

Counterexamples in Chemical Ring Perception[†]Franziska Berger,^{||} Christoph Flamm,[‡] Petra M. Gleiss,[‡] Josef Leydold,[⊥] and Peter F. Stadler^{*,#,\$}

Zentrum Mathematik, TU München, Boltzmannstrasse 3, D-85747 Germany, Institut für Theoretische Chemie und Molekulare Strukturbiologie, Universität Wien, Währingerstrasse 17, A-1090 Wien, Austria, Department for Applied Statistics and Data Processing, University of Economics and Business Administration, Augasse 2-6, A-1090 Wien, Austria, Bioinformatik, Institut für Informatik, Universität Leipzig, Kreuzstrasse 7b, D-04103 Leipzig, Germany, and The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501

Received June 11, 2003

Ring information is a large part of the structural topology used to identify and characterize molecular structures. It is hence of crucial importance to obtain this information for a variety of tasks in computational chemistry. Many different approaches for “ring perception”, i.e., the extraction of cycles from a molecular graph, have been described. The chemistry literature on this topic, however, reports a surprisingly large number of incorrect statements about the properties of chemically relevant ring sets and, in particular, about the mutual relationships of different sets of cycles in a graph. In part these problems seem to have arisen from a sometimes rather idiosyncratic terminology for notions that are fairly standard in graph theory. In this contribution we translate the definitions of concepts such as the Smallest Set of Smallest Rings, Essential Set of Essential Rings, Extended Set of Smallest Rings, Set of Smallest Cycles at Edges, Set of Elementary Rings, \mathcal{H} -rings, and β -rings into a more widely used mathematical language. We then outline the basic properties of different cycle sets and provide numerous counterexamples to incorrect claims in the published literature. These counterexamples may have a serious practical impact because at least some of them are molecular graphs of well-known molecules. As a consequence, we propose a catalog of desirable properties for chemically useful sets of rings.

1. INTRODUCTION

A critical step for many computer programs in chemistry is the search for molecular structures within a large database of chemical compounds. For this purpose, the molecular structure is conveniently represented as a graph. During the last years, the size of many chemical databases has increased dramatically. Speed, efficiency, and accuracy of the retrieval algorithms are thus essential. Retrieval, however, requires the solution of the *graph isomorphism problem*, that is, to check whether a bijective mapping exists between the vertices of an input graph and those of a graph in the database such that the edge adjacency relations are preserved. Similarly, the identification of a given substructure in a molecular graph requires the solution of the *subgraph isomorphism problem*. Since it is not known whether the graph isomorphism problem can be solved efficiently,³⁶ and since the subgraph isomorphism problem is hard,²⁵ it is important for practical purposes to have efficient methods that exclude a large fraction of graphs in a database quickly. The more precisely this preprocessing step works, the more efficiently a database search can be effected, by solving the (sub)graph isomorphism problem only for a few remaining candidate graphs. Methods to exclude candidates for the

graph isomorphism problem include fragment-based search and the comparison of graph descriptors.⁶⁴

Ring information is a large part of the structural topology used to identify and characterize molecular structures. Cycles, on the other hand, belong to the most important objects in graph theory.⁵⁷ It is not surprising therefore that a wealth of molecular descriptors has been developed that are based on various sets of cycles. Examples are the set Ω of all rings, the Smallest Set of Smallest Rings (SSSR), the Essential Set of Essential Rings (ESER), the Extended Set of Smallest Rings (ESSR), the \mathcal{H} -rings, the β -rings, and many others. For a review see ref 16. The different cycle sets arose from the attempt to ensure that all “chemically meaningful rings” are included. Of course, what is “chemically meaningful” depends on the application. A certain set of long cycles, for example, may be an important substructure for synthesis planning, while it is irrelevant for a search in patent databases.

Historically, the chemistry literature has to a certain extent developed an idiosyncratic terminology for notions that are fairly standard in graph theory. This has unfortunately lead to some confusion in the literature about the precise definitions of some of the widely used cycle sets. Presumably as a consequence, the published literature contains a surprisingly large number of incorrect statements about the properties of chemically relevant ring sets and, in particular, regarding the mutual relationships of different cycle sets.

