Cycles and Bicycles

Peter F. STADLER

Institut für Theoretische Chemie, Universität Wien

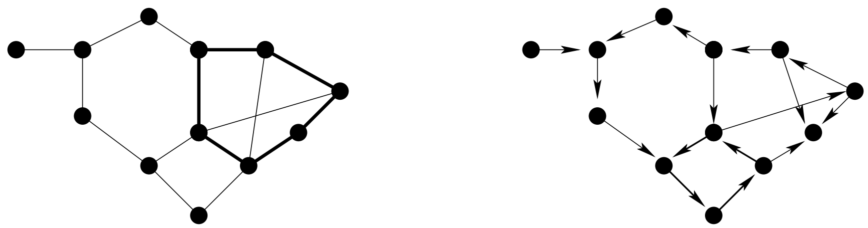
Santa Fe Institute, New Mexico

http://www.tbi.univie.ac.at/~studla

Bled, Slovenia, January 2002

Undirected Graphs

We consider cycles in simple undirected or directed graphs $G(V, E)$.
$E \ldots$ edge set
$V \ldots$ vertex set



In the directed case we distinguish
cycles ... without orientation
circuits ... following the orientation of the edges.

$U \subseteq E \ldots |E|$-dimension vector (indexed by the edges):

A cycle (=subgraph with even vertex degrees) is an edge-disjoint union of *elementary cycles*.

TWO-CONNECTED GRAPHS only: every edge is contained in a cycle.

Vector Spaces of Edges

$$U_e = \begin{cases} 1 & \text{if} \quad e \in U \\ 0 & \text{if} \quad e \notin U \end{cases}$$

Incidence matrix $\mathbf{H}$ of G:

$$H_{xe} = \begin{cases} 1 & \text{if} \quad x \in e \\ 0 & \text{if} \quad x \notin e \end{cases}$$

All cycles satisfy

$$\mathbf{H}U = 0 \text{ over GF}(2) = (\{0,1\}, \oplus, \cdot)$$

Cycle space $\mathfrak{C}$ = vector space spanned by cycles

Dimension:

$$\dim \mathfrak{C} = \gamma(G) = |E| - |V| + \text{components}(G)$$

Bases of a Vector Space

A set $\{x_1, \ldots, x_L\}$ of vectors is *linearly independent* if the linear equation

$$\lambda_1 x_1 + \lambda_2 x_2 + \ldots \lambda_l x_l = 0$$

has no solution except $\lambda_1 = \lambda_2 = \cdots = \lambda_L = 0$.

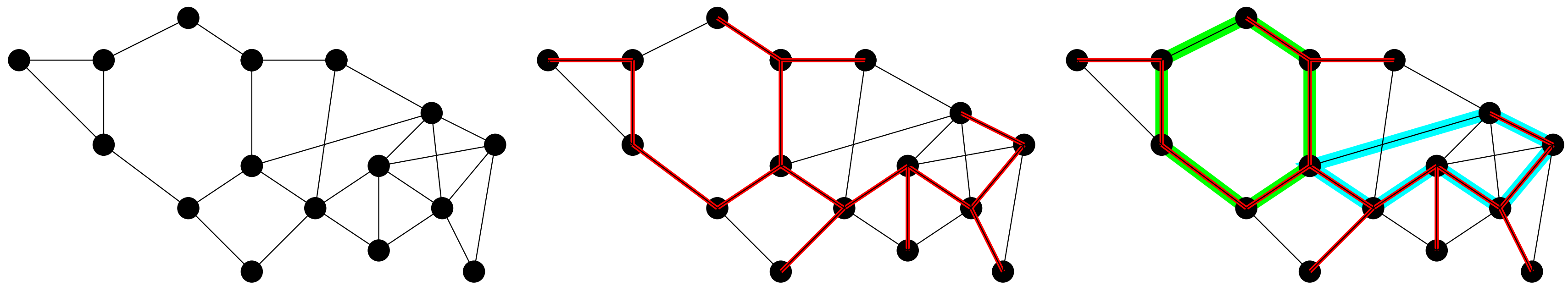
A basis of a vector space is a maximal set of linearly independent vectors.

Each vector x can be written as a linear combination of the basis elements $\mathcal{B} = \{y_1, y_2, \ldots, y_n\}$:

$$x = \lambda_1 y_1 + \lambda_2 y_2 + \cdots + \lambda_n y_n$$

Cycle Bases

Kirchhoff basis:



graph $\Rightarrow$ spanning tree $\Rightarrow$ cycles $C(T,e)$ for all $e \notin T$.

A basis of $\mathcal{B}$ is a Kirchhoff basis (or a *strictly fundamental* basis) of $\mathfrak{C}$ if there is a spanning tree T such that $\mathcal{B} = \{C(T,e) | e \in E \backslash T\}$.

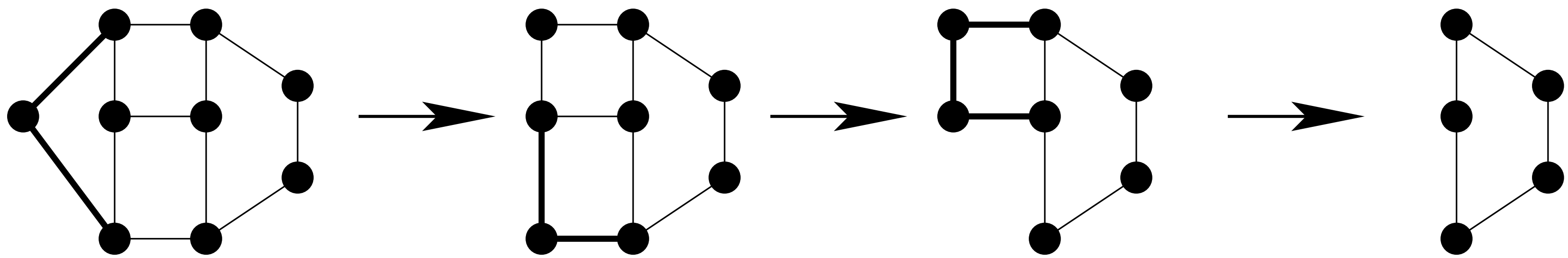
Fundamental Cycle Bases

A collection of $\nu(G)$ cycles in G is called *fundamental* if there is an ordering of these cycles such that

$$C_j \setminus (C_1 \cup C_2 \cup \cdots \cup C_{j-1}) \neq \emptyset \qquad \text{for } 2 \leq j \leq \nu(G)$$

Strictly Fundamental implies fundamental but not *vice versa*.

