edge is contained in at most two cycles of $\mathcal{B}$:

$$\forall\, e \in E :\ |\{C \in \mathcal{B} \mid e \in C\}| \leq 2.$$

Observation. For a planar graph $\mathcal{G}$, the cycles of the bounded faces of $\mathcal{G}$ form a two-basis of $\mathcal{G}$.

Theorem 4.11 (MAC LANE, 1937)**.** A graph $\mathcal{G}$ is planar if and only if it has a two-basis.

Before proving above theorem, we first need to provide the following lemma.

Lemma 4.12. Let $\mathcal{G}$ be a graph and $\mathcal{H}$ a minor of $\mathcal{G}$.

$$\mathcal{G} \text{ has a two-basis } \Rightarrow \mathcal{H} \text{ has a two-basis.}$$

forming minors preserves the property of having a two-basis

Proof. Let $\mathcal{G} = (V, E)$ be a graph with a two-basis $\mathcal{B}$.
We need to prove that the minor operations (Definition 4.3) do not affect the property of having a two-basis.

The operations (i) node deletion and (iii) edge contraction obviously preserve the preserve the property of minors having a two-basis as well.
Therefore, we only need to look at edge deletion:

Let $e \in E$ and $\mathcal{G}' := (V, E \setminus \{e\})$. By construction we know that $\mu(\mathcal{G}') = \mu(\mathcal{G}) - 1$.
We differentiate the following cases:

1) If e is contained in exactly one cycle $C \in \mathcal{B}$, then

$$\mathcal{B}' := \mathcal{B} \setminus \{C\} \subsetneq \mathcal{B}$$

is a linear independent set of cycles with length $|\mathcal{B}'| = \mu(\mathcal{G}')$.
$\Rightarrow$ $\mathcal{B}'$ is a cycle basis of $\mathcal{G}'$.
By construction, each edge in $\mathcal{G}'$ is still contained in at most two cycles of $\mathcal{B}'$.
$\Rightarrow$ $\mathcal{B}'$ is a two-basis.

2) If e is contained in exactly two cycles $C_1, C_2 \in \mathcal{B}$, then

$$\mathcal{B}' := \mathcal{B} \setminus \{C_1, C_2\} \cup \{\mathrm{ext}(\mathcal{G})\}$$

is a set of cycles with $|\mathcal{B}'| = \mu(\mathcal{G}')$ and as $e \in C_1$ and $e \in C_2$ we know that, due to the sum ($\oplus$), $e \notin \mathrm{ext}(\mathcal{G})$.
As every edge, that appears in two cycles in $\mathcal{B}$, does not appear in $\mathrm{ext}(\mathcal{G})$, due to the sum ($\oplus$), and every edge, that appears in only one cycle in $\mathcal{B}$, appears in $\mathrm{ext}(\mathcal{G})$, we know that $\mathcal{B}'$ is a two-basis of $\mathcal{G}'$.

Geometrical Visualization:
The joined face of C_1 and C_2, that we get after deleting edge e, can be made the outer face (Observation 4.6.1).
Therefore the previous outer face becomes a bounded face and can thus be used in the basis construction.

$\Rightarrow$ Every minor of $\mathcal{G}$ also has a two-basis. □

Now we can prove Theorem 4.11.

Proof. Let $\mathcal{G}$ be a graph.

('$\Rightarrow$') Assume $\mathcal{G}$ is planar.
$\Rightarrow$ $\exists$ two-basis of $\mathcal{G}$, namely the cycles bounding the inner faces of $\mathcal{G}$.

('$\Leftarrow$') Assume $\mathcal{G}$ has a two-basis $\mathcal{B}$.
Lemma 4.12 $\Rightarrow$ every minor of $\mathcal{G}$ has a two-basis.
Proof by contradiction:

Assume: $\mathcal{G}$ is not planar.

Theorem 4.6 $\Rightarrow$ K_5 or $K_{3,3}$ has a two-basis.

- The $K_5 = (V, E)$ has $|V| = 5$ and $|E| = 10$ with its cyclomatic number

$$\mu(K_5) = |E| - |V| + 1 = 10 - 5 + 1 = 6$$

$\Rightarrow$ 7 faces (including $\mathrm{ext}(K_5)$) with every face being definied by at least 3 edges.
However, $2 \cdot |E|$ is the number of edges that must appear in all faces together (including $\mathrm{ext}(K_5)$).

$$2 \cdot |E| = 20 \geq 21 = 7 \cdot 3 \qquad \lightning$$

- The $K_{3,3} = (V, E)$ has $|V| = 6$ and $|E| = 9$ with its cyclomatic number

$$\mu(K_{3,3}) = |E| - |V| + 1 = 9 - 6 + 1 = 4$$

$\Rightarrow$ 5 faces (including $\mathrm{ext}(K_{3,3})$) with every face being definied by at least 4 edges (because $K_{3,3}$ is bipartite).
However, $2 \cdot |E|$ is the number of edges that must appear in all faces together (including $\mathrm{ext}(K_{3,3})$).

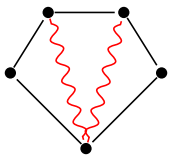
$$2 \cdot |E| = 18 \geq 20 = 5 \cdot 4 \qquad \lightning$$

$\Rightarrow$ $\mathcal{G}$ is planar.

□

Definition 4.13 (Maximal Planarity)**.** Let $\mathcal{G}$ be a planar graph.
$\mathcal{G}$ is called maximally planar if every inner face is bound by cycle of length 3 (triangular faces).

Example 4.4. The following pentagonal graph $\mathcal{G}$, can be *completed* into a maximally planar graph by adding the red edges:



Observation 4.13.1. Regarding a planar graph $\mathcal{G} = (V, E)$, Euler's formula can be rewritten as

$$|E| - |F| = |V| - 2.$$

Additionally, every edge is adjacent to exactly two faces and in turn every face (its descriptive cycle) is made up of at least 3 edges, resulting in

$$|E| \leq 3(|V| - 2).$$

This formula is an equality, if every face is a triangle, i.e. if $\mathcal{G}$ is maximally planar

$$|E| = 3(|V| - 2).$$

We know that the average vertex degree in $\mathcal{G}$ is

$$\deg(\mathcal{G}) = \frac{1}{|V|} \sum_{v \in V} \deg(x) = \frac{2 \cdot |E|}{|V|} \leq \frac{2 \cdot 3 \cdot (|V| - 2)}{|V|} = 6 - \frac{12}{|V|} < 6$$

$\Rightarrow$ There is at least one vertex x in a planar graph s.t. $\deg(x) \leq 5$.

Definition 4.14 (Deficiency)**.** Let $\mathcal{G} = (V, E)$ be a graph and $v \in V$.
The deficiency of vertex v is

$$\mathrm{def}(v) := 6 - \deg(v).$$

Corollary 4.15. In a maximally planar graph, the sum of the deficiencies is 12.

Proof. Let $\mathcal{G} = (V, E)$ be a maximally planar graph.
We know the equality

$$|E| = 3(|V| - 2)$$

holds. Thus we get

$$\sum_{v \in V} \mathrm{def}(v) = \sum_{v \in V} (6 - \deg(v)) = 6|V| - 6(|V| - 2) = 12.$$

□

As a refinement of our definition for graph drawings, where edge are drawn as curve functions, we introduce the following theorem:

Theorem 4.16 (FÁRY's Theorem)**.** Every planar graph has an embedding in the 2D-plane s.t. all edges are drawn as straight lines.

Proof. Let $\mathcal{G} = (V, E)$ be a planar graph.
W.l.o.g. we can assume all graphs in this proof (if not stated otherwise) to be maximally planar, as for every not-maximally planar subgraph, the deletion of an edge from the plane drawing (of the maximally planar graph) does not change the straight-line property of the other edges.
If $|V| = 1$, the embedding is trivial and if $|V| = 2$ there can exist at most one edge in $\mathcal{G}$ which can be trivially drawn as a straight line.
For graphs with more than two vertices we prove by induction over $|V|$.

