Minimal Cycle Bases of Outerplanar Graphs

Josef Leydold[a] and Peter F. Stadler[b,c,*]

[a]Dept. for Applied Statistics and Data Processing
University of Economics and Business Administration
Augasse 2-6, A-1090 Wien, Austria
Phone: **43 1 31336-4695 Fax: **43 1 31336-738
E-Mail: Josef.Leydold@wu-wien.ac.at
URL: http://statistik.wu-wien.ac.at/staff/leydold

[b]Institut für Theoretische Chemie, Universität Wien
Währingerstraße 17, A-1090 Wien, Austria
Phone: **43 1 40480 665 Fax: **43 1 40480-660
E-Mail: studla@tbi.univie.ac.at
*Address for correspondence

[c]The Santa Fe Institute
1399 Hyde Park Road, Santa Fe, NM 87501, USA
Phone: (505) 984 8800 Fax: (505) 982 0565
E-Mail: stadler@santafe.edu
URL: http://www.tbi.univie.ac.at/~studla

Submitted: July 18, 1997; Accepted: February 27, 1998.

Abstract

2-connected outerplanar graphs have a unique minimal cycle basis with length $2|E| - |V|$. They are the only Hamiltonian graphs with a cycle basis of this length.

Keywords: Minimal Cycle Basis, Outerplanar Graphs

AMS Subject Classification: Primary 05C38. Secondary 92D20.

1. Introduction

The description of cyclic structures is an important problem in graph theory (see e.g. [16]). Cycle bases of graphs have a variety of applications in science and engineering, among them in structural analysis [11] and in chemical structure storage and retrieval systems [7]. Naturally, minimal cycles bases are of particular practical interest.

In this contribution we prove that outerplanar graphs have a unique minimal cycle basis. This result was motivated by the analysis of the structures of biopolymers. In addition we derive upper and lower bounds on the length of minimal cycle basis in 2-connected graphs.

Biopolymers, such as RNA, DNA, or proteins form well-defined three dimensional structures. These are of utmost importance for their biological function. The most salient features of these structures are captured by their *contact graph* representing the set E of all pairs of monomers V that are spatially adjacent. While this simplification of the 3D shape obviously neglects a wealth of structural details, it encapsulates the type of structural information that can be obtained by a variety of experimental and computational methods. Nucleic acids, both RNA and DNA, form a special type of contact structures known as *secondary structures*. These graphs are outer-planar and subcubic, i.e., the maximal vertex degree is 3.

A particular type of cycles, which is commonly termed *loops* in the RNA literature, plays an important role for RNA (and DNA) secondary structures: the energy of a secondary structure can be computed as the sum of energy contributions of the *loops*. These *loops* form the unique minimal cycle basis of the contact graph. Experimental energy parameters are available for the contribution of an individual *loop* as a function of its size, of the type of bonds that are contained in it, and on the monomers (nucleotides) that it is composed of [8]. Based on this energy model it is possible to compute the secondary structure with minimal energy given the sequence of nucleotides using a dynamic programming technique [17].

2. Preliminaries

In this contribution we consider only finite simple graphs $G(V, E)$ with vertex set V and edge set E, i.e., there are no loops or multiple edges. $G(V, E)$ is 2-connected if the deletion of a single vertex does not disconnect the graph.

Let $G_1(V_1, E_1)$ and $G_2(V_2, E_2)$ be two sub-graphs of a graph $G(V, E)$. We shall write $G_1 \setminus G_2$ for the subgraph of G induced by the *edge set* $E_1 \setminus E_2$.

The set $\mathcal{E}$ of all subsets of E forms an m-dimensional vector space over GF(2) with vector addition $X \oplus Y := (X \cup Y) \setminus (X \cap Y)$ and scalar multiplication $1 \cdot X = X$, $0 \cdot X = \emptyset$ for all $X, Y \in \mathcal{E}$. A *cycle* is a subgraph such that any vertex degree is even. We represent a cycle by its edge set C. Sometimes it will be convenient to regard C as a subgraph (V_C, C) of $G(V, E)$. The set $\mathcal{C}$ of all cycles forms a subspace of $(\mathcal{E}, \oplus, \cdot)$ which is called the *cycle space* of G. A basis $\mathcal{B}$ of the cycle space $\mathcal{C}$ is called a *cycle basis* of $G(V, E)$ [2]. The dimension of the cycle space is the *cyclomatic number* or *first Betti number* $\nu(G) = |E| - |V| + 1$.

It is obvious that the cycle space of graph is the direct sum of the cycle spaces of its 2-connected components. It will be sufficient therefore to consider only 2-connected graphs in this contribution.

A *elementary cycle* is a cycle C for which (V_C, C) is a connected minimal subgraph such that every vertex in V_C has degree 2. We say that a cycle basis is elementary if all cycles are elementary. A cycle C is a *chordless cycle* if (V_C, C) is an induced subgraph of $G(V, E)$, i.e., if there is no edge in $E \setminus C$ that is incident to two vertices of V_C. We shall say that a cycle basis is chordless if all its cycles are chordless.

The length $|C|$ of a cycle C is the number of its edges. The length $\ell(\mathcal{B})$ of a cycle basis $\mathcal{B}$ is sum of the lengths of its cycles: $\ell(\mathcal{B}) = \sum_{C \in \mathcal{B}} |C|$. A *minimal cycle basis* is a cycle basis with minimal length. Let $c(\mathcal{B})$ be the length of the longest cycle in the cycle basis $\mathcal{B}$. Chickering [3] showed that if $\ell(\mathcal{B})$ is minimal then $c(\mathcal{B})$ is minimal, i.e., a minimal cycle basis has a shortest longest cycle.