Ear Decomposition



Every two-connected graph has an ear decomposition. Each ear decomposition defines a basis of the cycle space $\mathfrak{C}$.

Minimal Length Cycle Bases

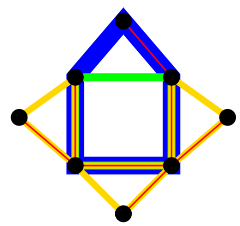
Length $|C|$ of cycle C = number of edges

Length of a cycle basis $\ell(\mathcal{B}) = \sum_{C \in \mathcal{B}} |C|$.

Relevant Cycles (Plotkin '71, Vismara '97);

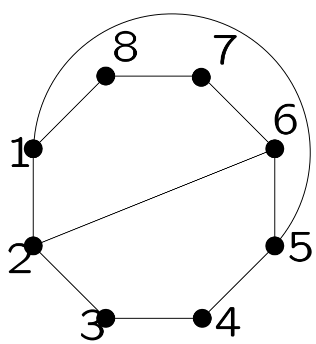
C is contained in a minimal cycle basis
$\Longleftrightarrow$
C cannot be written as a $\oplus$-sum of shorter cycles

Some Counter examples

- Not every MCB is strictly fundamental (Horton, Deo) 

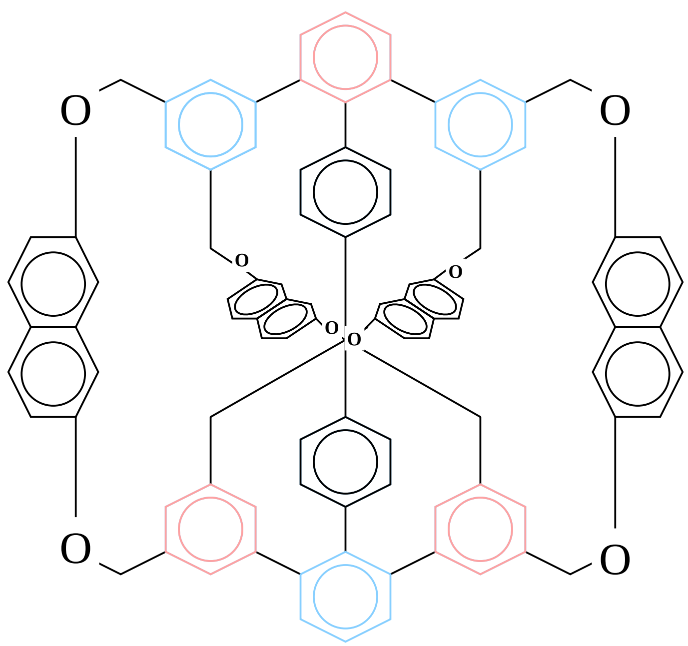
- Not every MCB is fundamental (examples are quite complicated)

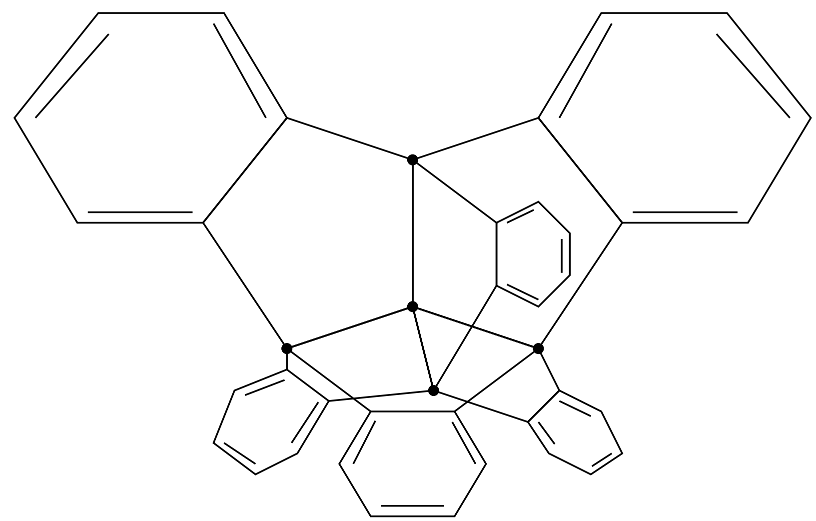
- The MCB of a planar graph not necessarily consists of faces