This contribution is an attempt to compile the definitions of the chemically most important cycle sets in a more standard graph-theoretical language, to discuss their basic

* Corresponding author e-mail: peter.stadler@bioinf.uni-leipzig.de.

[†] Dedicated to George W. A. Milne, a former long-term Editor-in-Chief of *JCICS*.

[‡] Universität Wien.

[§] The Santa Fe Institute.

^{||} TU München.

[⊥] University of Economics and Business Administration Wien.

[#] Universität Leipzig.

properties, and, in particular, to provide counterexamples for many of the incorrect claims we found in the published literature. We show that these discrepancies are of potential practical importance because in some cases the chemical graphs of known molecules serve as counterexamples.

2. PRELIMINARIES

In this section we collect the definitions of the basic graph-theoretical objects that are used throughout this contribution. Our language is largely based on the book⁵⁴ by Thulasiraman and Swamy. We will consider simple undirected graphs $G(V, E)$ with vertex set V and edge set E . We regard an edge as an unordered pair of (distinct) vertices, i.e., $E \subseteq V \times V$. A subgraph $H(V', E')$ of $G(V, E)$ is a graph satisfying $V' \subseteq V$ and $E' \subseteq E$. A subgraph H is *spanning* if $V' = V$.

For a set \mathcal{H} of subgraphs of G we write $E[\mathcal{H}]$ for the union of all their edges, i.e.,

$$E[\mathcal{H}] := \bigcup_{H \in \mathcal{H}} \bigcup_{e \in H} \{e\} \quad (1)$$

We say that \mathcal{H} covers a subgraph $F(V', E')$ of G if $E' \subseteq E[\mathcal{H}]$. Let $v \in V$ be a vertex. The number of edges incident with v is called *degree* of v . A *path* P in G is a sequence of edges $(e_1, e_2, \dots, e_{k-1})$ with $e_i = \{v_i, v_{i+1}\}$, $i = 1, \dots, k-1$ such that the vertices v_i , $i = 1, \dots, k$ are all different, except possibly v_1 and v_k . A graph G is *connected* if for any two vertices $v, w \in V$, there is a path P such that $v_1 = v$ and $v_k = w$. The *connected components* of a graph G are its maximal connected subgraphs.

A *cycle* C in G is a subgraph in which every vertex has even degree. It will be convenient to identify a cycle with its edge set C and also with the binary vector that describes the edge-incidence structure of C . These incidence vectors (and we can hence say, the cycles in G) form a vector space over $\text{GF}(2)$ where the vector addition of two cycles C' , C'' corresponds to their symmetric difference

$$C' \oplus C'' = (C' \cup C'') \setminus (C' \cap C'') \quad (2)$$

This vector space is known as the *cycle space* $\mathbf{C}(G)$, and it has dimension $\dim \mathbf{C}(G) = \mu(G) = |E| - |V| + c(G)$, where $c(G)$ is the number of connected components of G . The parameter $\mu(G)$ is called the *cyclomatic number* or *first Betti number* of the graph G . A basis \mathcal{B} of the cycle space $\mathbf{C}(G)$ is called a *cycle basis*. Given a cycle basis \mathcal{B} there is a one-to-one relation between the cycles C and subsets $\mathcal{B}_C \subseteq \mathcal{B}$ such that

$$C = \bigoplus_{Z \in \mathcal{B}_C} Z \quad (3)$$

A cycle C is *elementary* if it is connected and each vertex has degree two. Elementary cycles meet our expectation of rings in a molecular graphs. A graph G is *2-connected* if any two vertices are contained in a common elementary cycle. The cycle space of G is the direct sum of the cycle spaces of the 2-connected components of G . Hence we restrict our attention to 2-connected graphs from now on.