A cycle C is *relevant* [13] if it is contained in a minimal cycle basis. Vismara [15] proved the following

Proposition 1. *A cycle C is relevant if and only if it cannot be represented as a $\oplus$-sum of shorter cycles.*

An immediate consequence is

Corollary 2. *A relevant cycle is chordless. Hence a minimal cycle basis is chordless (and of course elementary).*

3. Fundamental Cycle Bases

In what follows let $G(V, E)$ be a 2-connected graph.

A collection of $\nu(G)$ cycles in G is called *fundamental* if there exists an ordering of these cycles such that [9, 18]

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

Of course such a collection is a cycle basis. Not all cycle bases are fundamental [9].

Lemma 3. *An elementary fundamental cycle basis can be ordered such that*

(i) C_1 is an elementary cycle and

(ii) $C_j \setminus (C_1 \cup \cdots \cup C_{j-1}) = P_j$ is a nonempty path for $2 \leq j \leq \nu(G)$.

Proof. Let $G_i = C_1 \cup \cdots \cup C_i$. Then $\nu(G_i) \geq \nu(G_{i-1}) + 1$ for $i \geq 2$ and consequently $\nu(G) = \nu(G_{\nu(G)}) \geq \nu(G_{\nu(G)-1}) + 1 \geq \cdots \geq \nu(G_1) + (\nu(G) - 1) = \nu(G)$. Therefore equality holds and we have $\nu(G_i) = i$, i.e. $\mathcal{B}_i = \{C_1, \ldots, C_i\}$ is a cycle basis for G_i.

Next notice that there exists an ordering for which (1) holds such that G_i is connected for all $i \geq 1$, i.e. $C_i \cap G_{i-1} \neq \emptyset$. Otherwise there exists a j such that $C \cap G_j = \emptyset$ for all $C \in \mathcal{B} \setminus \mathcal{B}_{j-1}$ for all orderings satisfying (1). But then $C_j \cup \cdots \cup C_{\nu(G)}$ has empty intersection with $G_{j-1} = C_1 \cup \cdots \cup C_{j-1}$, a contradiction, since $G = C_1 \cup \cdots \cup C_{\nu(G)}$ is 2-connected. G_i is connected since by assumption all C_j are elementary.

An immediate consequence is that $C_j \setminus G_{j-1}$ must be either a path as claimed, or an elementary cycle which has one vertex in common with G_{j-1}. Otherwise we would have $\nu(G_j) > j$. If $C_j \setminus G_{j-1}$ is a cycle, this one vertex must be a cut vertex of G_j. In this case, there must be a list of cycles $C_{k_1}, C_{k_2}, \ldots, C_{k_p}$ in $B \setminus B_j$ such that: C_{k_1} has edges in common with G_{j-1}, C_{k_q} has edges in common with $C_{k_{q+1}}$ for all $q \in [1, p-1]$, and C_{k_p} has edges in common with C_j. Then we can reorder the basis by exchanging C_j and C_k. □

A weaker result holds for non-fundamental cycle bases:

Lemma 4. *Any elementary non-fundamental cycle basis can be ordered such that $C_1 \cup \ldots \cup C_i$ is 2-connected for all $i \geq 1$.*

Proof. Analogously to the proof of lemma 3 there exists an ordering such that G_i is connected for all $i \geq 1$, i.e. $C_i \cap G_{i-1} \neq \emptyset$. Otherwise there exists a j such that $C \cap G_j = \emptyset$ for all $C \in \mathcal{B} \setminus \mathcal{B}_{j-1}$ for all orderings. But then $C_j \cup \cdots \cup C_{\nu(G)}$ has empty intersection with $G_{j-1} = C_1 \cup \cdots \cup C_{j-1}$, a contradiction, since $G = C_1 \cup \cdots \cup C_{\nu(G)}$ is 2-connected. G_i is connected since by assumption all C_j are elementary.

Similarly there exists an ordering such that G_i is 2-connected for all $i \geq 1$. Obviously $G_1 = C_j$ is 2-connected, since C_j is elementary. Assume G_{j-1} is 2-connected. If $C_j \cap G_{j-1}$ consists of at least two vertices, G_j is 2-connected. If C_j and G_{j-1} have only one vertex in common (there must be at least one such vertex), then there must be a cycle $C_k \in \mathcal{B} \setminus \mathcal{B}_j$ which has edges in common with G_{j-1} and with P_j. Otherwise G cannot be 2-connected. Then we can reorder the basis by exchanging C_j and C_k. □

If $\mathcal{B}$ is a non-fundamental cycle basis of G then there is subgraph G' with cycle basis $\mathcal{B}' \subseteq \mathcal{B}$ such that each edge of G' is contained in at least two cycles of $\mathcal{B}'$ [9, prop. 4.2]. Furthermore, the examples of non-fundamental bases in [9] are much longer than the minimal cycles bases. One might be tempted therefore to conjecture that every minimal cycle basis is fundamental. Although this statement is easily verified for planar graphs (see corollary 13), it is not true in general: Consider the complete graph K_9 with 9 vertices. It is straightforward (we used `Mathematica`) to check that the following 28 cycles are independent and thus are a basis of the cycle space, since $\nu(K_9) = 28$.

$$\begin{gathered}(1,2,3),(2,3,4),(1,3,5),(3,4,5),(2,4,5),(2,5,6),(1,5,6),\\(1,4,6),(3,4,6),(2,6,7),(3,6,7),(1,3,7),(1,4,7),(4,5,7),\\(2,7,8),(5,7,8),(1,5,8),(1,6,8),(4,6,8),(2,3,8),(3,8,9),\\(4,8,9),(1,4,9),(1,2,9),(2,5,9),(5,6,9),(6,7,9),(3,7,9)\end{gathered}$$

Here $(1,2,3)$ denotes the 3-cycle $\{(v_1,v_2),(v_2,v_3),(v_3,v_1)\}$. This basis is minimal, since every cycle has length 3. But it is non-fundamental, since every edge is covered at least two times.