- *Induction Predicate* $P(n)$:
  Every planar graph $\mathcal{G} = (V, E)$ of $|V| = n \geq 3$ vertices can be re-embedded in the plane s.t. the corners $\{a, b, c\} \subseteq V$ of a triangular face delimit the external face and all edges are straight lines.

- *Basic step*:
  For a graph $\mathcal{G}_3 = (V, E)$ with $|V| = 3$, the graph can be re-embedded into the plane with identical coordinates and straight lines as edges:

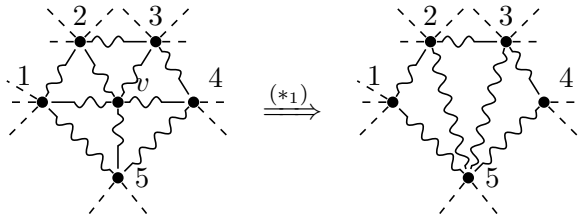
  $\Rightarrow$ $P(3)$ holds.

- *Induction requirement (IR)*:
  Let $P(k)$ be true for some $k \geq 3$.

- *Induction hypothesis (IH)*:
  $P(k) \Rightarrow P(k+1)$
- *Induction step (IS):*
  Let $\mathcal{G}_{k+1} = (V, E)$ be a planar graph with $|V| = k+1$ and $\{a, b, c\} \subsetneq V$ a set of distinct vertices.

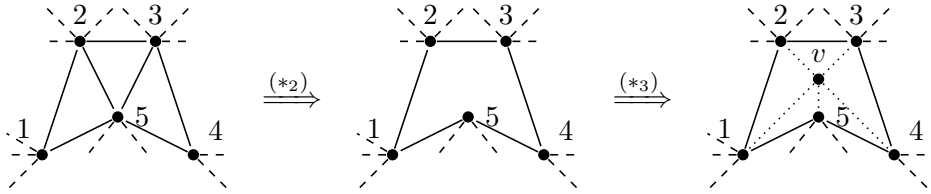
> We claim that there exists a vertex $v \in V$ with at most 5 neighbors and $v \notin \{a, b, c\}$.
> From Observation 4.13.1 we know that there exists at least one vertex with at most 5 neighbors.
> As $\mathcal{G}_{k+1}$ is maximally planar, we also know
>
> $$\forall\, x \in V:\ \deg(x) \geq 3 \text{ and hence } \operatorname{def}(x) = 6 - \deg(x) \leq 3$$
>
> As the sum of deficiencies in $\mathcal{G}_{k+1}$ is 12 (cf. Corollary 4.15), there are at least 4 vertices with a deficiency larger than zero and, therefore, a degree less than or equal to 5.
> Thus, there must exists a vertex $v \notin \{a, b, c\}$ with $\deg(v) \leq 5$, which proves our claim.

We then find $v \notin \{a, b, c\}$ with $3 \leq \deg(v) \leq 5$. Under the exhibition theorem (see later), we can assume the most difficult case that $\deg(v) = 5$.
Consider the embedding of $\mathcal{G}_{k+1}$



($*_1$) remove v and make maximally planar

resulting in the RHS embedded graph $\mathcal{G}'$ with k vertices.
(IR) $\Rightarrow$ $\mathcal{G}'$ can be re-embedded with only straight edges.



($*_2$) remove the previously added edges
($*_3$) reinsert v

Under the exhibition theorem, $\mathcal{G}_{k+1}$ can therefore be drawn as a planar graph with only straight lines.
$\Rightarrow P(k+1)$ is true.

□

Above prove leverages a result from geometry:

Theorem 4.17 (Specification of the exhibition theorem)**.** In an arbitrary, regular pentagon there is always a point in the interior from which all corners can be reached with a straight line without crossing any edges of the pentagon.

5. Graph Coloring

Literature lists a lot of graph coloring definitions and problems thereby, however, we will only look at one type of graph coloring problems.

Definition 5.1 (Proper Vertex Coloring)**.** Let $\mathcal{G} = (V, E)$ be an undirected graph and $\{1, \ldots, c\}$ a set of c *colors* (usually actual colors or $\subset \mathbb{N}^*$).
A graph coloring is a function

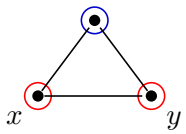
$$\sigma : V \to \{1, \ldots, c\}.$$

The graph coloring is called proper if

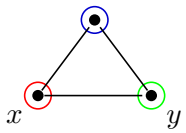
$$\forall \{x, y\} \in E : \ \sigma(x) \neq \sigma(y).$$

all vertices connected by edges have different colors

Example 5.1. For the example graph $\mathcal{G} = (V, E)$ with the (not-proper) coloring σ



because $\sigma(x) = \sigma(y)$ for $\{x, y\} \in E$, and a proper coloring σ'



Definition 5.2 (Graph Coloring Problem)**.** Let $\mathcal{G} = (V, E)$ be an undirected graph. The instance of the graph coloring problem on this graph describes finding a *proper vertex coloring*

$$\sigma : V \to \{1, \ldots, c\}$$

of $\mathcal{G}$, that uses the minimal number of necessary colors, i.e.

$$c = 1 \ \vee \ \forall \text{ proper } \sigma' : V \to \{1, \ldots, c'\} \text{ holds } c \leq c'.$$

This minimum number of colors for $\mathcal{G}$ is called the

$$\text{CHROMATIC number } \chi(\mathcal{G}).$$

Remark. *The decision problem for Graph Coloring is* NP*-complete.*

Observation. Regarding the chromatic number of a graph $\mathcal{G} = (V, E)$, we know certains bounds based on the type of $\mathcal{G}$:

- bipartite graph:
  - general: $\chi(\mathcal{G}) \leq 2$,
  - if $|E| \geq 1$: $\chi(\mathcal{G}) = 2$
- graph consisting of a single elementary cycle:
  - if $|V|$ even: $\chi(\mathcal{G}) = 2$,
  - if $|V|$ odd: $\chi(\mathcal{G}) = 3$
- complete graph:
  - $\chi(K_n) = n$

Observation. Let $\mathcal{G} = (V, E)$ be an undirected graph.
Reminding on the average degree in a graph $\mathcal{G}$ (cf. Definition 1.6)

$$\Delta(\mathcal{G}) := \max_{x \in V} \deg(x)$$

then we get an upper bound for the chromatic number

$$\chi(\mathcal{G}) \leq \Delta(\mathcal{G}) + 1$$

which is an equality, i.e. $\chi(\mathcal{G}) = \Delta(\mathcal{G}) + 1$, if and only if $\mathcal{G}$ is complete or a vertex disjoint union of elementary cycles containing at least one cycle of odd length.
Hence we cannot improve on this upper bound any further, at least for general graphs.

One slight improvement on the bound is through the following theorem, which we for now state without proof:

Theorem 5.3 (BROOK's theorem)**.** Let $\mathcal{G}$ be a graph with $\Delta(\mathcal{G}) \geq 3$.

$$\forall\, n \in \mathbb{N} :\ \mathcal{G} \neq K_n \quad \Rightarrow \quad \chi(\mathcal{G}) \leq \Delta(\mathcal{G})$$

if $\mathcal{G}$ is not complete, the upper bound is $\Delta(\mathcal{G})$

5.1. Coloring in Planar Graphs

Theorem 5.4 (6-Colorability)**.** Every planar graph can colored with at most 6 colors.

Proof. We prove by induction over the number of vertices.