A *chord* of C is an edge $e = \{x, y\} \in E$ such that $e \notin C$, but both x and y are vertices of C . (In the terminology of ref 16 x and y are “Nachbarpunkte”.) A cycle C is *simple* or

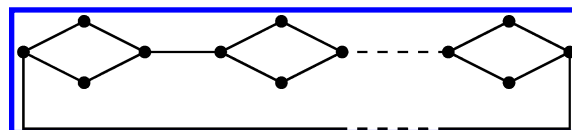


Figure 1. The set of relevant cycles of this graph contains $2^{|V|/4}$ cycles with $3|V|/4$ vertices and all $|V|/4$ quadrangles.⁵⁶ Each relevant cycle is a shortest cycle through one of its edges, thus $\mathcal{J}(G) = \mathcal{R}(G)$.

chordless if it is elementary and has no chord. A cycle C is *tied* if it is elementary and has exactly one chord.

A plane graph \hat{G} consists of a set V of points in \mathbb{R}^2 and a set E of line segments in \mathbb{R}^2 connecting exactly two points in V such that two lines intersect only in the points that they connect. A graph G is *planar* if there is a plane graph \hat{G} that is isomorphic to G . We say that \hat{G} is an *embedding* of G in the plane. The connected sets of $\mathbb{R}^2 \setminus \hat{G}$ are the *regions* of \hat{G} . There is exactly one unbounded region that will be called the infinite region; all other regions are bounded. Plane embeddings are equivalent to embeddings on the surface of a sphere \mathbb{S} . Here all regions are of course bounded. Each region of the spherical embedding can be made the infinite region of the plane embedding by a suitable projection. If \hat{G} is 2-connected, then each region is delimited by a unique elementary cycle in G which we call a *face*. Note that faces are not necessarily simple.

Since each edge of \hat{G} appears in exactly two faces the \oplus -sum of all faces is 0. The set of all faces except one, on the other hand, forms a basis of the cycle space. In particular, the faces belonging to the bounded regions of every plane embedding of G define a cycle basis for G which will be called a *planar cycle basis*.

The *length* $|C|$ of a cycle C is the number of its edges. The length of a cycle basis \mathcal{B} is the total length of its cycles, $l(\mathcal{B}) = \sum_{C \in \mathcal{B}} |C|$. A *minimum cycle basis* is a cycle basis of minimal length. A cycle is *relevant* if it cannot be written as \oplus -sum of shorter cycles. Equivalently, a cycle is relevant if it is contained in at least one minimum cycle basis.^{45,56} We write $\mathcal{R}(G)$ for the set of relevant cycles. The number of relevant cycles may grow exponentially with the number of vertices, see Figure 1 for an example.⁵⁶ Each cycle in a minimum cycle basis, i.e., each cycle in $\mathcal{R}(G)$, is simple.³³

For each edge $e \in E$ we write $\mathcal{J}(e)$ for the set of shortest cycles that contain e . The set of shortest cycles through the edges of G is $\mathcal{J}(G) = \bigcup_{e \in E} \mathcal{J}(e)$. For all graphs holds $\mathcal{J}(G) \subseteq \mathcal{R}(G)$.⁴⁹ The set $\mathcal{J}(G)$ may also grow exponentially since in the example in Figure 1 we have $\mathcal{J}(G) = \mathcal{R}(G)$.

A *spanning tree* T of a connected graph G is a spanning subgraph of G that does not contain cycles. It is easy to see that a spanning tree always exists in a connected graph. Kirchhoff³⁵ introduced the following construction for cycle bases of a graph in 1847: Suppose T is a spanning tree of G . Then for each edge $e \notin T$ there is a unique cycle in $T \cup \{e\}$ which is called a *fundamental cycle w.r.t. T* . The set of fundamental cycles belonging to a given spanning tree forms a basis of the cycle space which is called the *fundamental basis w.r.t. T* . For more details see e.g. ref 54. A cycle basis \mathcal{B} that is fundamental w.r.t. to some spanning tree T of G is called *Kirchhoff-fundamental*.