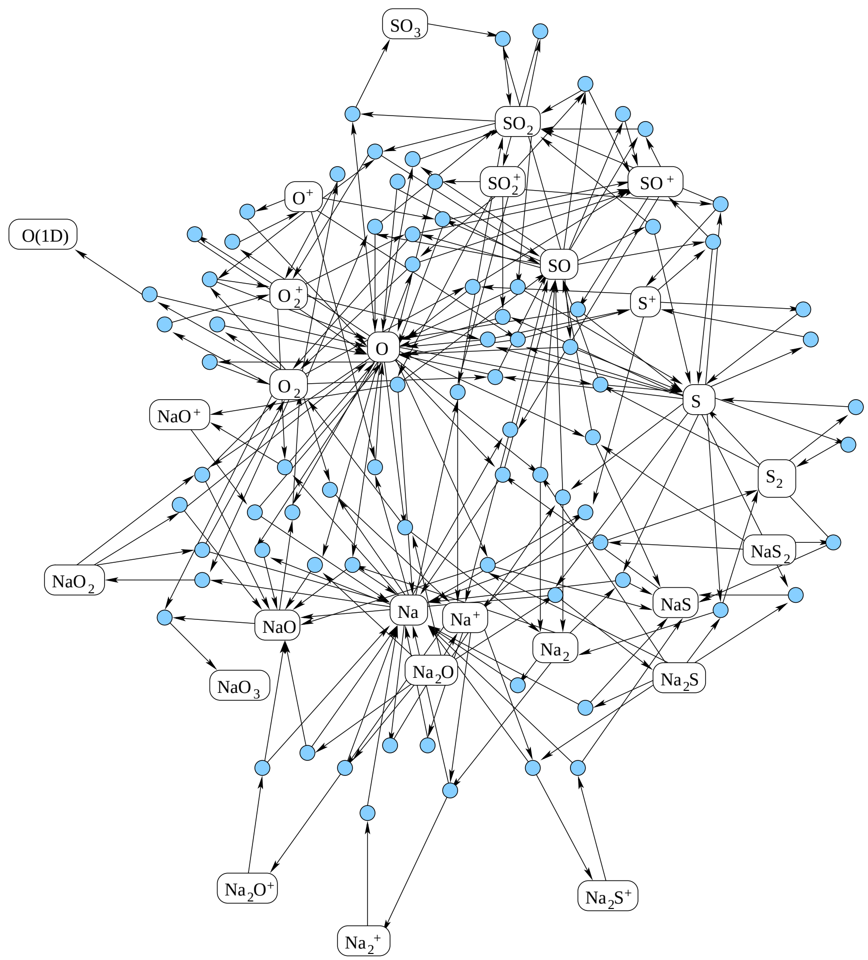
Who cares about MCBs?

Chemical Ring Perception (SSSR).





Analysis of chemical reaction networks.



Reaction network of Io's athmosphere

Matroid Property

The cycles of G form a matroid $\Longrightarrow$
A minimal cycle basis is obtained from the set of all cycles by a greedy procedure:

1. Sort set $\mathcal{C}$ of cycles by length

   $\mathcal{B} \leftarrow \emptyset$

2. while($\mathcal{C} \neq \emptyset$)

   $\mathcal{C} \leftarrow \mathcal{C} \setminus \{C\}$

   if $\mathcal{B} \cup \{C\}$ independent: $\mathcal{B} \leftarrow \mathcal{B} \cup \{C\}$.

Problem: exponentially many cycles.

Necessary conditions:

elementary (all vertices have degree 2)

short (isometric) for all vertices x, y in C, the cycle C contains a shortest paths between x and y.

Horton's Polynomial Time Algorithm

A cycle is *edge short* if C contains an edge $e = \{x, y\}$ and a vertex z such that

$$C = \{x, y\} \cup P(x, z) \cup P(y, z)$$

where $P(x, z)$ and $P(y, z)$ are shortest paths.

If C is relevant then it is edge-short
(Horton'87).

Construct (at most) $|E| \times |V|$ edge-short cycles.

Horton showed that even if $P(x, z)$ is not unique one may choose any shortest path, i.e., the $|E| \times |V|$ cycles contain a minimal cycle basis.

Alternative trick [Hartvigsen'94]: small perturbation of edge length to make minimum weight cycle basis unique.

Graph Operations

The length of minimal cycle bases does not behave "well" under many simple graph operations:

G_1 G_2

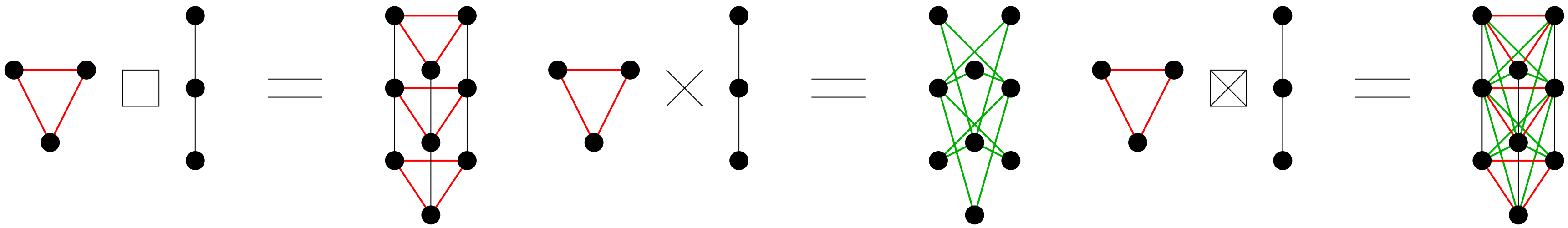
G_1 has $\nu(G_1) = 3$ and $\ell(G_1) = 38$. Deletion of a single edge leads to G_2 with $\nu(G_2) = 2$ but $\ell(G_2) = 44$.

Similar: other graph minor operations.