The concept of *fundamental* has originally been introduced by Kirchhoff in 1847 [12] in the following way:

Suppose T is a spanning tree of G. Then for each edge $\alpha \notin T$ there is unique cycle in $T \cup \{\alpha\}$ which is called a *fundamental cycle*. The set of fundamental cycles belonging to a given spanning tree form a basis of the cycle subspace which is called the *fundamental basis w.r.t. T*. For details see [14]. It is obvious that a fundamental basis w.r.t. a spanning tree is a special case of the fundamental collections defined at the beginning of this section, see also [9].

4. Outerplanar Graphs

A graph $G(V, E)$ is *outer-planar* if it can be embedded in the plane such that all vertices lie on the boundary of its exterior region. Given such an embedding we will refer to the set of edges on the boundary to the exterior region as the *boundary* B of G. A graph is outerplanar if and only if it does not contain a K_4 or $K_{3,2}$ minor [1]. An algebraic characterization in terms of a spectral invariant is discussed in [4].

Lemma 5. *An outerplanar graph $G(V, E)$ is Hamiltonian if and only if it is 2-connected.*

Proof. A Hamiltonian graph is always 2-connected. Suppose G is outerplanar, 2-connected, but not Hamiltonian. 2-connectedness implies that there is no cut-vertex. Thus the boundary B of G is a closed path containing an edge at most once. A vertex x that is incident with more than 2 edges of B must be a cut-vertex of G since it partitions B into (at least) two edge-disjoint closed paths B' and B''. Let V' and V'' the vertices incident with the edges in B' and B'', respectively. Outerplanarity implies that there are no edges connecting a vertex $y \in V' \setminus \{x\}$ with a vertex $z \in V'' \setminus \{x\}$. Thus x is a cut vertex, contradicting 2-connectedness. □

Lemma 6. *A 2-connected outerplanar graph $G(V, E)$ contains a unique Hamiltonian cycle H.*

Proof. If G is a cycle graph, there is nothing to show. Otherwise denote by H the Hamiltonian cycle forming the boundary of G and consider an arbitrary edge $\alpha \notin H$. By construction G is embedded in the plane such that $\alpha = (p, q)$ divides G into two subgraphs G_1 and G_2 with vertex sets V_1 and V_2 satisfying $|V_i| \geq 3$ and $V_1 \cap V_2 = \{p, q\}$. Now consider two vertices $x \in V_1 \setminus \{p, q\}$ and $y \in V_2 \setminus \{p, q\}$. Since G is outerplanar, each path from x to y passes through p or q. Each elementary cycle containing both x and y therefore consists of two disjoint paths, one of which passes only through p while the other one passes only through q. Thus the edge (p, q) cannot be part of any elementary cycle containing both x and y, and hence G contains no Hamiltonian cycle different from H. □

As a consequence there is a unique partition of the edge set E of an outerplanar graph G into the Hamiltonian cycle H and the *set of chords* $K = E \setminus H$. It will be convenient to label the vertices such that the edges in H are $(i, i+1)$ for $1 \leq i \leq n-1$ and $(1, n)$. Without loosing generality we may assume that n is a vertex of degree 2.

It will be useful to introduce the following partial order on K:

$$\alpha = (i_\alpha, j_\alpha) \prec \beta = (i_\beta, j_\beta) \quad \text{if and only if} \quad i_\beta \leq i_\alpha < j_\alpha \leq j_\beta \text{ and } \alpha \neq \beta\,. \tag{2}$$

We say that α is *interior to* β. If there is no $\gamma \in K$ such that $\alpha \prec \gamma \prec \beta$ we say that α is immediately interior to β, $\alpha \prec\!\!\prec \beta$. For each alpha in K we set $Y_\alpha = \{\beta \in K | \beta \prec\!\!\prec \alpha\}$. Y_* denotes the set of $\prec$-maximal elements in K, i.e., the set of contacts that are not interior to any other contact. Yan's [19] *bamboo shoot graphs* are exactly those outer-planar graphs for which $(K, \prec)$ is an ordered set.

Nucleic acids, both RNA and DNA, form a special type of contact structure known as *secondary structure*. A graph $G(V,E)$, with $V=\{1,\ldots,n\}$, is a secondary structure if it satisfies

(i) The so-called backbone $T=\{(i,i+1)|1\leq i<n\}$ is a subset of E.
(ii) For each $i\in V$ there is at most one contact $\alpha\in E\setminus T$ incident with i.
(iii) If $(i,j),(k,l)\in E\setminus T$ and $i<k<j$ then $i<l<j$.

The contacts in nucleic acids are usually called *base pairs*. Note that the backbone T is a spanning tree of G.

Lemma 7. *A secondary structure graph is connected, outerplanar, and subcubic.*

Proof. By properties (i) and (ii) it is clear that a secondary structure graph is subcubic. Property (iii) implies that, when the vertices are arranged along a circle then one may draw the chord $E\setminus T$ in the interior of this circle without intersection, i.e., G is outerplanar. (This is a common representation for drawing RNA secondary structures.) $\square$

The converse is not true since outerplanar subcubic graphs do not necessarily have unbranched spanning trees T.

5. Minimal Cycle Bases of Outerplanar Graphs

Let (G,V) be a 2-connected outerplanar graph. The set $T=H\setminus\{(1,n)\}$ is a spanning tree of $G(V,E)$. The fundamental basis $\mathcal{F}$ of the cycle space w.r.t. T (in the sense of Kirchhoff) therefore consists of the uniquely determined cycles F_α in $T\cup\{\alpha\}$, $\alpha\in K$ and the Hamiltonian cycle $H=T\cup\{(1,n)\}$. We define

$$C_\alpha=\left[\bigoplus_{\beta\in Y_\alpha}F_\beta\right]\oplus F_\alpha \qquad\text{and}\qquad C_*=\left[\bigoplus_{\beta\in Y_*}F_\beta\right]\oplus H\,. \tag{3}$$

Furthermore we set $\mathcal{M}=\{C_\alpha|\alpha\in K\}\cup\{C_*\}$.