A collection \mathcal{B} of $\mu(G)$ cycles in G is called *fundamental*^{32,62} if there exists an ordering of these cycles such that

$$C_j \setminus (C_1 \cup C_2 \cup \dots \cup C_{j-1}) \neq \emptyset \text{ for } 2 \leq j \leq \mu(G) \quad (4)$$

Obviously, a fundamental collection \mathcal{B} is a cycle basis. A cycle basis \mathcal{B} is *strictly fundamental* if eq 4 is satisfied for *every* ordering of \mathcal{B} . It is shown in ref 30, Prop.2.1 that a cycle basis is strictly fundamental if and only if it is Kirchhoff-fundamental.

A sequence (Z_1, Z_2, \dots, Z_k) of elementary cycles is *cyclically well-arranged* if each partial sum is an elementary

$$Q_j = \bigoplus_{i=1}^{j-1} Z_i, \quad 1 \leq j \leq k \quad (5)$$

cycle. If (Z_1, Z_2, \dots, Z_k) is cyclically well-arranged, then the intersection $Z_j \cap Q_{j-1}$ consists of a path P_j . A basis \mathcal{B} consisting of elementary cycles is *cyclically robust* if for every elementary cycle C the set \mathcal{B}_C can be cyclically well-arranged.³⁴

3. SSSR: SMALLEST SET OF SMALLEST RINGS

Originally, the **S**mallest **S**et of **S**mallest **R**ings was defined as a minimum length Kirchhoff-fundamental basis, see e.g. ref 16, p 173. Deo et al.¹⁰ showed in 1982 that the problem of finding a Kirchhoff-fundamental cycle basis with minimum length is \mathcal{NP} -hard. In the more recent chemistry literature^{13,14,16} the term SSSR is also used as a synonym for *minimum cycle basis*. This tacit shift of definition reflects the wide-spread misconception that every cycle basis or at least every minimum cycle basis is strictly fundamental, i.e., derivable from a suitable spanning tree. In the same context the literature often speaks of “smallest fundamental cycles”, see e.g. ref 16, p 180. Below we show that the above-mentioned change of definition is indeed necessary to obtain a set of rings that deserves the designation SSSR in accordance with the chemical intuition. The reason is that a minimum Kirchhoff-fundamental basis, besides being hard to compute, often does not contain all smallest cycles of the graph.

Let us first clarify the concept of a “fundamental cycle” that is often used in the ring perception literature. Every *elementary* cycle appears in some Kirchhoff-fundamental cycle basis of G . To see this, observe that an elementary cycle C can be decomposed into a path P and an edge e that completes the cycle. The path P can be extended to a spanning tree T of G . By construction C is part of the Kirchhoff basis derived from T . The term “fundamental cycle” therefore simply means “elementary cycle”, when used without reference to a particular spanning tree T .

Not all cycle bases of a graph are fundamental.³² It is shown in ref 39, Cor.13, however, that every minimum cycle basis of a *planar* graph is fundamental. Nevertheless, there are planar graphs, such as the one in Figure 2, that do not have a strictly fundamental minimum cycle basis, i.e., for which no minimum cycle basis can be derived from a spanning tree.

For nonplanar graphs an even stronger negative result holds: There are graphs with nonfundamental minimum cycle bases. A nonfundamental minimum cycle basis of the complete graph K_9 is described in ref 39. The example in Figure 3, which is due to Champetier,⁷ has a unique minimum cycle basis consisting entirely of triangles that is nonfundamental.

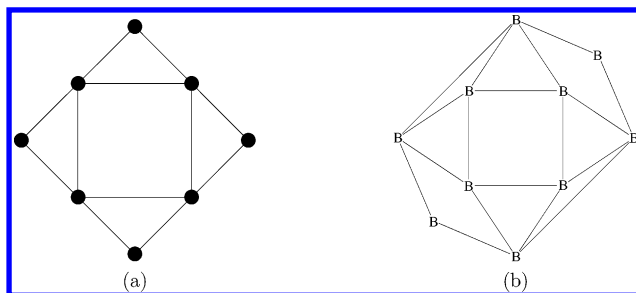


Figure 2. (a) The minimum cycle basis of this planar graph consists of the four triangles C_1 – C_4 and the central square C_5 . The ordering $(C_1, C_2, C_3, C_4, C_5)$ does not satisfy eq 4: in fact, $C_5 \subseteq C_1 \cup C_2 \cup C_3 \cup C_4 = E$. Alternatively one easily checks directly that no spanning tree generates the minimum cycle basis. (b) A chemical realistic example of the same type (extended by two “outer” quadrangles, respectively, triangles) is the decaborane molecule.²³