- *Induction Predicate* $P(n)$:
  6 colors suffice for all planar graphs $\mathcal{G} = (V, E)$ with $|V| = n$ vertices.
- *Basic step*:
  For any graph $\mathcal{G} = (V, E)$ with $|V| \leq 6$ the proof is trivial.
- *Induction requirement (IR)*:
  Let $P(k)$ be true for some $k \geq 6$.
- *Induction hypothesis (IH)*:
  $P(k) \Rightarrow P(k + 1)$

- *Induction step (IS):*
  Let $\mathcal{G}_{k+1} = (V, E)$ be a planar graph with $|V| = k + 1$.
  We already know there exists at least one vertex $v \in V$ with $\deg(v) \leq 5$.
  Defining $\mathcal{G}' := \mathcal{G}|_{V \setminus \{v\}}$ we know from (IR) that $\mathcal{G}'$ is 6-colorable.
  Thus, let σ be a 6-coloring of $\mathcal{G}'$.
  As v has ≤ 5 neighbors with ≤ 5 distinct colors given by σ, so use the sixth color for v, which yields a 6-coloring $\tilde{\sigma}$ for $\mathcal{G}$ extending σ.
  $\Rightarrow P(k+1)$ is true.

□

The 6-colorability for planar graphs can be further improved:

Theorem 5.5 (5-Colorability)**.** Every planar graph is 5-colorable.

Proof. We again prove by induction over the number of vertices.

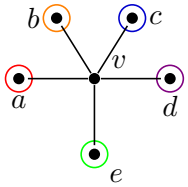
- *Induction Predicate* $P(n)$:
  5 colors suffice for all planar graphs $\mathcal{G} = (V, E)$ with $|V| = n$ vertices.

- *Basic step*:
  For any graph $\mathcal{G} = (V, E)$ with $|V| \leq 5$ the proof is trivial.

- *Induction requirement (IR)*:
  Let $P(k)$ be true for some $k \geq 5$.

- *Induction hypothesis (IH)*:
  $P(k) \Rightarrow P(k+1)$

- *Induction step (IS):*
  Let $\mathcal{G}_{k+1} = (V, E)$ be a planar graph with $|V| = k + 1$ and let $v \in V$ be a vertex with $\deg(v) \leq 5$.
  If v has less than 5 neighbors, the coloring is trivial/ analogous to the previous proof. Therefore, we only need to consider the case of $\deg(v) = 5$.
  Defining $\mathcal{G}' := \mathcal{G}|_{V \setminus \{v\}}$ we know from (IR) that $\mathcal{G}'$ is 5-colorable.
  Thus, let σ be a 5-coloring of $\mathcal{G}'$.
  W.l.o.g. we can assume that all neighbors of v have distinct colors, as the coloring again would be trivial otherwise. Hence, the neighborhood of v can be seen as follows



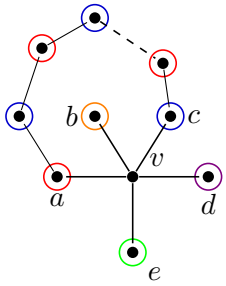
as a visual aid the colors $\in \mathbb{N}$ are drawn As actual colors: 1, 2, 3, 4, 5.

Consider the subgraph of $\mathcal{G}'$ induced by the vertices with color $\sigma(a) = 1$ or $\sigma(c) = 3$, i.e.

denote $\mathcal{G}'[1,3] := \mathcal{G}'|_{V'}$ with $V' := \{x \in \mathcal{G}' \mid \sigma(x) = \sigma(a) \ \vee \ \sigma(x) = \sigma(c)\}$.

Case 1) Vertices a and c are in different connected components of $\mathcal{G}'[1,3]$.
Then, swap colors assigned 1 and 3 in the component containing c and color $\sigma(v) = 3$.

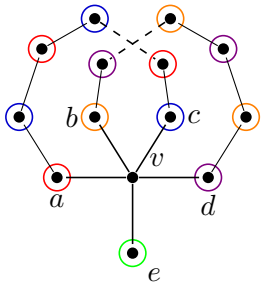
Case 2) Vertices a and c are in the same connected component of $\mathcal{G}'[1,3]$.
Therefore, there exists a 1,3-colored path connecting a and c:



This path plus $c - v - a$ closes a cycle that contains b but not d in the planar graph embedding.
We now consider the subgraph of $\mathcal{G}'$ induced by the vertices with color $\sigma(b) = 2$ or $\sigma(d) = 4$, i.e. $\mathcal{G}'[2,4]$.

Case 2a) Vertices b and d are in different connected components of $\mathcal{G}'[2,4]$.
Then, swap colors assigned 2 and 4 in the component containing d and color $\sigma(v) = 4$.

Case 2b) Vertices a and c are in the same connected component of $\mathcal{G}'[2,4]$.
Therefore, there exists a 2,4-colored path connecting b and d:



However, this path crosses the 1,3-path from a to c.
Since $\mathcal{G}$ is planar embedded, edges do not cross outside of a vertex, hence both paths must have a vertex in common. ↯
$\Rightarrow$ Case 2b) does never happen for planar embeddings.

$\Rightarrow$ $\mathcal{G}$ has a proper 5-coloring.
$\Rightarrow$ $P(k+1)$ is true.

□

There also exist a further improvement on the 5-colorability for planar graphs:

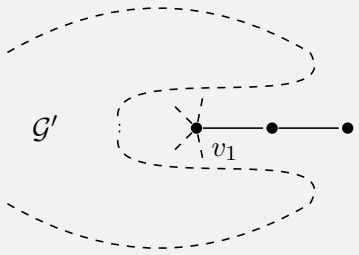
Theorem 5.6 (4-Colorability)**.** Every planar graph is 4-colorable.

Remark. *The "proof" of above theorem was done through computer calculation for testing the finite number of minimal possible counter examples and, to date, no proof via purely pen-&-paper exists.*

As preparation for proving Brook's theorem we introduce the following construction/ procedure PATHCOLOR:

Construction 5.7. Let $\mathcal{G} = (V, E)$ be a graph.
Suppose we have a partial (proper) coloring of $\mathcal{G}$ and let $v_1 v_2 \dots v_r$ be a path in $\mathcal{G}$ that has remained uncolored.

Now consider $\mathcal{G}$ with maximum degree $\Delta(\mathcal{G}) \leq k$ and suppose we have a partial k-coloring and $P = v_1 v_2 \dots v_r$ is a maximum, non-colored path in $\mathcal{G}$.
Then the starting node of the path, v_1, has at most $k-1$ neighbors in $\mathcal{G}' := \mathcal{G} - P$, which leaves at least one color left to color v_1:



This process can then be extended along P up until the very last vertex.
$\Rightarrow$ PATHCOLOR extends the given coloring to all vertices of the path, except the last vertex (v_r).

Proof of Brook's Theorem. Let $\mathcal{G} = (V, E)$ be a graph with $\Delta(\mathcal{G}) \geq 3$ and $\forall\, n \in \mathbb{N} : \mathcal{G} \neq K_n$. If $\mathcal{G}$ is made up of a single elementary cycle, we know already that Brook's theorem holds. Therefore we can assume $\mathcal{G}$ to not be a single elementary cycle and especially not a cycle of odd length.

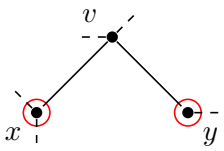
It is possible to require graph connectedness for each following step, as it would otherwise be a disjoint union of smaller subgraphs, however, the same can also be concluded later on (see Case 2.a).