Theorem 8. *Let $G(V,E)$ be a 2-connected and outerplanar graph. Then $\mathcal{M}$ is the unique minimal cycle basis of G.*

Proof. Consider an edge $\alpha\in K$ such that $Y_\alpha=\emptyset$, that is, a minimal element of the poset $(K,\prec)$. We observe that $F_\alpha=C_\alpha$ in this case.

Let G' be the graph obtained from G by deleting the edges $F_\alpha\setminus\{\alpha\}$ and all vertices that are isolated as a consequence. It is clear that G' is again a 2-connected outerplanar graph: Its boundary is the Hamiltonian cycle $H'=H\oplus C_\alpha$. The set of chords of G' is $K'=K\setminus\{\alpha\}$. The fundamental basis $\mathcal{F}'$ of G' w.r.t. $T'=H'\setminus\{(1,n)\}$ consists of H' and the cycles F'_α, $\alpha\in K'$, which are obtained by the rule $F'_\beta=F_\beta\oplus C_\alpha$ if $\alpha\prec\beta$ and $F'_\beta=F_\beta$ if $\alpha\not\prec\beta$. Furthermore, we have $Y'_\beta=Y_\beta\setminus\alpha$ and $F'_\beta=C_\beta$ if and only if $Y'_\beta=\emptyset$.

Consider an arbitrary cycle basis $\mathcal{B}$ of G. We can construct a cycle basis $\tilde{\mathcal{B}}$ of G from $\mathcal{B}$ that consists of C_α and a cycle basis $\mathcal{B}'$ of G' by the following procedure: For each $Z\in\mathcal{B}$ we define $Z'=Z$ if $Z\cap C_\alpha=\emptyset$ or if $Z\cap C_\alpha=\{\alpha\}$. In the remaining cases, where $Z \cap C_\alpha \neq \{\alpha\}$ or $\emptyset$ because $Y_\alpha \neq \emptyset$, we set $Z' = Z \oplus C_\alpha = (Z \setminus C_\alpha) \cup \{\alpha\}$. We have to distinguish three cases:

(i) $C_\alpha \in \mathcal{B}$ and $Z = Z'$ for all other cycles. Then $\mathcal{B} = \tilde{\mathcal{B}} = \mathcal{B}' \cup \{C_\alpha\}$ and $\ell(\mathcal{B}) = \ell(\mathcal{B}') + |C_\alpha|$.

(ii) $C_\alpha \in \mathcal{B}$, but there is at least one cycle $Z' \in \mathcal{B}$ satisfying $Z' \neq Z$. The length of this cycle is $|Z'| = |Z| - |C_\alpha| + 2 < |Z|$, i.e., $\ell(\tilde{\mathcal{B}}) < \ell(\mathcal{B})$.

(iii) $C_\alpha \notin \mathcal{B}$. Then there is at least one cycle $Z' \neq Z$ and all Z' are non-empty. Since C_α is independent of all Z' there must be at least on dependent cycle in the set $\{Z'|Z \in \mathcal{B}\}$, which must be removed in order to obtain the basis $\mathcal{B}'$. The length of this cycle is of course at least 3. Thus

$$\ell(\tilde{\mathcal{B}}) \leq \ell(\mathcal{B}) + |C_\alpha| - |Z| + |Z'| - 3 = \ell(\mathcal{B}) + |C_\alpha| - |C_\alpha| + 2 - 3 = \ell(\mathcal{B}) - 1\,,$$

and $\tilde{\mathcal{B}}$ is strictly shorter than $\mathcal{B}$ in this case, too.

Thus, if $\mathcal{B}$ is a minimal cycle basis of G, then cases (ii) and (iii) cannot occur, i.e., a minimal cycle basis of G consists of C_α and a minimal cycle basis $\mathcal{B}'$ of G'.

Repeating this argument $|K|$ times shows that each cycle C_β, $\beta \in K$, must be contained in any minimal cycle basis of G. The remainder G^* of G after all cycles C_β, $\beta \in K$ are removed by the above procedure is composed of Y_* and those edges of H that are not contained in any of the cycles C_α. The edge set of G^* is the chordless cycle C_*. Thus $\{C_*\} \cup \{C_\alpha | \alpha \in K\} = \mathcal{M}$ is therefore the only minimal cycle basis of Γ. □

Let $G(V,E)$ be a planar graph, and let $\{\hat{Q}_j\}$ be the collection of faces in a given embedding in the plane. Each face $\hat{Q}_j$ uniquely defines the cycle Q_j which forms its boundary. The collection of cycles $\{Q_j\}$, $j = 1, \ldots, \nu(G)$, is a cycle basis of G. Any cycle basis obtained in this way is called a *planar cycle basis*.

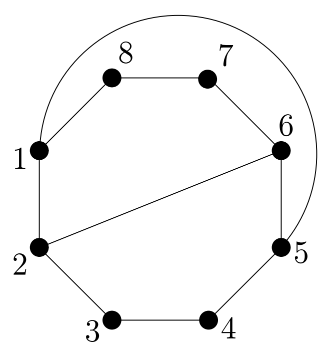


FIGURE 1. Hamiltonian planar graph with a non-planar minimal cycle basis. It is easy to verify that this graph has no planar embedding with the face $Q = (1,2,6,5)$. A minimal cycle basis contains Q and two of the cycles $(2,3,4,5,6)$, $(1,2,6,7,8)$, $(1,2,3,4,5)$, and $(1,5,6,7,8)$. Hence $\ell(\mathcal{M}) = 14$ while the planar bases have $\ell(\mathcal{M}) = 15$.