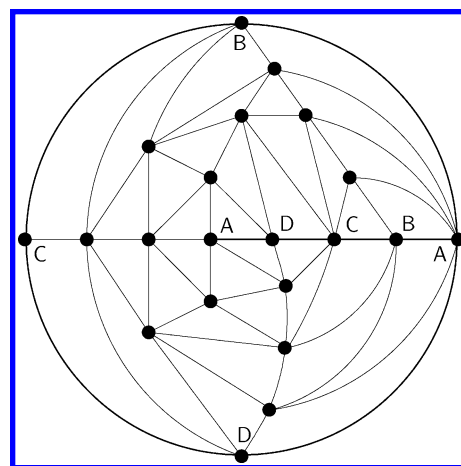


Figure 3. Champetier's graph.⁷ The pairs of points labeled A, B, C, and D, respectively, are identified, hence the 4-cycle ABCD appears twice in the drawing. This graph is null-homotopic (i.e., it has a cycle basis consisting exclusively of triangles), and all triangles are part of the unique minimum cycle basis. Each edge is contained in two triangles with the exception of the four edges of the 4-cycle ABCD which are contained in three triangles. Thus there is no ordering of the triangles satisfying eq 4.

Table 1. Worst Case Behavior of Minimum Cycle Basis Algorithms^a

algorithm		general	planar
Horton	1987 ^{33,28}	$O(E ^{2.376} V)$	$O(V ^{3.376})$
Balducci and Pearlman ^c	1994 ¹	$O(D\mu E ^2 V)$	$O(V ^5)$
Hartvigsen and Mardon (planar)	1994 ³¹		$O(V ^2 \log V)$
Vismara ^b	1997 ⁵⁶	$O(\mu E ^3)$	$O(V ^4)$
Berger, Gritzmam, de Vries	2004 ³	$O(E ^3)$	$O(V ^3)$

^a The maximum vertex degree is denoted by D ; the cyclomatic number is $\mu = |E| - |V| + 1$. ^b Vismara's algorithm computes “prototypes” for the set of all relevant cycles and produces a minimum cycle basis as a byproduct. ^c The estimate of the worst case complexity in ref 1 is incorrect. We give here the bound derived in ref 55.

Consequently, to compute an SSSR correctly in all cases, an algorithm has to be used that generates a minimum cycle basis. A number of polynomial-time algorithms for this task have been published in recent years. Their worst-case complexity is rather high. Nevertheless, the average performance on problems of practical interest seem to be much more favorable. In Table 1 we summarize the performance bounds of some of these approaches for general graphs and for planar graphs, where $|E| \leq 3|V| - 6$.

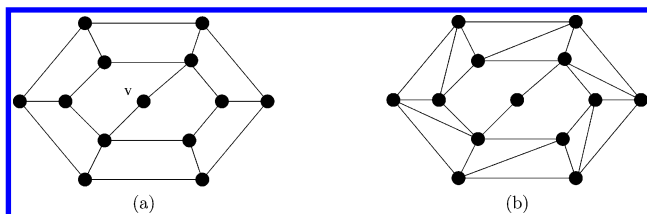


Figure 4. Counterexamples to the correctness of Figueras' algorithm: (a) For the central vertex v , one cycle of length five is found. After v is removed, the remaining cycles found during the second step of the algorithm must include a cycle of length six. On the other hand, a minimum cycle basis consists of quadrangles and two five cycles. (b) The algorithm stops after the first step because only vertices of degree four remain.