Denote $k := \Delta(\mathcal{G})$. We prove by induction over $|V|$:

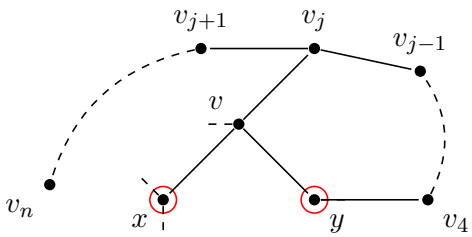
- *Induction Predicate* $P(n)$:
  For any graph $\mathcal{G} = (V, E)$ with $|V| = n$, $k = \Delta(\mathcal{G})$ and above restrictions, there exists a proper k-coloring for $\mathcal{G}$.
- *Basic step*:
  For any graph $\mathcal{G} = (V, E)$ with $|V| \leq k$ the proof is trivial.
- *Induction requirement (IR)*:
  Let $P(n')$ be true for some $n' \geq k$.
- *Induction hypothesis (IH)*:
  $P(n') \Rightarrow P(n' + 1)$
- *Induction step (IS):*
  Let $\mathcal{G} = (V, E)$ be a graph with mentioned restrictions and $|V| = n' + 1$.
  - Assume $\mathcal{G}$ is not regular, i.e. $\exists\, v \in V$ with $\deg(v) < \Delta(\mathcal{G}) = k$.
    By (IR) there exists a k-coloring of $\mathcal{G}' := \mathcal{G}|_{V \setminus \{v\}}$.
    $\Rightarrow$ As v has at most $k - 1$ neighbors, there is at least one color left to color v.
  - Now we assume that $\mathcal{G}$ is regular, i.e. $\forall\, v \in V :\ \deg(v) = \Delta(\mathcal{G}) = k$.
    Let $v \in V$ be an arbitrary vertex in $\mathcal{G}$.
    Then v must have two neighbors x and y that are not adjacent, as by assumption, $\mathcal{G}$ does not contain a clique on $k + 1$ (or more) vertices.
    Let $P = xvyv_4v_5 \dots v_r$ be a path that is maximal in $\mathcal{G}$ (i.e. there exists no longer path starting with x-v-y).

    We can distinguish two cases:

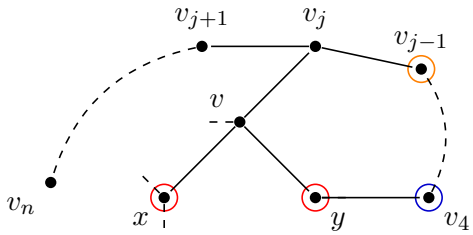
  Case 1) P covers all vertices in $\mathcal{G}$. $\Rightarrow\ v_r = v_n$ with $|V| = n$
  We start by picking the same color for $x = v_1$ and $y = v_3$.



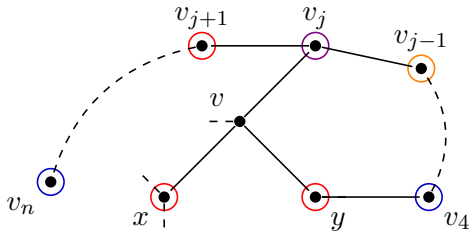
  We determine v_j with $3 < j \leq n$ s.t. v_j is a neighbor of v. We know that such a v_j must exists, because $\deg(v) = \Delta(\mathcal{G}) \geq 3$ by assumption.



Next, we call PATHCOLOR on $P_1 = v_4 \dots v_j$ which colors the path up to and including v_{j-1}:

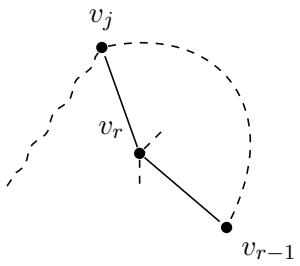


Afterwards, we call PATHCOLOR on $P_2 = v_n v_{n-1} \dots v_{j+1} v_j v$ so only v remains uncolored:



As v has $\Delta(\mathcal{G}) = k$ neighbors but x and y have the same color and therefore there remains at least one color to color v.

Case 2) P does not cover all vertices in $\mathcal{G}$, i.e. P "gets stuck" in v_r with $r < |V|$.
$\Rightarrow$ all neighbors of v_r are in P.
Let v_j be the vertex with the smallest $j < r$ that is a neighbor of v_r.



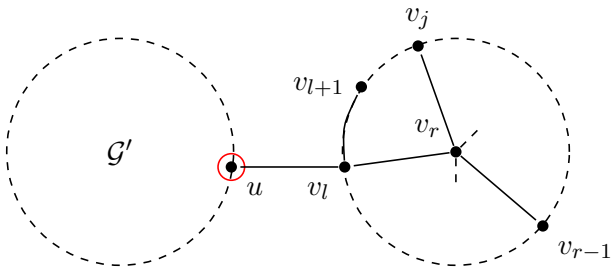
Then the subpath $P' = v_j \dots v_{r-1} v_r$ of P with the additional edge $\{v_r, v_j\}$ forms a cycle

$$C = P' \cup \{\{v_r, v_j\}\}.$$

Consider $\mathcal{G}'$ as $\mathcal{G}$ without C.
By (IR) there exists a k-coloring of $\mathcal{G}'$, which is a partial coloring of $\mathcal{G}$.
We can (possibly) distinguish two cases:

Case 2.a) C and $\mathcal{G}'$ are not connected by an edge, hence $P = P'$.
Then $\mathcal{G}$ is the disjoint union of C and $\mathcal{G}'$ and therefore both get separate k-colorings, which exist by (IR).

Case 2.b) C and $\mathcal{G}'$ are connected by an edge.
Then, let v_l in C be the vertex with the highest index, s.t. v_l is connected to $\mathcal{G}$. We know $v_r \neq v_l$.



By construction, v_{l+1} has no neighbors in $\mathcal{G}'$.
We assign v_{l+1} to have the same color as u and then call PATHCOLOR on

$$v_{l+2} \dots v_r v_j \dots v_{l-1} v_l$$

which colors the entire cycle except v_l.
As v_l has $\Delta(\mathcal{G}) = k$ neighbors but u and v_{l+1} have the same color and therefore there remains at least one color to color v_l.

□

6. CoGraphs

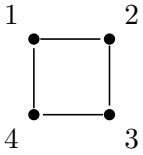
Definition 6.1 (Complement)**.** Let $\mathcal{G} = (V, E)$ be an undirected graph.
Then $\overline{\mathcal{G}} = (V, E^C)$ with

$$E^C := (V \times V) \setminus E$$

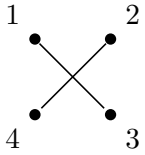
is the complement of $\mathcal{G}$.

Example 6.1. For some connected graphs, their complement graph consists of multiple connected components.

$\mathcal{G}$:

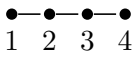


$\overline{\mathcal{G}}$:

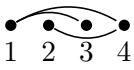


There also exist graphs that are isomorphic to its complement:

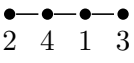
$\mathcal{G}$:



$\overline{\mathcal{G}}$:



$\simeq$



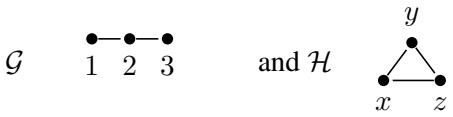
Definition 6.2. CoGraphs (abrv. complement-reducible graphs) are defined inductively as follows:

- K_1 is a cograph.
- The disjoint union of two (or more) cographs is a cograph.
- The complement of a cograph is a cograph.

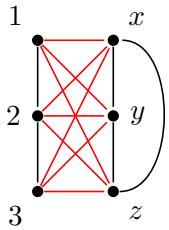
Definition 6.3 (Graph Join)**.** Let $\mathcal{G} = (V, E)$ and $\mathcal{H} = (V', E')$ be two (disjoint) graphs. The join $\mathcal{G} \triangledown \mathcal{H} := (V^*, E^*)$ has

- the vertex set $V^* = V \cup V'$ and
- the edge set $E^* = E \cup E' \cup (V \times V')$.