It is natural to ask whether every planar graph has a minimal cycle basis that is also planar. The answer to the question is negative in general, as figure 1 shows.

Corollary 9. *$\mathcal{M}$ is planar cycle basis with length $\ell(\mathcal{M}) = 2|E| - |V|$.*

Proof. The cycle basis $\mathcal{M}$ is the planar basis obtained by embedding G in such a way in the plane that the Hamiltonian cycle H becomes the outer boundary. By construction we have $\ell(\mathcal{M}) = |H| + 2|K|$. Using $|H| = |V|$ and $|K| = |E| - |V|$ leads to the desired result. □

We now turn to an algorithm for finding the unique minimal cycle basis of an outerplanar graph. Since our investigation is motivated by RNA secondary structures, we assume that the backbone of the outerplanar graph, i.e. the Hamiltonian-cycle is already given.

The basic idea of algorithm 1 is to find the minimal cycles C_α described in the proof of theorem 8. When such a cycle is found, it is added to the cycle basis and "chopped off" the graph. Step 1 generates an ordered list of V (along the Hamiltonian cycle). It is best implemented as linked list of pointers to the vertices. Steps 8 and 9 push every contact (and $(1, n)$) that have not already been processed on the stack. Steps 3 and 4 pop all contacts incident to the current vertex i from the stack. By corollary 9 the ordering of the edges used in step 8 ensures that the cycle generated in steps 5 and 6 are chordless and all vertices except i and k_j in P have degree 2. Hence they are part of $\mathcal{M}$. In step step 7 they are "chops off" taking advantage of the fact that V is a linked list. It is easy to see that this algorithm is of order $O(|V|)$. We illustrate the algorithm in Figure 2.

It is interesting to note that the fundamental basis $\mathcal{F}$ can be easily expressed in terms of the minimal cycle basis $\mathcal{M}$:

$$\begin{aligned} F_\alpha &= C_\alpha \oplus \left[\bigoplus\nolimits_{\beta\in Y_\alpha} F_\beta\right] = C_\alpha \oplus \left[\bigoplus\nolimits_{\beta\in Y_\alpha} C_\beta \oplus \left[\bigoplus\nolimits_{\gamma\in Y_\beta} F_\gamma\right]\right] \\ &= C_\alpha \oplus \left[\bigoplus\nolimits_{\beta\in Y_\alpha} C_\beta \oplus \left[\bigoplus\nolimits_{\gamma\in Y_\beta} C_\gamma \oplus \left[\bigoplus\nolimits_{\delta\in Y_\gamma} F_\delta\right]\right]\right] = \dots \end{aligned}$$

The expansion eventually stops if $Y_\psi = \emptyset$ and hence $F_\psi = C_\psi$. Clearly, the nested sums contain each bond in $W_\alpha = \{\beta \in K | \beta \prec \alpha\}$, the set of contacts interior to α, and α itself exactly once. Therefore we have

$$F_\alpha = \bigoplus_{\beta\in W_\alpha\cup\{\alpha\}} C_\beta\,.$$

Analogously one finds

$$H = \left[\bigoplus_{\beta\in Y_*} F_\beta\right] \oplus C_* = \left[\bigoplus_{\beta\in K} C_\beta\right] \oplus C_*\,.$$

6. Upper Bounds on $\min \ell(\mathcal{B})$

In [10, theorem 6] an upper bound for the length of a minimal cycle basis $\mathcal{M}$ of an arbitrary graph $G(V, E)$ is given:

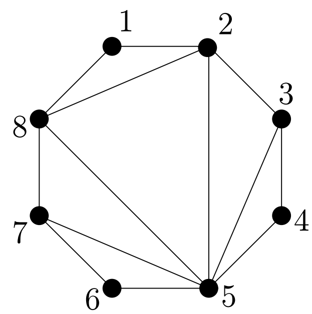
$$\ell(\mathcal{M}) \le 3(|V| - 1)(|V| - 2)/2. \tag{4}$$

While this bound is sharp for complete graphs [5], it can be improved substantially for planar graphs.

algorithm 1 find minimal cycle basis of outerplanar graphs

Input: adjacency matrix, Hamiltonian cycle $\{1, \dots, n\}$

1: $V \leftarrow (1, \dots, n)$.
2: **for all** vertices i, $1 \le i \le n$ **do**
3:     **while** there is an edge (i, k_j) at the top of the stack **do**
4:         pop edge (i, k_j) from stack.
5:         $P \leftarrow$ path from k_j to i in V.
6:         add cycle $P \cup \{(i, k_j)\}$ to cycle basis.
7:         remove vertices in $P \setminus \{i, k_j\}$ from V.
8:     **for all** edges (i, k_j), $n \ge k_1 > k_2 > \dots > i + 1$ **do**
9:         push edge (i, k_j) on stack.



i	step	action	bottom stack top	V
	0	beginn	empty	empty
	1	make list	empty	$(1,2,3,4,5,6,7,8)$
1	9	push $(1,8)$	$(1,8)$	$(1,2,3,4,5,6,7,8)$
2	9	push $(2,8)$	$(1,8),(2,8)$	$(1,2,3,4,5,6,7,8)$
2	9	push $(2,5)$	$(1,8),(2,8),(2,5)$	$(1,2,3,4,5,6,7,8)$
3	9	push $(3,5)$	$(1,8),(2,8),(2,5),(3,5)$	$(1,2,3,4,5,6,7,8)$
4		none	$(1,8),(2,8),(2,5),(3,5)$	$(1,2,3,4,5,6,7,8)$
5	4–7	create $(3,4,5,3)$	$(1,8),(2,8),(2,5)$	$(1,2,3,5,6,7,8)$
5	4–7	create $(2,3,5,2)$	$(1,8),(2,8)$	$(1,2,5,6,7,8)$
5	9	push $(5,8)$	$(1,8),(2,8),(5,8)$	$(1,2,5,6,7,8)$
5	9	push $(5,7)$	$(1,8),(2,8),(5,8),(5,7)$	$(1,2,5,6,7,8)$
6		none	$(1,8),(2,8),(5,8),(5,7)$	$(1,2,5,6,7,8)$
7	4–7	create $(5,6,7,5)$	$(1,8),(2,8),(5,8)$	$(1,2,5,7,8)$
8	4–7	create $(5,7,8,5)$	$(1,8),(2,8)$	$(1,2,5,8)$
8	4–7	create $(2,5,8,2)$	$(1,8)$	$(1,2,8)$
8	4–7	create $(1,2,8,1)$	empty	$(1,8)$
		stop		