It is well-known that a minimum cycle basis of a graph is in general not unique. Plotkin⁴⁵ therefore introduced the set of \mathcal{H} -rings as the union of all minimum cycle bases (SSSRs). This set is also known as the set of *relevant cycles* $\mathcal{R}(G)$ of the graph G . It is shown in ref 56 that a cycle is relevant if and only if it cannot be represented as a \oplus -sum of strictly shorter cycles. The set of relevant cycles is in a sense the smallest set of short basis cycles that is *uniquely* defined for a graph, hence avoiding the ambiguities of a minimum cycle basis. Vismara⁵⁶ describes an algorithm for computing $\mathcal{R}(G)$, an efficient implementation of this algorithm in C++²⁶ is available at <http://www.tbi.univie.ac.at/~pmg/cycdeco/>.

Figueras' algorithm²⁰ for computing an SSSR has been celebrated as significant advance because of its computational speed. It is implemented e.g. in the Chemistry Development Kit (CDK).⁴⁸ The algorithm belongs to a class of methods whose idea is to iteratively remove edges and/or vertices from the graph after forming the shortest cycles through them, until no more cycles remain. Champetier's graph in Figure 3 shows that this algorithmic scheme must fail in general.

Figueras' method, however, may fail also for simple planar graphs:

(i) In the first step of the algorithm the set of vertices of degree two (the N_2 nodes in ref 20) is investigated. For each vertex of this type, the algorithm generates a shortest cycle (note that this cycle need not be unique) and adds it to a cycle set. Then all vertices of degree two are simultaneously removed from the graph. This procedure is repeated until no further N_2 nodes remain. The simultaneous removal of all degree two nodes was proposed as a remedy for the obviously erroneous strategy of removing them iteratively. Figure 4(a) shows, however, that the problem is by no means solved, even in graphs with a single vertex of degree two. Furthermore, the problem that the same cycle may be produced for two different vertices of degree two during Breadth First Search is not addressed in ref 20; this is a serious issue because it could result in a cycle set which is too small.

(ii) The algorithm assumes that after stage (i) there is always a vertex of degree 3 which is obviously not the case in general, see Figure 4(b).

Let us suppose, nevertheless, that there is such a vertex. In the second step, one of the edges in a shortest cycle containing this vertex is removed from the graph. Here, the edge is chosen that is incident with a longest shortest cycle through one of its endpoints. However, in Figure 4(a), no such choice can lead to a minimum cycle basis.

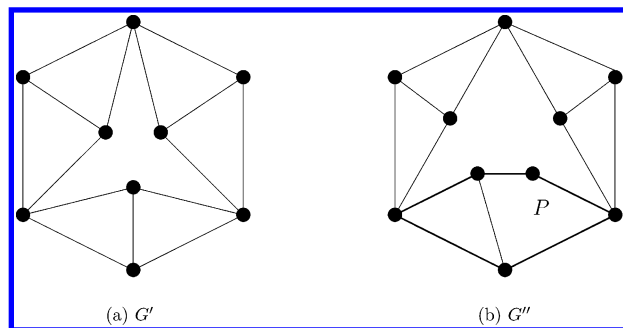


Figure 5. DESER and ESER are unrelated. The hexagonal outline H of the graph G' is ESER-dependent and DESER-independent. The situation is reversed in G'' . The outline hexagon is contained in minimum cycle bases of both G' and G'' , hence neither ESER nor DESER is a superset of a minimum cycle basis in general. For details see text.

4. ESER: ESSENTIAL SET OF ESSENTIAL RINGS

The **Essential Set of Essential Rings** was introduced by Fujita.^{21,22} Depending on atom types cycles are classified by the atoms that they contain as *carbon*, *heteroatom* (N, O, S, P), and *abnormal* (all other atoms). The definitions below are a rephrasing of those given in the review¹⁶ and in Vismara's dissertation,⁵⁵ respectively. For simplicity we assume here that all atoms are of the same type or at least comparable in the sense of Fujita.

For each simple cycle C define $\mathcal{T}[C]$ as the set of all tied cycles C' (belonging to the same atom-type class as C with at most the same number of heteroatoms and abnormal atoms, respectively) as C , that satisfy (i) $|C'| \leq |C|$ and (ii) $C' \cap C \neq \emptyset$. If (ii) is replaced by the stronger condition (iii) $2|C' \cap C| \geq |C'|$, i.e., at least half of the edges of C' are in C , we write $\mathcal{T}^*[C]$.