Example 6.2. Given the graphs



their join is (new added edges in red):



Observation 6.3.1. The correspondence between the *disjoint union* and the *join* of two graphs $\mathcal{G}$ and $\mathcal{H}$ is

$$\mathcal{G} \cup \mathcal{H} = \overline{\overline{\mathcal{G}} \triangledown \overline{\mathcal{H}}} \quad \text{and} \quad \mathcal{G} \triangledown \mathcal{H} = \overline{\overline{\mathcal{G}} \cup \overline{\mathcal{H}}}.$$

This also concludes that the join of two cographs is a cograph.

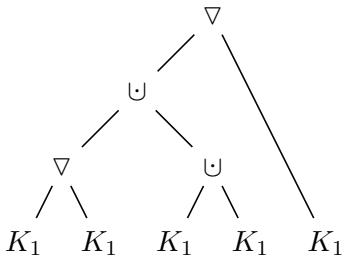
Algorithm 6.1. CoGraphs cen be recognized via the following recursive algorithm:

- Input: a graph $\mathcal{G}$
- If $\mathcal{G}$ is not connected:
  - Run algorithm on each connected component
- Else, if $\mathcal{G}$ is connected:
  - If $\mathcal{G} = K_1$: `return true`

- Else, construct $\overline{\mathcal{G}}$
- If $\overline{\mathcal{G}}$ is also connected: `return false`
- Else:
  - Run algorithm on $\overline{\mathcal{G}}$

Remark. *Both the Definition 6.2 and Algorithm 6.1 associate a co-tree to the cograph.*

Example 6.3. Given the following co-tree



starting with five K_1 graphs, denoted through vertices $a - e$, we get a step-wise construction of the cograph by

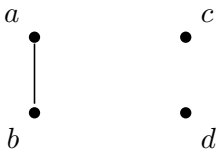
$K_1 \triangledown K_1$:



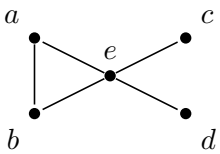
$K_1 \cup K_1$:



then, the combination to $(K_1 \triangledown K_1) \cup (K_1 \cup K_1)$



and the final cograph $((K_1 \triangledown K_1) \cup (K_1 \cup K_1)) \triangledown K_1$:



Observation 6.3.2. The coloring of cographs again is given inductively.

- K_1 is trivially 1-colorable
- disjoint union ($\cup$):
  determine the connected component with the maximum number of colors, then re-color all other connected components with this set of colors
- joins ($\triangledown$):
  for a join of two cographs $\mathcal{G} = (V, E)$ and $\mathcal{H} = (V', E')$, their vertex sets V and V' must have disjoint color sets as all nodes from $\mathcal{G}$ are connected by an edge to all nodes of $\mathcal{H}$.

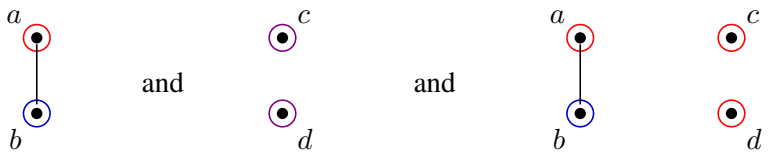
Therefore, the chromatic numbers are:

- $\chi(K_1) = 1$
- $\chi(\mathcal{G} \cup \mathcal{H}) = \max\{\chi(\mathcal{G}), \chi(\mathcal{H})\}$
- $\chi(\mathcal{G} \triangledown \mathcal{H}) = \chi(\mathcal{G}) + \chi(\mathcal{H})$

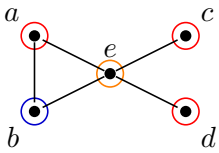
Example 6.4. Considering the same co-tree from Example 6.3, we start with 5 K_1 graphs, which all have their own 1-coloring:



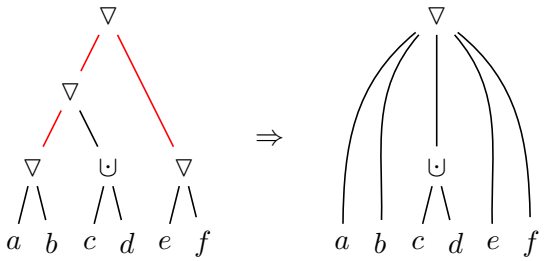
The coloring for all the steps are then constructed as by above observation:



resulting in the final, colored cograph:



Observation 6.3.3. Due to the commutativity of both the join and the disjoint union, co-trees can be converted into so called discriminating co-trees by contracting edges between vertices with the same operation:



7. Graph Isomorphisms

Definition 7.1 (Graph isomorphism)**.** Let $\mathcal{G} = (V, E)$ and $\mathcal{H} = (V', E')$ be graphs. A mapping (or function)

$$\varphi : V \to V'$$

such that

(1) $\forall\, x, y \in V: \quad \{x, y\} \in E \;\Leftrightarrow\; \{\varphi(x), \varphi(y)\} \in E'$ and

(2) φ is bijective

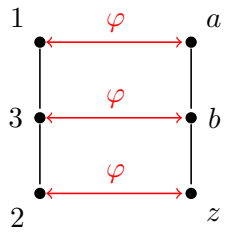
then φ is called a graph isomorphism.

Two graphs $\mathcal{G}$ and $\mathcal{H}$ are called isomorphic if there exists such a graph isomorphism between them and we write

$$\mathcal{G} \simeq \mathcal{H}.$$

$\simeq$ is an equivalence relation.

Example 7.1. Two isomorphic graphs $\mathcal{G}$ (LHS) and $\mathcal{H}$ (RHS):



Definition 7.2 (Graph Invariant)**.** Let $\mathfrak{G}$ be the set of all graphs and X a set of structures (e.g. $\mathbb{N}$ or strings over the alphabet $A = \{0, 1\}$).
A function

$$f : \mathfrak{G} \to X$$

is called a graph invariant if for all two graphs $\mathcal{G}, \mathcal{H} \in \mathfrak{G}$ holds

$$\mathcal{G} \simeq \mathcal{H} \quad \Rightarrow \quad f(\mathcal{G}) = f(\mathcal{H}).$$

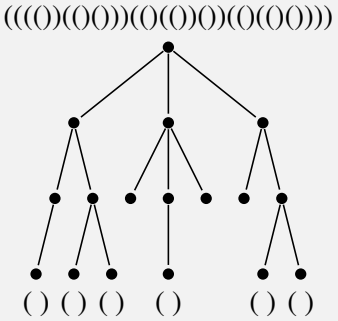
Example. Some examples of graph invariants are:

- the number of vertices and the number of edges
- the cyclomatic number
- the laplacian spectrum of a graph
- the number of cycles of a specific length in a MCB
- the "degree sequence", i.e. the list of the number of vertices for each degree, ordered by degree
- the chromatic number
- whether a graph is planar
- whether a graph is a cograph

Algorithm 7.1 (Isomorphism of Trees)**.** An isomorphism between trees can be constructed through the following tree invariant:

1) Isomorphism of rooted ordered trees:
   - Assign a pair of parantheses to each leaf vertex.
   - For every parent vertex, the label is

     (*labels of the children in order*)



The label of the root is an invariant of the rooted ordered tree.

Often the label is binary encoded through $'('\to 0$ and $')' \to 1$.

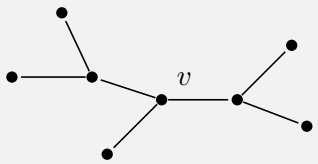
2) Isomorphism of rooted unordered trees:
The procedure is identical to rooted ordered trees, however, to create unique labels for all vertices, the label of each parent vertex is

(*lexicographically ordered list of the labels of the children*)

3) Isomorphism of trees:
For an arbitrary tree $T = (V, E)$ and a vertex $v \in V$, the eccentricity of v is the length of the longest path to all other vertices

$$\mathrm{ecc}(v) := \max_{x \in V} \mathrm{dist}(v, x).$$

We can then define the center of the tree $\mathrm{center}(T)$ as a vertex of minimal eccentricity.