Cartesian and Strong Graph Products

Given two non-empty graphs $G = (V_G, E_G)$ and $H = (V_H, E_H)$:

Cartesian product $G \Box H$:
Vertex set $V_G \times V_H$
Edges: $(x_1, x_2)(y_1, y_2)$ is an edge in $E_{G \Box H}$ iff either $x_2 = y_2$ and $x_1 y_1 \in E_G$ or if $x_1 = y_1$ and $x_2 y_2 \in E_H$



Direct product $G \times H$:
Vertex set $V_G \times V_H$
Edges: $(x_1, x_2)(y_1, y_2)$ is an edge if $x_1 y_1 \in E_G$ and $x_2 y_2 \in E_H$

Strong product $G \boxtimes H$:

Vertex set $V_G \times V_H$

Edges: those of the direct product and those of the Cartesian product

Relevant Cycles in Product Graphs

$G = (V_G, E_G)$ and $H = (V_H, E_H)$ be two non-empty graphs,
T_G and T_H spanning trees,
$\mathcal{B}_G$ and $\mathcal{B}_H$ cycle bases of G and H.

Hammack's Basis for Cartesian Products (1999):

$$\mathcal{H}_1 = \{e\Box f | e \in T_G,\, f \in T_H\}$$
$$\mathcal{H}_2 = \{C^y | C \in \mathcal{B}_G,\, y \in V_H\}$$
$$\mathcal{H}_3 = \{{}^xC | x \in V_G,\, C \in \mathcal{B}_H\}$$
$$\mathcal{B}^* = \mathcal{H}_1 \cup \mathcal{H}_2 \cup \mathcal{H}_3$$

$\mathcal{B}^*$ is i general **not** minimal even if $\mathcal{B}_G$ and $\mathcal{B}_H$ are minimal.
- *Counterexample:* $C_5 \Box K_2$

Hammack basis: 4 squares and 2 pentagons.
Minimal length basis: 5 squares and 1 pentagon.

Hammack's basis is minimal if $\mathcal{B}_G$ and $\mathcal{B}_H$ consist of triangles and squares only.

Consider a triangle-free graph for the moment.

Idea. Start with the Hammack basis and replace as many cycles in the fibres as possible by squares from
$\mathcal{C}_\Box = \{e \Box f | e \in E_G, f \in E_H\}$.

Lemma. For all $C \in \mathcal{B}_G$ and all $x, y \in V_H$ there is a collection of squares in $\mathcal{C}_\Box$ such that
$C^x = C^y \oplus$ squares.

It is hence sufficient to have one copy $\mathcal{B}_G$ and one copy of $\mathcal{B}_H$ in one G and one H-fibre. The rest of the basis can be completed from $\mathcal{C}_\Box$.

To show that there are no relevant cycles that are not contained in $\mathcal{C}_\square$ or a fibre we consider the following procedure:

Set $\delta(x)$ = sum of distance of $x \in G\square H$ from two fixed fibres. Define for any cycle C^* in $\square H$:

$$\delta(C^*) = \sum_{x \in V_{C^*}} \frac{\mathrm{deg}_{C^*}(x)}{2} \delta(x),$$

We show that we keep adding squares from $\mathcal{C}_\square$ to C^* until we arrive at $\delta(C^{(k)}) = 0$ and $|C^{(k)}| \leq |C^*|$. Since C^* is either strictly shorter than C^* or it is the edge-disjoint union of a cycle in ${}^xH$ and a cycle in G^y C^* cannot be relevant.

It remains to show that it is impossible to replace any further basis cycle by squares from $\mathcal{C}_\square$. (not hard)

For graphs with triangles: retain triangles in each fibre and use the longer basis cycles in a single fibre only.

Total Basis Length in Iterated Products

$$\begin{aligned}\ell(G\Box H) = \ell(G) + \ell(H) + \\ 3\big[t_G(|V_H|-1) + 3t_H(|V_G|-1)\big] \\ + 4\big[(|E_G|-t_G)(|V_H|-1) + |E_H|-t_H)(|V_G|-1) - \\ (|V_H|-1)(|V_G|-1)\big]\end{aligned}$$

Substitute

$$G^n = G\Box G^{n-1} \simeq G^{n-1}\Box G, \qquad G^1 = G$$

and set $a = |E|/|V|$ and $\tau = t_G/|V|$, where t_G is the number of triangles in the MCB of G.

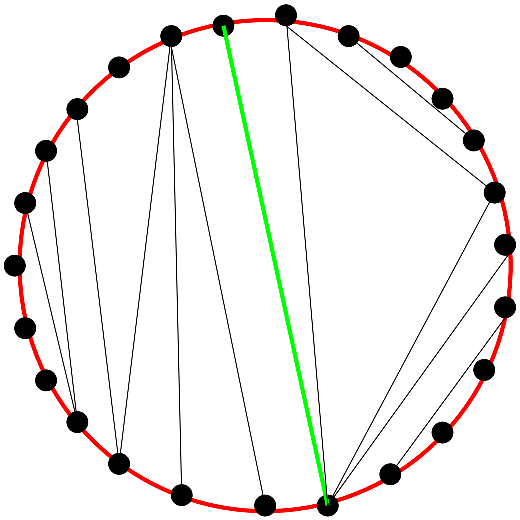
$$\begin{aligned}\ell(G^{n+1}) = \ell(G) + \ell(G^n) + 3\tau\big[|V|(|V^n|-1) + n|V|^n(|V|-1)\big] \\ + 4(a-\tau)\big[|V|(|V|^n-1) + n|V|^n(|V|-1) \\ - (|V|^n-1)(|V|-1)\big].\end{aligned}$$

Dividing by $\nu(G^{n+1})$ and setting $\xi = 1/V$ eventually yields:

$$L_\infty = \lim_{n\to\infty} L_n = 3\frac{\tau}{a} + 4\frac{a-\tau}{a}.$$

Graphs with a Unique MCB

Outerplanar graphs
Two-connected o.p. graphs have a Hamiltonian Cycle H and
each chord separates G into 2 two-connected outerplanar graphs G_1 and G_2.