A simple cycle C is *ESER-dependent* if there is a subset $\mathcal{T}' \subseteq \mathcal{T}^*[C]$ such that $C \subseteq E[\mathcal{T}']$. The reviews^{13,16} give a slightly different definition: A simple cycle C is *DESER-dependent* if there is a subset $\mathcal{T}' \subseteq \mathcal{T}[C]$ such that (i) $C \subseteq E[\mathcal{T}']$, i.e., \mathcal{T}' covers C , and (ii) $2|C| < |E[\mathcal{T}']|$. Finally, $\text{ESER}(G)$ ($\text{DESER}(G)$) is the set of all simple cycles in G that are not ESER-dependent (DESER-dependent).

Consider the "outer" hexagon, H , of the graph G' in Figure 5(a). This example is taken from ref 55, p 78. The set $\mathcal{T}[H]$ consists of the three squares. The only subset of $\mathcal{T}[H]$ that covers H is $\mathcal{T}[H]$ itself. Furthermore $\mathcal{T}^*[H] = \mathcal{T}[H]$. Each square has two of its edges in common with H , hence H is ESER-dependent. On the other hand, $|E[\mathcal{T}[H]]| = 3 \times 4 = 12 \not> 2|H| = 12$, hence $H \in \text{DESER}(G)$.

Now consider the outer hexagon, H , of the graph G'' in Figure 5(b). The set $\mathcal{T}[H]$ now consists of the two tied squares Q_1 and Q_2 and the tied pentagon P . We have $|E[\mathcal{T}[H]]| = 2 \times 4 + 5 = 13 > 2|H| = 12$, thus H is DESER-dependent. However, $\mathcal{T}^*[H] = \{Q_1, Q_2\}$ does not cover H , hence $H \in \text{ESER}(G)$.

Thus the definitions of ESER and DESER give unrelated cycle sets. Downs¹³ mentions that "the ESER is in general always a superset of an SSSR". Similarly, Fujita [see ref 21, Figure 2] claims that $\text{ESER}(G)$ contains an SSSR. This is incorrect as the graph G' shows for ESER and the graph G'' shows for DESER. The relevant cycles of G' are the 6 triangles (all of which are essential in the sense of they are contained in every MCB) and all $2^3 = 8$ interchangeable (in the sense of²⁷) hexagons, of which one is contained in every

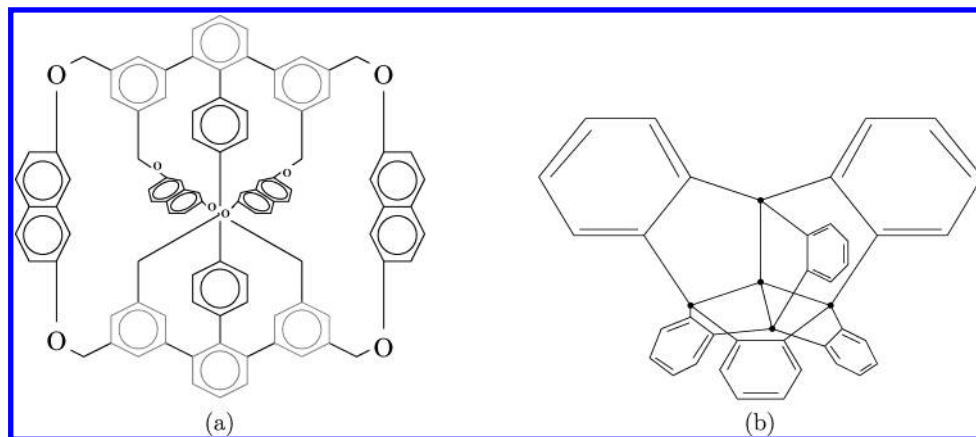


Figure 6. Two nonplanar chemical graphs: (a) kuratowskiphane,⁸ containing a $K_{3,3}$ as a minor, and (b) centrohexaindane³⁸ containing a K_5 as a minor. The two graphs K_5 and $K_{3,3}$ are the forbidden minors for planarity in Wagner's theorem.⁵⁸