The center must be on the longest path and also must be the center of this path, either the exact central vertex or the two vertices incident to the central edge.
$\Rightarrow$ 1 or 2 unique centers of the tree.

Using the unique label string of the unordered tree rooted at the center (or the lexicographically smaller one of the unique label strings from both possible roots), we get an invariant for trees in general.

$\Rightarrow$ polynomial-time algorithm for tree isomorphism.

Remark. *The discriminating co-trees of isomorphic cographs are also isomorphic. Additionally, as co-trees can be computed from cographs in linear time, the cograph isomorphism can be solved in linear time.*

Another graph category with a poly-time algorithm for isomorphism tests are planar graphs. For those, we utilize the following concept:

Algorithm 7.2 (WEINBERG's Code)**.** Let $\mathcal{G} = (V, E)$ be a planar 3-connected graph. The idea is to find a code for any edge on the planar drawing of the graph.

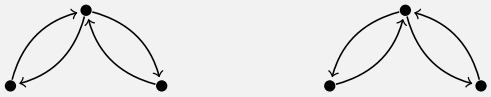
We start with a plane drawing of the graph. This construction works on any embedding but due to $\mathcal{G}$ being planar we can assume a plane drawing w.l.o.g.

- Replace every edge by 2 directed edges in opposite directions resulting in $|E'| = 2|E|$ (creates an even number of edges and even degrees, required for Euler-Tour)

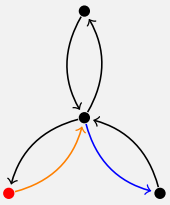


- choose either clockwise or anti-clockwise orientation for the edges (typically anti-clockwise is chosen)

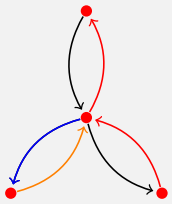
clockwise: anti-clockwise:



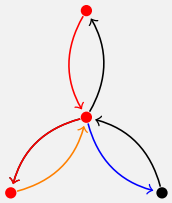
- $\forall\, e \in E'$ create tour code as follows:
  - (i) start by going along e and keep track of visited edges and vertices
  - (ii) When a new vertex is reached ($\rightarrow$), exit ($\rightarrow$) to the first unused edge on the right (from entering perspective):



(iii) When a previously encountered vertex (•) is reached (→), and the "back-edge" (i.e. the edge opposing the current one) is unused, exit (→) along the "back-edge":



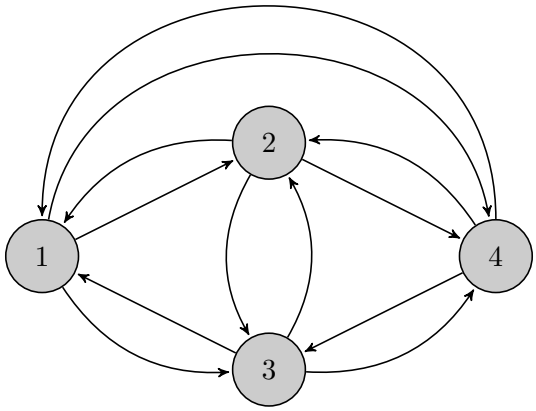
(iv) When a previously encountered vertex (•) is reached (→), and the "back-edge" (i.e. the edge opposing the current one) is has already been used (→), exit (→) to the first unused edge on the right:



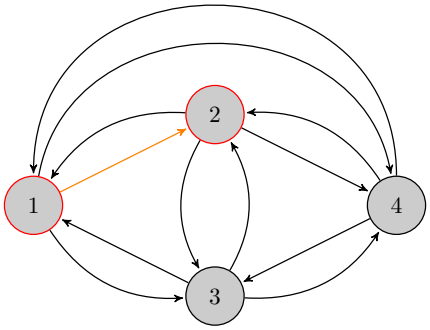
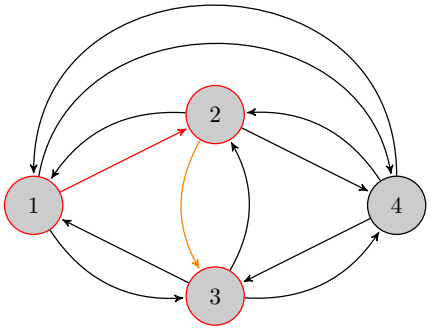
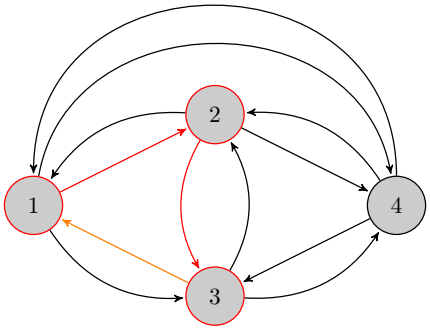
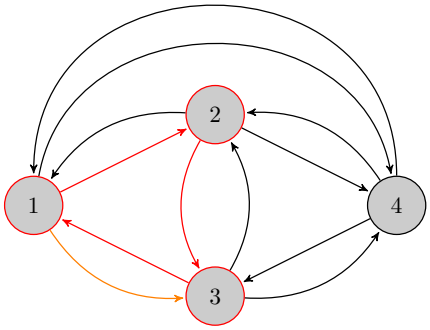
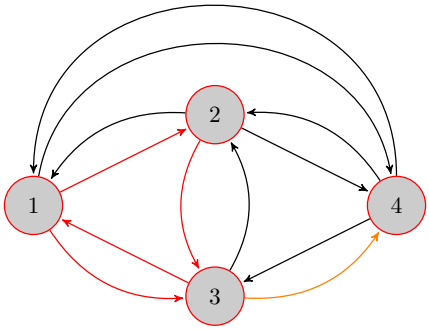
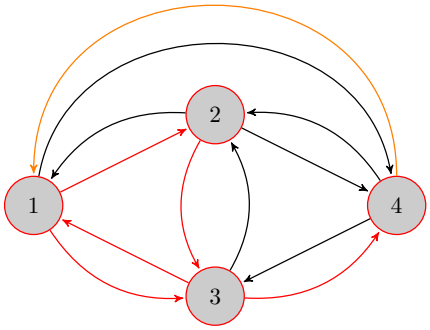
(v) The code is the sequence of visited vertices.

Applying this gives us the "right code" (anti-clockwise decisions during the tour) for each edge. Repeat same process for all edges using left instead of right, i.e. with clockwise decisions during the tour.