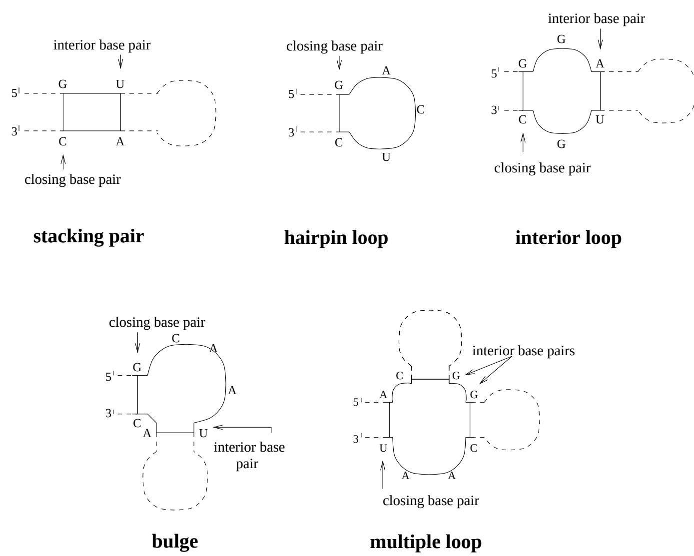


Take an edge e in H. The shortest cycle C through e contains at least one chord, hence C is a member of the MCB. Split C along the chords and repeat.

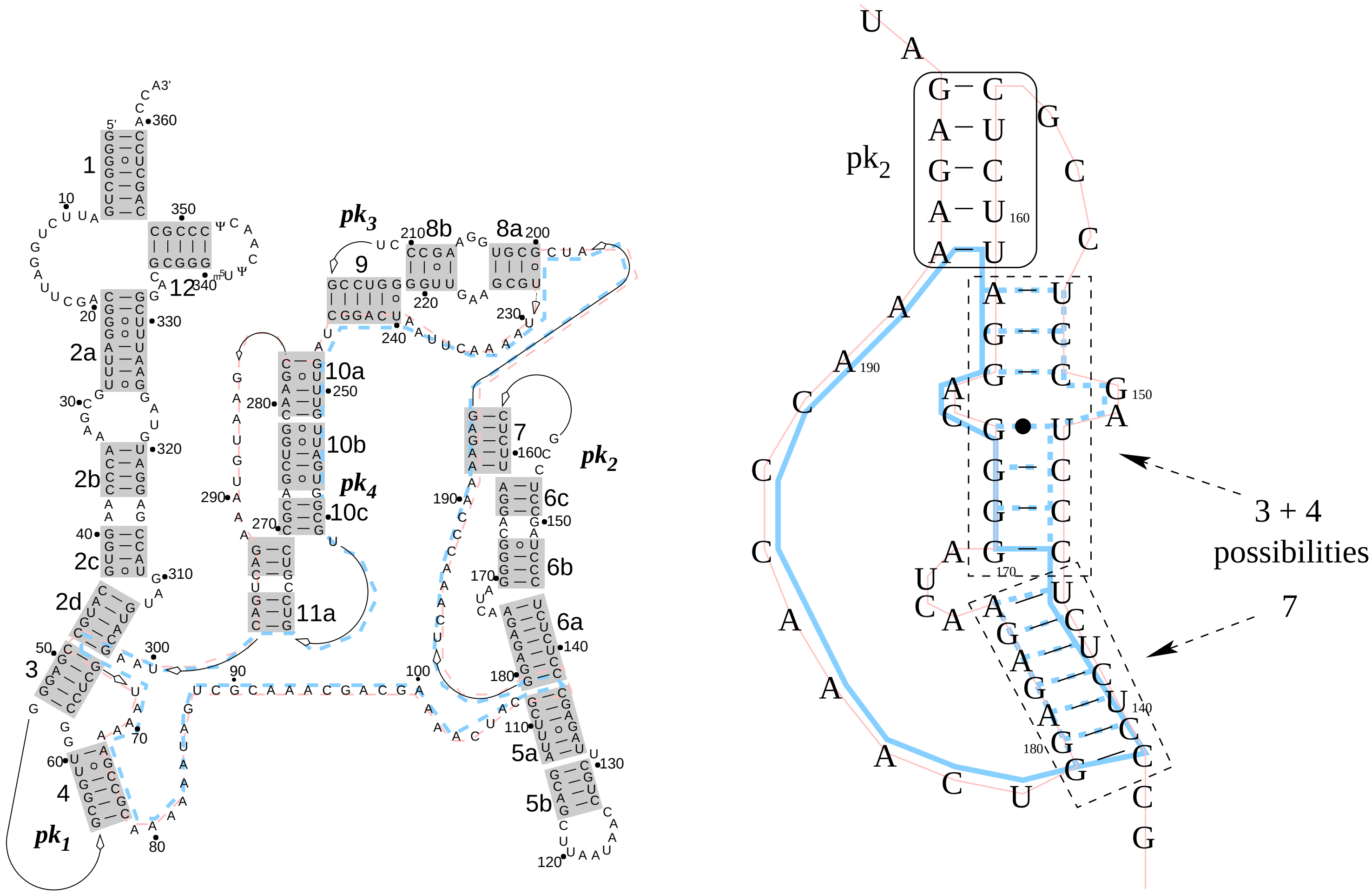
RNA Secondary Structures

outerplanar.

"loops" = cycles of the **unique** MCB.



Pseudoknots: MCB usually not unique.



tmRNA from E.coli with its pseudoknots

Exchangability of Relevant Cycles

Set $\mathcal{R}$ of relevant cycles of an undirected graph can be computed efficiently by Vismara's algortithm (1997).

Def.: $C, C' \in \mathcal{R}$ are exchangable, $C \leftrightarrow C'$, if there is a set $\mathcal{Q}$ of relevant cycles such that
(i) $|C''| \leq |C| = |C'|$ for all $C'' \in \mathcal{Q}$,
(ii) $\mathcal{Q} \cup \{C'\}$ is linearly independent, and
(iii) $C' = C \cup \bigoplus \mathcal{Q}$.