FIGURE 2. Example for algorithm 1

First we need the following simple observations:

Proposition 10. *Let $G(E,V)$ be a 2-connected graph. Then*

$$|E| \leq 3\,|V| - 6 \qquad \textit{if } G \textit{ is planar.} \tag{5}$$

$$|E| \leq 2\,|V| - 3 \qquad \textit{if } G \textit{ is outerplanar.} \tag{6}$$

These bounds are sharp for all $|V| \geq 3$.

The result on planar graphs is an immediate corollary of Euler's formula for polyhedra. The upper bound on outerplanar graphs follows from a theorem by G.A. Dirac [6] stating that for any graph not containing K_4 as a minor we have $|E| \leq 2|V| - 3$.

A bamboo-shoot graph [19] consisting of n triangles has $n' = n + 2$ vertices and $2n + 1 = 2n' - 3$ edges. Consider the graph G_n recursively obtained by adding a vertex n which is connected to the three vertices labeled $n-1$, 1, and 2 of G_{n-1}. We set $G_3 = K_3$, the cycle of length 3. It is obvious that these graphs are all planar, and G_n has 3 edges and 1 vertex more than G_{n-1}. Thus G_n has n vertices and $3(n-3) + 3 = 3n - 6$ edges.

We can translate the above result into upper bounds for the lengths of a minimal cycle bases that depend only on the number of vertices:

Theorem 11. *Let $G(E,V)$ be a 2-connected planar graph with a minimal cycle basis $\mathcal{M}$. Then*

$$\ell(\mathcal{M}) \leq 6\,|V| - 15 \qquad \textit{if } G \textit{ is planar.} \tag{7}$$

$$\ell(\mathcal{M}) \leq 3\,|V| - 6 \qquad \textit{if } G \textit{ is outerplanar.} \tag{8}$$

Proof. Analogously to the proof of proposition 10 we find for the planar case $\ell(\mathcal{M}) \leq 2\,|E| - 3 \leq 2\,(3\,|V| - 6) - 3 = 6\,|V| - 15$ by (5) as claimed. Similarly for the outer planar case: $\ell(\mathcal{M}) = 2\,|E| - |V| \leq 2(2\,|V| - 3) - |V| = 3\,|V| - 6$ which is (8). $\square$

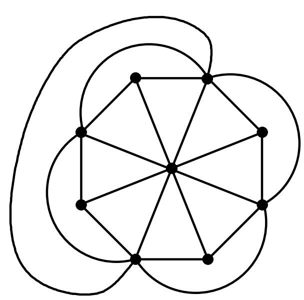


FIGURE 3. A Hamiltonian planar graph for which inequality (7) is sharp.

It is not possible to improve the bound (7) for planar Hamiltonian graphs, see the example in figure 3. Similar examples for $|V| = 2^m + 1$ can be constructed by the following recipe:

1. Start with a 2^m-gon.
2. Insert the center as additional vertex c and add edges (c, i) for $1 \leq i \leq 2^m$.

3. Insert edges $(1, 3)$, $(3, 5)$, $(5, 7)$, ...
4. Insert edges $(1, 5)$, $(5, 9)$, $(9, 13)$, ... , and so on.

Lemma 12. *Let $G(V, E)$ be a 2-connected planar graph. Then $\ell(\mathcal{M}) \leq 2|E| - g(G)$, where $g(G)$ denotes the girth of G.*

Proof. A planar cycle basis contains each "interior" edge twice, while the edges of the outer boundary appear only once. The length of the outer boundary is at least $g(G)$. □

As an immediate consequence we have

Corollary 13. *Every minimal cycle basis of a planar graph is fundamental.*

Proof. Suppose $\mathcal{B}$ is not fundamental. By [9, prop. 4.2] we can assume that here is a subset $\mathcal{B}' \subseteq \mathcal{B}$ such that (i) $\mathcal{B}'$ is a minimal cycle basis for G', the subgraph of G induced by $\mathcal{B}'$, and (ii) $\mathcal{B}'$ covers every edge of G' two or more times. We can assume that G' is 2-connected; otherwise consider G'' with minimal cycle basis $\mathcal{B}''$ where G'' is a 2-connected subgraph of G' with minimal cycle basis $\mathcal{B}'' \subseteq \mathcal{B}'$ is a subset. (Note that $\mathcal{B}''$ necessarily covers every edge of G'' two or more times.) Thus $\ell(\mathcal{B}') \geq 2|E'|$. But since every planar cycle basis has length $\ell(\mathcal{B}) < 2|E|$ by lemma 12, $\mathcal{B}'$ cannot be a minimal cycle basis of G', and the proposition follows. □

7. Lower Bounds on $\ell(\mathcal{B})$

Theorem 14. *Let $G(V, E)$ be a 2-connected graph and $\mathcal{B}$ a cycle basis of G. Then*

$$\ell(\mathcal{B}) \geq 2|E| - |V| \tag{9}$$

If equality holds, then the cycle basis $\mathcal{B}$ is minimal. Equality holds for a basis if and only if for every vertex $v \in V$ the number of cycles $c \in \mathcal{B}$ through v is $d_v - 1$, where d_v denotes the degree of v.