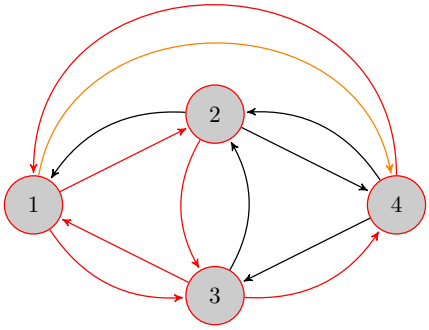
Example 7.2. Consider the following graph where the undirected edges have already been replaced by 2 opposing directed edges each:



The "right code" for the edge $(1, 2)$ is constructed as follows:

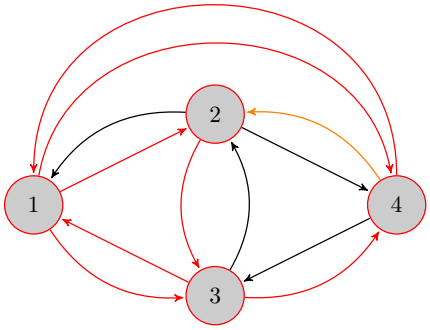
Step 1	Step 2
	
Tour-Code 1,2	Tour-Code 1,2,3
Step 3	**Step 4**
	
Tour-Code 1,2,3,1	Tour-Code 1,2,3,1,3
Step 5	**Step 6**
	
Tour-Code 1,2,3,1,3,4	Tour-Code 1,2,3,1,3,4,1

Step 7



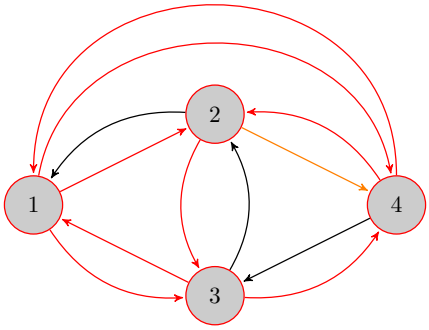
Tour-Code
1,2,3,1,3,4,1,4

Step 8



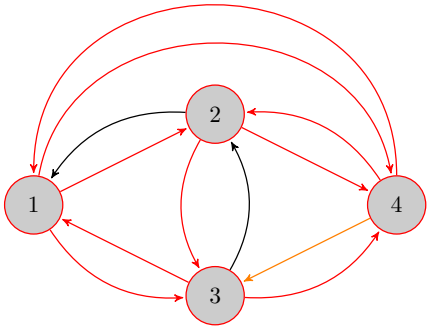
Tour-Code
1,2,3,1,3,4,1,4,2

Step 9



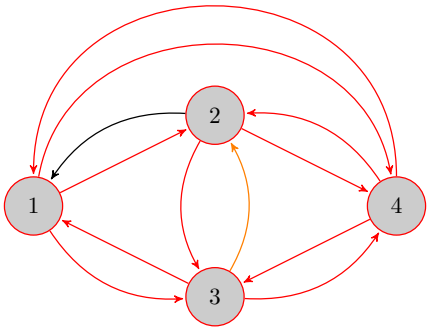
Tour-Code
1,2,3,1,3,4,1,4,2,4

Step 10



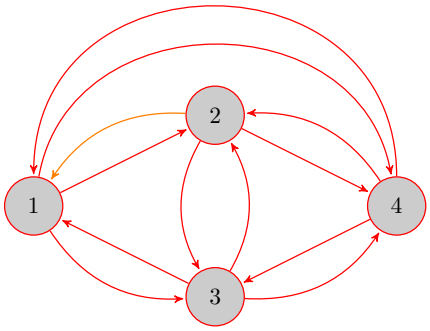
Tour-Code
1,2,3,1,3,4,1,4,2,4,3

Step 11



Tour-Code
1,2,3,1,3,4,1,4,2,4,3,2

Step 12



final Tour-Code
1,2,3,1,3,4,1,4,2,4,3,2,1

This process needs to be repeated for every other edge and then entirely again for all edges in clockwise direction.

Algorithm 7.3. Let $\mathcal{G} = (V, E)$ be a planar 3-connected graph.
By construction, $\mathcal{G}$ has $4 \cdot |E|$ Weinberg's codes, which each have a length of $2|E| + 1$.
As $\mathcal{G}$ is planar we know that $3|V| - 6 \geq |E|$.
$\Rightarrow$ we can use number of vertices for complexity estimation.

- Store all codes in a matrix of size $O(|V|^2)$.
- Sort the matrix rows using radir or distribution sort (assuming that all vertices are unique integers).
- 2 graphs $\mathcal{G}$ and $\mathcal{H}$ are isomorphic if and only if their sorted matrices are identical.

Algorithm 7.4 (WEISFEILER-LEHMANN Isomorphism Test)**.**
A heuristic for general graph isomorpism, i.e. regarding two graphs $\mathcal{G}_1$ and $\mathcal{G}_2$

- if the test fails, $\mathcal{G}_1$ and $\mathcal{G}_2$ are not isomorphic,
- if the test is successful, $\mathcal{G}_1$ and $\mathcal{G}_2$ might be isomorphic.

For the general idea:
Through an iterative process are the neighborhood characteristics (with incrementing distance) compared between the two graphs. As these features are stored in each vertex, the algorithm checks for at least one existing bijection by mapping the histograms of the feature codes onto each other.

Onto the actual algorithm:

- In each step i, every vertex n is annotated by a tuple $C_{i,n}$, containing the old compressed label and a multi-set of the old compressed labels of the neighboring vertices

  multi-set: the element order is irrelevant but elements may appear more than once
- on every step we need to recompute the (compressed) labels through a suitable hash function (e.g. "increasing integers")
- for notation we will use $C_{i,n}$ as tuples and $\mathfrak{C}_{i,n}$ as compressed labels
- each step is calculated on both graphs in parallel and they share the same hash function

Steps:

1) $\forall\, k \in V :$ define $\mathfrak{C}_{0,k} = 1$

2) for every iteration $i \geq 1$ and $\forall\ k \in V$:

$$C_{i,k} = (\mathfrak{C}_{i-1,k}, \{\mathfrak{C}_{i-1,x} \mid x \in \mathrm{N}(k)\})$$
$$\mathfrak{C}_{i,k} = \mathrm{hash}(C_{i,k})$$

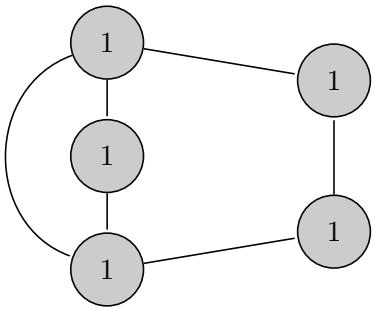
3) after each iteration step:

- partition the nodes by their compressed labels $\mathfrak{C}_{i,k}$, i.e. construct histogram over all present labels
- test if both graphs have the same partition, i.e. is the multi-set of compressed labels identical for both graphs
- if the partitions are not the same/ diverge, then reject $\mathcal{G}_1$ and $\mathcal{G}_2$ as not isomorphic

4) otherwise repeat steps 2) and 3) until $i = |V|$ or partitions don't change.

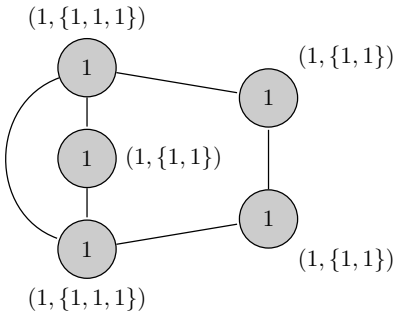
The following showcases the iterative process for one graph. For a full example on two graphs see Example 7.4.

Example 7.3. Consider the following graph $\mathcal{G} = (V, E)$ annotated initial compressed labels $\mathfrak{C}_{0,k}$ for each vertex $k \in V$.



For simplicity, the old compressed label will be written inside the nodes and the current annotated tuple in its proximity. On the right hand side, current continuation of the hash table.

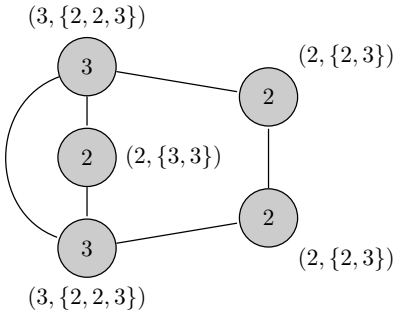
$i = 1$



hash	
initial	1
$(1, \{1, 1\})$	2
$(1, \{1, 1, 1\})$	3

$\Rightarrow$ Compare multi-set $\{2, 2, 2, 3, 3\}$ to the second graph. Halt on mismatch.