Theorem. $C \leftrightarrow C'$ is an equivalence relation.

Surprisingly tedious to prove ...

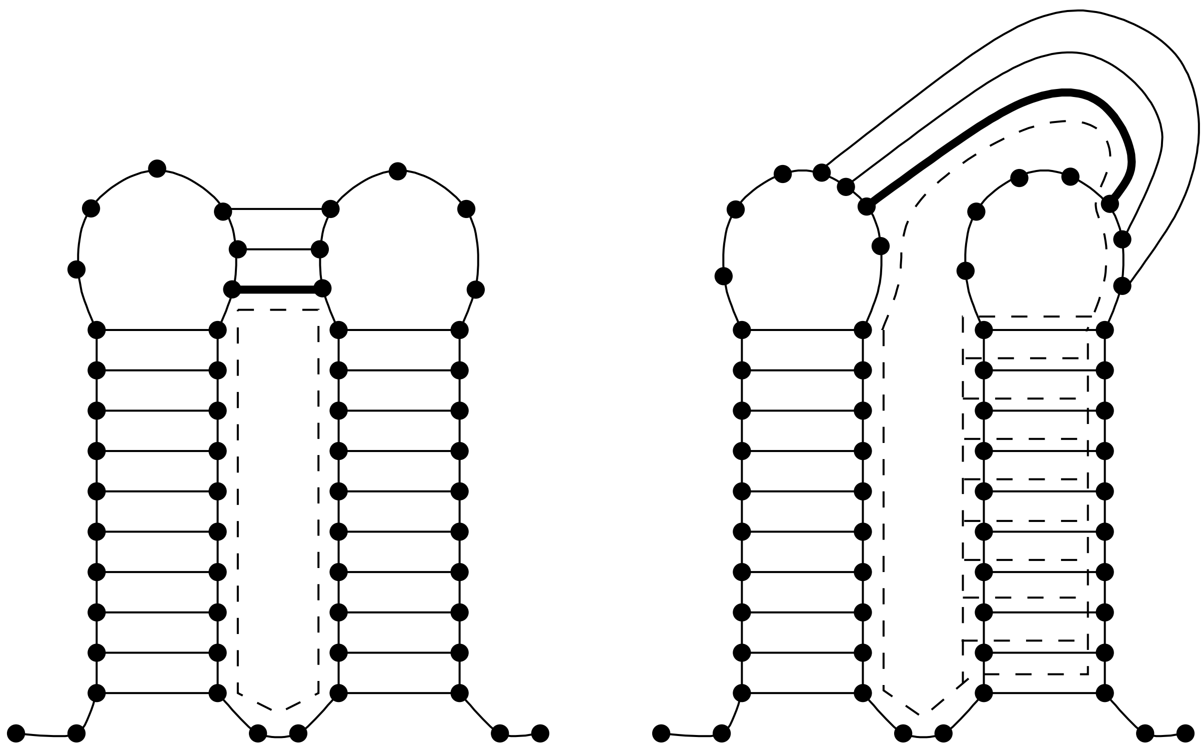
uses explicitly that we work over GF(2), i.e.,

does not work for general matroids.

Theorem. Let $\mathcal{W}$ be a $\leftrightarrow$-class and let $\mathcal{M}$ be a minimal cycle basis. Then

$$\mathrm{knar}(\mathcal{W}) = |\mathcal{M} \cap \mathcal{W}|$$

is independent of the choice of the minimal cycle basis $\mathcal{M}$.



Directed Graphs

Let $G(V, A)$ be a directed graph and a U a cycle in G.
Associated vector:

$$U_e = \begin{cases} +1 & \text{if} \quad e \in U \quad \text{in forward direction} \\ -1 & \text{if} \quad e \in U \quad \text{in backward direction} \\ 0 & \text{if} \quad e \notin U \end{cases}$$

Incidence matrix $\mathbf{H}$ of G:

$$H_{xe} = \begin{cases} -1 & \text{if} \quad x \text{ is inital point of arc } e \\ +1 & \text{if} \quad x \text{ is terminal point of arc } e \\ 0 & \text{if} \quad x \notin e \end{cases}$$

All cycles satisfy

$$\mathbf{H}U = 0 \text{ over } \mathbb{R}$$

Circuit cycle in forward direction, $C_e = 0, +1$.

Circuit Bases

Theorem. (Berge) If $G(V, A)$ is strongly connected if it has a cycle basis consisting of (elementary) circuits.

Remark. Elementary circuits generate the extremal rays of the convex cone

$$\mathbb{K} := \{U : \mathbf{H}U = 0 \quad \text{and} \quad U(e) \geq 0\}$$

How to compute a minimum length circuit basis?

Circuits again form a matroid (linear independence over $\mathbb{R}$).

$\Longrightarrow$ Greedy Algorithm.

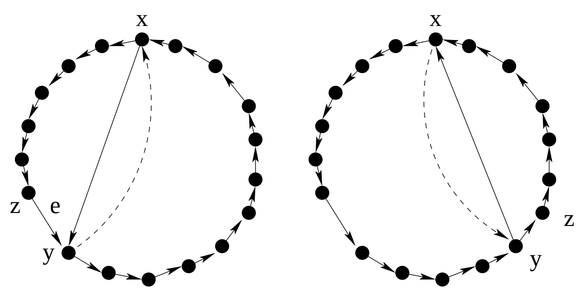
Again exponentially many circuits.

Def. A circuit C is *short* if for all vertices x and y it contains a shortest path $S[x, y]$ *or* a shortest path $S[y, x]$.