Proof. Let S_v denote the graph induced by all edges incident to v (S_v is a star of diameter 2 with d_v edges). The edge set E_v of S_v with the addition $\oplus$ forms a d_v-dimensional vector space. Let $\mathcal{T}_v$ denote its sub space where each element consists of an even number of edges. As can easily be verified, $\dim(\mathcal{T}_v) = d_v - 1$.

Let $\mathcal{C}_v$ denote the vector space $(\{C \cap S_v \colon C \in \mathcal{C}\}, \oplus, \cdot)$. It is obvious that $\mathcal{C}_v \subseteq \mathcal{T}_v$. Moreover for all $C \in \mathcal{C}_v$ that consists of exactly 2 edges, there exists a $C' \in \mathcal{C}$ with $C = C' \cap S_v$ and thus $\mathcal{C}_v \supseteq \mathcal{T}_v$. Otherwise every cycle C in G has vertex v as double point and hence v was a cut vertex of G, a contradiction to G being 2-connected.

Therefore $\mathcal{C}_v = \mathcal{T}_v$ is spanned by $\mathcal{B}_v = \{C \cap S_v \colon C \in \mathcal{B}, C \cap S_v \neq \emptyset\}$. Consequently $|\mathcal{B}_v| \geq \dim(\mathcal{T}_v) = d_v - 1$ and $\ell(\mathcal{B}_v) \geq 2\,d_v - 2$, since every $C \in \mathcal{B}$ contains at least two edges. Notice that $|E| = \frac{1}{2}\sum_{v \in V} d_v$ (every edge is incident to 2 vertices) and $\ell(\mathcal{B}) = \frac{1}{2}\sum_{v \in V} \ell(\mathcal{B}_v)$ (every edge is contained in the stars S_{v_i} centered at two vertices v_1 and v_2). Hence we find

$$\ell(\mathcal{B}) = \frac{1}{2}\sum_{v \in V} \ell(\mathcal{B}_v) \geq \sum_{v \in V} (d_v - 1) = 2\,|E| - |V|$$

i.e., inequality (9). Equality holds if and only if $\mathcal{B}_v$ is a basis of $\mathcal{T}_v$. Thus the statement follows. □

Theorem 15. *Let $G(V,E)$ be a 2-connected graph with a elementary cycle basis $\mathcal{B}$. Then $\ell(\mathcal{B}) = 2|E| - |V|$ if and only if $\mathcal{B}$ is a fundamental cycle basis such that $|C_i \cap (C_1 \cup \cdots \cup C_{i-1})| = 1$ for all $2 \le i \le \nu(G)$ (i.e. consists of exactly one edge). In this case $\mathcal{B}$ is a minimal cycle basis.*

Proof. By lemmata 3 and 4 we can order the cycle basis $\mathcal{B}$ such that $C_1 \cup \cdots \cup C_i$ is 2-connected. Thus $C_i \cap (C_1 \cup \cdots \cup C_{i-1})$ consists of at least one edge.

Let $\phi(G) = 2|E| - |V|$. Let E_i denote the edge set of G_i and let V_i and $V(C_j)$ denote the vertex sets of G_i and C_j, respectively. Then we find

$$\begin{aligned}\phi(G_i) &= 2|E_i| - |V_i| \\ &= 2(|E_{i-1}| + |C_i| - |E_{i-1} \cap C_i|) - (|V_{i-1}| + |V(C_i)| - |V_{i-1} \cap V(C_i)|) \\ &= (2|E_{i-1}| - |V_{i-1}|) + (2|C_i| - |V(C_i)|) - (2|E_{i-1} \cap C_i| - |V_{i-1} \cap V(C_i)|) \\ &= \phi(G_{i-1}) + |C_i| - (|E_{i-1} \cap C_i| - \eta_i)\end{aligned}$$

since $|C_i| = |V(C_i)|$ and $|V_{i-1} \cap V(C_i)| = |E_{i-1} \cap C_i| + \eta_i$, where η_i denotes the number of connected components (i.e. paths) in $C_i \setminus E_{i-1}$. The latter equality holds because C_i is elementary. Notice that the number of components in $C_i \cap E_{i-1} = \eta_i$ if $\eta_i \ge 1$ and 1 otherwise, since C_i is elementary. Thus $|E_{i-1} \cap C_i| \ge \eta_i$ and

$$\phi(G_i) \le \phi(G_{i-1}) + |C_i|$$

Equality holds if and only if $|E_{i-1} \cap C_i| = \eta_i$.

Let $\mathcal{B}_i = \{C_1, \ldots, C_i\}$. We have $\ell(\mathcal{B}_1) = |C_1| = 2|E_1| - |V_1| = \phi(G_1)$ since C_1 is an elementary cycle. By induction we then find $\ell(\mathcal{B}_i) \ge \phi(G_i) = 2|E_i| - |V_i|$:

$$\ell(\mathcal{B}_i) = \ell(\mathcal{B}_{i-1}) + |C_i| \ge \phi(G_{i-1}) + |C_i| \ge \phi(G_i)$$

Equality holds if and only if all $|E_{i-1} \cap C_i| = \eta_i$. If $\mathcal{B}$ is fundamental then $\eta_i = 1$ for all i by lemma 3 and $E_{i-1} \cap C_i$ is a single edge for all i as claimed. If $\mathcal{B}$ is not fundamental, than we always have an j such that $C_j \subseteq E_{j-1}$ and thus $|E_{j-1} \cap C_j| = |C_j| > 0 = \eta_j$. Thus $\ell(\mathcal{B}_j) > \phi(G_j)$.