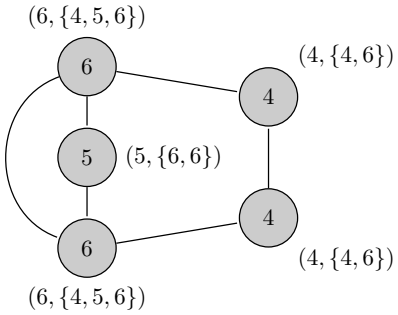
$i = 2$



hash	
$(2, \{2, 3\})$	4
$(2, \{3, 3\})$	5
$(3, \{2, 2, 3\})$	6

$\Rightarrow$ Compare multi-set $\{4, 4, 5, 6, 6\}$ to the second graph. Halt on mismatch.

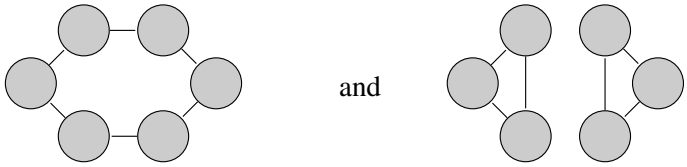
$i = 3$



hash	
$(4, \{4, 6\})$	7
$(5, \{6, 6\})$	8
$(6, \{4, 5, 6\})$	9

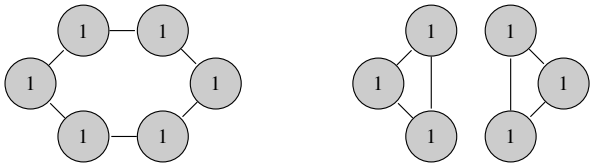
. . . Pattern should be understandable from here.

Example 7.4. Considering the two graphs



For better readability (and as they are identical for all vertices on each step), the calculated tuples for each vertex are omitted.

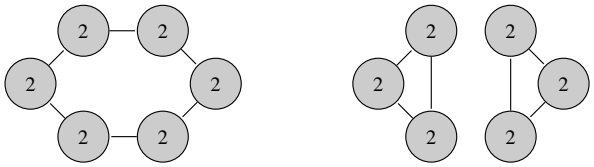
$i = 1$



hash	
initial	1
$(1, \{1, 1\})$	2

where all vertices on both graphs have the same label $(1, \{1, 1\})$.
Thus the multi-set over all new compressed labels is $\{2, 2, 2, 2, 2, 2\}$ for both graphs.

$i = 2$



hash	
$(2, \{2, 2\})$	3

where all vertices on both graphs have the same label $(2, \{2, 2\})$.
Thus the multi-set over all new compressed labels is $\{3, 3, 3, 3, 3, 3\}$ for both graphs.

$\Rightarrow$ for all steps up to $i = |V| = 6$ the multi-set over the present compressed edges is identical between the two graphs, therefore, the Weisfeiler-Lehmann Isomorpism Test is successful.
However, both graphs are obviously not isomorphic.

8. Random Graphs

Definition 8.1. Let $\mathcal{G} = (V, E)$ be a graph with $|V| = n$ and $V = \{0, 1, \dots, n-1\}$. Define the binomial coefficient $\binom{n}{k}$ with $n \geq k \geq 0$ as

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}$$

possibilities to draw k elements from n, ignoring order

Definition 8.2 (Uniform Random Graphs, aka. original Erdős-Rényi model)**.**
Let $\mathfrak{G}_{n,m}$ be the family of all graphs with $|V| = n$ and $|E| = m$.
For any $\mathcal{G} \in \mathfrak{G}_{n,m}$, we assign the same probability

$$\mathbb{P}(\mathcal{G}) = \binom{\binom{n}{2}}{m}^{-1}.$$

Definition 8.3 (Binomial Random Graphs, aka. Erdős-Rényi-Gilbert model)**.**

- start with graph with n vertices and no edges
- every edge has the same probability $p \in [0, 1]$ to be added

Let $\mathfrak{G}_{n,p}$ be the family of all graphs with $|V| = n$ and probability p for an edge to appear.
For any $\mathcal{G} = (V, E) \in \mathfrak{G}_{n,p}$, we assign the probability

$$\mathbb{P}(\mathcal{G}) = p^m \cdot (1-p)^{\binom{n}{2}-m} \quad \text{with } |E| = m.$$

Lemma 8.4. Any binomial graph with n vertices and m edges has the same probability

$$\binom{\binom{n}{2}}{m}^{-1}.$$

Proof (simplified). Let $\mathcal{G} = (V, E)$ be a graph.
The calculation of *conditional* probability is given by

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

The probability for the graph then comes out to

$$\begin{aligned}\mathbb{P}(\mathcal{G} \in \mathfrak{G}_{n,p} \mid |E| = m) &= \frac{\mathbb{P}(\mathcal{G} \in \mathfrak{G}_{n,p} \mid |E| = m)^{(*)}}{\mathbb{P}(\mathcal{G} \in \mathfrak{G}_{n,m})} \\ &= \frac{p^m \cdot (1-p)^{\binom{n}{2}-m}}{\binom{\binom{n}{2}}{m} \cdot p^m \cdot (1-p)^{\binom{n}{2}-m}}\end{aligned}$$

$(*)$ reduced superset of possibilities

□

Observation (Properties of $\mathfrak{G}_{n,p}$)**.**

- graphs have on average $\binom{n}{2} \cdot p$ edges
- distribution, that any particular vertex has degree k is

$$\mathbb{P}(\deg(v) = k) = \binom{n-1}{k} \cdot p^k \cdot (1-p)^{n-1-k}.$$

As this is a binomial distribution we know the expected value and variance:

$$\begin{aligned}\mathbb{E}(\deg(v) = k) &= p \cdot (n-1) \\ \sigma^2 &= p \cdot (1-p) \cdot (n-1)\end{aligned}$$

Definition 8.5 (Independence Number)**.** Let $\mathcal{G} = (V, E)$ be a graph.
The independence number

$$\alpha(\mathcal{G}) = \max\{|V'| \mid V' \subseteq V \ \wedge \ \forall\, x, y \in V' : \ xy \notin E\}$$

is the size of the largest set of vertices, s.t. no two vertices are adjacent.

The opposite of the independence number is the (maximum) clique in a graph.

Definition 8.6 ((Maximum) Clique)**.** Let $\mathcal{G} = (V, E)$ be a graph.
The (maximum) clique

$$\mathrm{cli}(\mathcal{G}) = \max\{|V'| \mid V' \subseteq V \ \wedge \ \forall\, x, y \in V' : \ xy \in E\}$$

is the size of the largest set of vertices, where every pair of vertices is adjacent.

Observation. Interesting properties of graphs $\mathcal{G} \in \mathfrak{G}_{n,p}$:

- the probability that the independence number is at least k:

$$\mathbb{P}(\alpha(\mathcal{G}) \geq k) \leq \binom{n}{k} \cdot (1-p)^{\binom{k}{2}}$$

- the probability that the (maximum) clique is at least k:

$$\mathbb{P}(\mathrm{cli}(\mathcal{G}) \geq k) \leq \binom{n}{k} \cdot p^{\binom{k}{2}}$$

- What is the expected number of cycles of length k for any $\mathcal{G} \in \mathfrak{G}_{n,p}$?
  - Let $(n)_k := n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-k+1)$.
    possibilities to draw k elements ordered from n
  - Every k-cycle has $2k$ sequences (k random starting points, 2 walking directions).
  - Let $|C_k|$ be the number of *possible* k-cycles in a graph of $\mathfrak{G}_{n,p}$. We know

$$|C_k| = \frac{(n)_k}{2k}$$

    resulting in the expected value

$$\mathbb{E}(X) = \frac{(n)_k}{2k} \cdot p^k.$$

Bibliography