Def.: A circuit C is *arc-short* if C contains a vertex x and an arc $e = (v, w)$ such that $C = P[w, x] + P[x, v] + (v, w)$ where $P[w, x]$ and $P[x, v]$ are shortest directed paths.

Lemma. If C is short, it is arc-short

Proof.



Lemma. If C is relevant, then C is short.

Proof. C relevant but not short $\Longrightarrow$
$\exists x, y$ in C: C contains neither shortest paths $S[x, y]$ nor $S[y, x]$. Then $C^1 = C[x, y] + S[y, x]$, $C^2 = S[x, y] + S[y, x]$, and $C^3 = S[x, y] + S[y, x]$ are closed paths in G and hence are sums of (shorter) circuits. Furtermore

$$C = C[x, y] + C[y, x] = C^1 + C^2 - C^3$$

and $|C^i| < |C|$

Minimum Circuit Base

1: Compute directed distances and shortest paths with perturbed edge length. $\mathcal{O}(|V|^3)$

2: Construct $|A| \times |V|$ candidates for arc-short cycles.

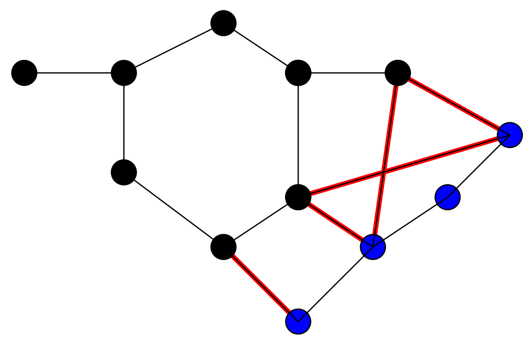
3: Check that the cycles are elementary. $\mathcal{O}(|V|)$ for each cycle, i.e., $\mathcal{O}(|A| \times |V|^2)$

4: Greedy step. At most $|A| \times |V|$ Gauss eliminations on a $(\nu(G) + 1) \times |E|$ matrix, i.e., at most $\mathcal{O}(\nu(G)|E|^2 \times |V|)$.

For most graphs probably much faster.

Cuts

Let (V_1, V_2) be a partition of the vertex set V, i.e., $V_1, V_2 \neq \emptyset$ and $V_1 \cup V_2 = V$.



A **cut** or **cocycle** $K = \langle V_1, V_2 \rangle$ is the set of all edges in G that have one end in V_1 and one end in V_2.

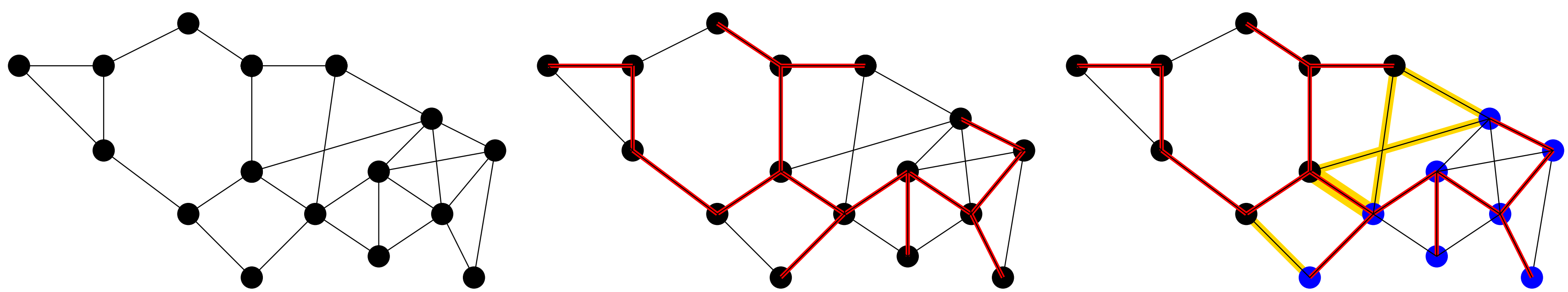
The cuts form a vector space $\mathfrak{K}$ over $(\{0,1\}, \oplus, \cdot)$ with dimension $|V| - 1$.

Fundamental Cuts

A basis is again obtained from a spanning tree: Let $b \in T$. Removal of b disconnects the tree T into exactly two subtrees with vertex sets $V_1^{T,b}$ and $V_2^{T,b}$. The cut

$$\text{cut}(T,b) := \langle V_1^{T,b}, V_2^{T,b} \rangle$$

is *fundamental cut* of G.



T has $|V|-1$ edges, thus there are $|V|-1$ linearly independent fundamental cuts.

Cut Sets

A cut is a *cut set of* G if both V_1 and V_2 are connected.

$\Rightarrow$ every fundamental cut cut(T, b) is a cut set of G. ($T \setminus \{b\}$ consists of two trees)

Size of a cut: $|K|$

Length of a cut basis $\ell(\mathcal{B}) = \sum_{K \in \mathcal{B}} |K|$

Minimal basis of the cut space?

The Cut Tree

Let $G(V, E)$ be a graph, possibly with edge weights $w(e)$.

A *cut tree* $T^{\#}$ of G is a tree with vertex set V with the following property:
For every pair of distinct vertices $s, t \in V$, let e be a minimum weight edge on the unique path from s to t in $T^{\#}$. Deleting e from $T^{\#}$ separates $T^{\#}$ into two connected components $V_1^{st,e}$ and $V_2^{st,e}$ such that

$$\text{cut}(T^{\#}; e) = \langle V_1^{st,e}, V_2^{st,e} \rangle$$

is a minimum weight cut separating s, t.

The algorithms by Gomory and Hu (1961) and Gusfield (1991) compute a cut tree $T^{\#}$ and the sets cut$(T^{\#}; e)$ in $\mathcal{O}(|E||V|^2 \log |V|)$ steps.