Moreover, if $\ell(\mathcal{B}) = 2|E| - |V|$ then, by theorem 14, $\mathcal{B}$ is a minimal cycle basis. □

In the following we derive some weaker conditions for which (9) is sharp.

Lemma 16. *Let $G(V,E)$ be a 2-connected graph. If $\ell(\mathcal{B}) = 2|E| - |V|$ for a cycle basis $\mathcal{B}$, then G is planar.*

Proof. By theorems 15 equality in equation (9) implies that $\mathcal{B}$ is minimal, elementary and fundamental. Thus there exists an ordering of $\mathcal{B}$ as described in lemma 3. By theorem 15 every cycle C_i has exactly one edge in common with $C_1 \cup \cdots \cup C_{i-1}$. Thus by induction we can add the "ear" P_i from C_i into the planar drawing of $C_1 \cup \cdots \cup C_{i-1}$, for all $i \ge 2$. □

Lemma 17. *Let $G(V,E)$ be a Hamiltonian graph. Then there exists a cycle basis $\mathcal{B}$ for which $\ell(\mathcal{B}) = 2|E| - |V|$ holds if and only if G is outerplanar.*

Proof. By corollary 9 a minimal cycle basis of an outerplanar graph has length $2|E| - |V|$.

If equality holds in equ.(9) then G is planar (lemma 16) and $\mathcal{B}$ is fundamental (theorem 15). Moreover $\mathcal{B}$ can be ordered such that every cycle C_i has exactly one edge in common with $C_1 \cup C_{i-1}$. Obviously $H_1 = C_1$ is Hamiltonian cycle in C_1. Then by induction $H_{i+1} = H_i \oplus C_{i+1}$ is a Hamiltonian cycle in $C_1 \cup C_{i+1}$. Furthermore we can draw the "ears" P_{i+1} of the cycles C_i (lemma 3) into the outside of $C_1 \cup C_i$. Thus H_i is the boundary to the exterior region of G_i and the proposition follows by induction. $\square$

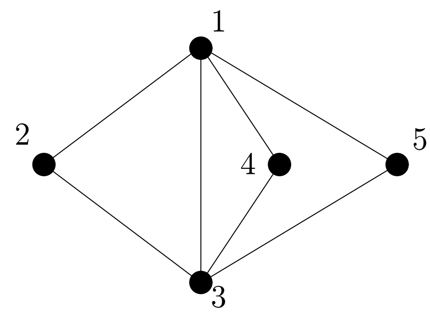


FIGURE 4. Non-Hamiltonian planar graph for which equality holds in lemma 16.

The condition "Hamiltonian" in lemma 17 cannot be relaxed. An example of a planar (non-Hamiltonian) graph with $\ell(\mathcal{B}) = 2\,|E| - |V|$ is shown in figure 4.

References

[1] G. Chartrand and F. Harary. Planar permutation graphs. *Ann. Inst. Henri Poincarè B*, 3:433–438, 1967.

[2] W.-K. Chen. On vector spaces associated with a graph. *SIAM J. Appl. Math.*, 20:525–529, 1971.

[3] D. M. Chickering, D. Geiger, and D. Heckerman. On finding a cycle basis of with a shortest maximal cycle. *Inform. Processing Let.*, 54:55–58, 1994.

[4] Y. Colin de Verdière. Sur un novel invariant des graphes et un critère de planarité. *J. Comb. Theory B*, 50:11–21, 1990.

[5] N. Deo, G. M. Prabhu, and M. S. Krishnamoorty. Algorithms for generating fundamental cycles in a graph. *ACM Trans. Math. Software*, 8:26–42, 1982.

[6] G. A. Dirac. In abstrakten Graphen vorhandene vollständige 4-Graphen und ihre Unterteilungen. *Math. Nachr.*, 22:61–85, 1960.

[7] G. M. Downs, V. J. Gillet, J. D. Holliday, and M. F. Lynch. Review of ring perception algorithms for chemical graphs. *J. Chem. Inf. Comput. Sci.*, 29:172–187, 1989.

[8] S. M. Freier, R. Kierzek, J. A. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson, and D. H. Turner. Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci. (USA)*, 83:9373–9377, 1986.

[9] D. Hartvigsen and E. Zemel. Is every cycle basis fundamental? *J. Graph Theory*, 13:117–137, 1989.

[10] J. D. Horton. A polynomial-time algorithm to find the shortest cycle basis of a graph. *SIAM J. Comput.*, 16:359–366, 1987.

[11] A. Kaveh. *Structural Mechanics: Graph and Matrix Methods.* Research Studies Press, Exeter, UK, 1992.

[12] G. Kirchhoff. Über die Auflösung der Gleichungen, auf welche man bei der Untersuchungen der linearen Verteilung galvanischer Ströme geführt wird. *Poggendorf Ann. Phys. Chem.*, 72:497–508, 1847.

[13] M. Plotkin. Mathematical basis of ring-finding algorithms in CIDS. *J. Chem. Doc.*, 11:60–63, 1971.

[14] K. Thulasiraman and M. N. S. Swamy. *Graphs: Theory and Algorithms*. J. Wiley & Sons, New York, 1992.

[15] P. Vismara. Union of all the minimum cycle bases of a graph. *Electronic J. Comb.*, 4:#R9 (15 pages), 1997.

[16] H.-J. Voss. *Cycles and Bridges in Graphs*. Kluwer, Dordrecht, 1991.

[17] M. S. Waterman. Secondary structure of single-stranded nucleic acids. *Adv. Math. Suppl. Studies*, 1:167–212, 1978.

[18] H. Whitney. On abstract properties of linear dependence. *Am. J. Math.*, 57:509–533, 1935.

[19] L. Yan. A family of special outerplanar graphs with only one triangle satisfying the cycle basis interpolation property. *Discr. Math.*, 143:293–297, 1995.