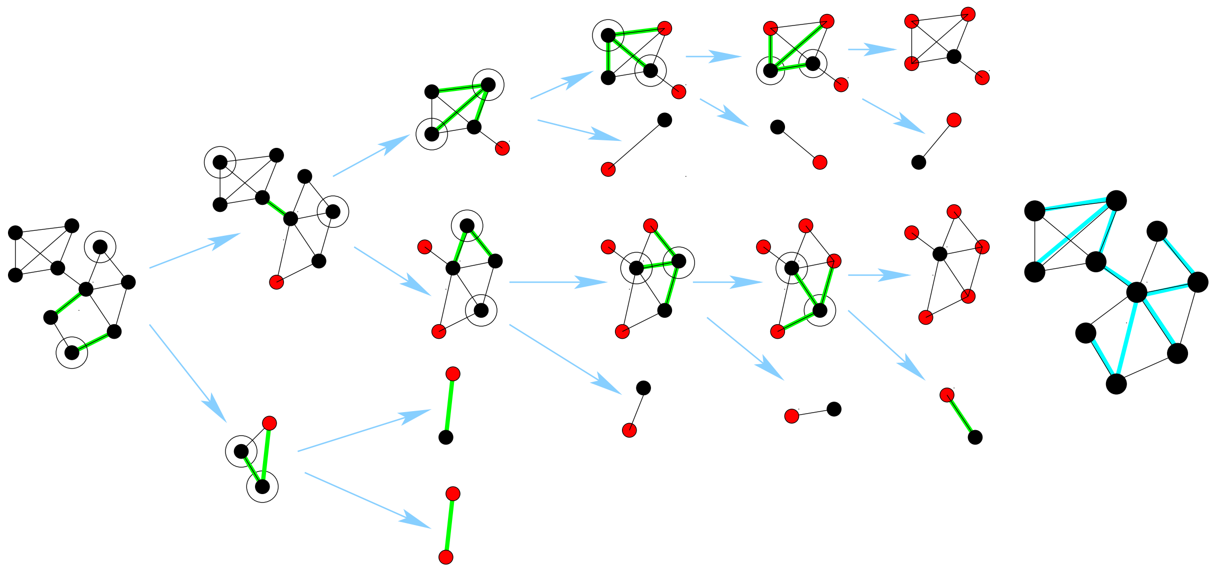
The Gomory-Hu Algorithm

In a nutshell:

(1) Pick two vertices s and t (at random) and find the minimum weight cut (V_1, V_2) that separates s and t.

(2) Form two graphs G_1 and G_2 by contracting V_2 and V_1, repectively.

(3) Repeat with both graphs until only graphs with two vertices as left.



Lemma. (see e.g. Golynski, Horton 2001)
If $T^{\#}$ is a cut tree then

$$\mathcal{M} = \{\text{cut}(T^{\#}; e) | e \in T^{\#}\}$$

is a minimum weight basis of the cut space $\mathfrak{K}$.

Proof. We use the edge-weight perturbation trick to make the Gomory-Hu tree and all cut weights unique.

Suppose $\mathcal{Q}$ is a minimal cut basis and let $H = \text{cut}(T^{\#}; e)$ be the minimum weight cut separating s and t. Then H is a $\oplus$-sum of cuts in $\mathcal{Q}$. This sum must contain a cut H' which separates s and t. Suppose $H \neq H'$. Of course, H' cannot be shorter than the cut H, hence $\mathcal{Q}' = \mathcal{Q} \setminus \{H'\} \cup \{H\}$ is shorter than $\mathcal{Q}$, a contradiction to minimality. Thus $\text{cut}(T^{\#}; e) \in \mathcal{Q}$. This holds for each of the $|V| - 1$ cuts associated with $T^{\#}$, which are linearly independent, and the lemma follows from $\dim \mathfrak{K} = |V| - 1$.

Open Question: How to compute the set of relevant cuts in the unweighted (or degenerate) case?

Relationships of Cycles and Cuts

$C \in \mathfrak{C}$ if $|C \cap K|$ is even for *all* $K \in \mathfrak{K}$.

$K \in \mathfrak{K}$ if $|C \cap K|$ is even for *all* $C \in \mathfrak{C}$.

Thus, for all $C \in \mathfrak{C}$ and all $K \in \mathfrak{K}$ $|C \cap K|$ is even, i.e.,

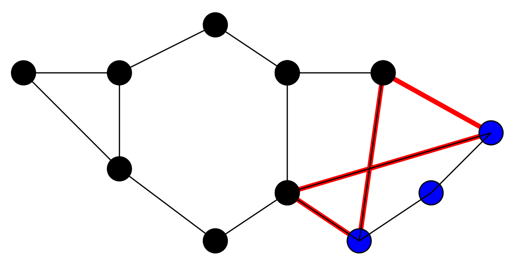
$$\bigoplus_{e \in E} C_e \cdot K_e = 0\,.$$

In other words $\mathfrak{C}$ and $\mathfrak{K}$ are "orthogonal" **over GF(2)**.

$\mathfrak{C}$ and $\mathfrak{K}$ are *orthogonal complements* iff $\mathfrak{C} \cap \mathfrak{K} = \{\emptyset\}$.

So what is a "bicycle"?

Def. A *bicycle* B is a subset of E that is both a cycle and a cocycle (cut).



Thus the *bicycle space* is $\mathfrak{B} = \mathfrak{C} \cap \mathfrak{K}$.

Some graphs have bicycles, some don't . . .

QUESTION: How can we compute (minimal) Bicycle Bases ???

Thanx!

Joint work with

Petra Gleiss
(Uni Wien, Dept. of Theoretical Chemistry)

Josef Leydold
(WU Wien, Dept. of Statistics)

Wilfried Imrich
(Montan Uni Leoben, Dept. of Mathematics)