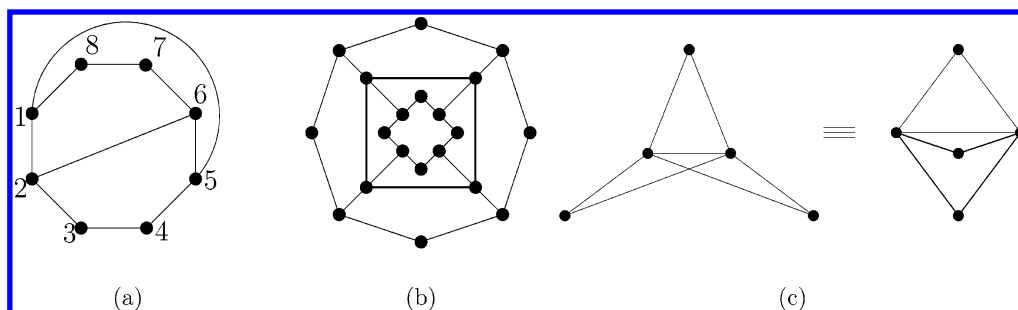


Figure 7. Three examples of planar graphs for which all minimum cycle bases are nonplanar. (a) This graph is taken from ref 39. No plane embedding has the face $Q = (1, 2, 6, 5)$. A minimum cycle basis contains Q and two of the cycles $(2, 3, 4, 5, 6)$, $(1, 2, 6, 7, 8)$, $(1, 2, 3, 4, 5)$, and $(1, 5, 6, 7, 8)$ and hence has length $\neq 14$. The planar bases have length 15. (b) The bold square Q in this graph, taken from ref 17, Figure 12, is the shortest cycle and hence contained in every minimum cycle basis. It cannot appear as a face, however. (c) The [1.1.1]propellane molecule (tricyclo[1.1.1.0^{1,3}]pentane, see e.g. ref 63). The relevant cycles are the three triangles, while the faces of every planar embedding are two triangles and one quadrangle.

minimum cycle basis. The relevant cycles of G'' are the 5 triangles and the square contained in P (each of these cycles is essential), and the $2^2 = 4$ hexagons (one of which is contained in every minimum cycle basis). Thus, neither $\text{ESER}(G')$ nor $\text{DESER}(G'')$ contain a minimum cycle basis. In fact, they even do not contain any cycle basis.

5. MINIMUM PLANAR CYCLE BASES

Some authors focus entirely on planar graphs in the context of chemical ring perception, see e.g. refs 17 and 19. It should be noted, however, that there are nonplanar chemical graphs. Three spectacular examples are given in Figures 6 and 10, other examples are described e.g. in refs 2, 37, 44, 47, and 53.

Recall that a cycle basis of G is *planar* if it consists of the faces belonging to the bounded regions of a plane embedding of G . A basis has a minimal length among all planar cycle bases if and only if the length of face F_∞ belonging to the unbounded region is maximal, because $l(\mathcal{B}) = 2|E| - |F_\infty|$ for any planar cycle basis. We will call such a basis a *minimal planar cycle basis*.

Plane embeddings can be computed in $O(|V|)$ time, see e.g. refs 9 and 46. Unfortunately, the embedding of a planar graph on the sphere S is in general not unique, unless G is 3-connected (that is, any subset $S \subseteq V$ for which $G \setminus S$ is not connected has cardinality at least 3).⁶⁰ Algorithms are available that can produce all embeddings of G in the plane.⁴ The computation of one or all minimum planar cycle bases

can be achieved e.g. by the integer linear programming approach outlined in ref 41. The number of planar embeddings of a 2-connected graph may grow exponentially with the number of vertices.⁴³ In the graph in Figure 1, for instance, every relevant cycle is a face in some embedding.

We denote by $\text{Faces}(G)$ the set of all possible faces in G , i.e., the set of all cycles that are faces in some embedding on S .

It is well-known that $\text{Faces}(G)$ in general does not contain a minimum cycle basis. In Figure 7 we give three examples. Two are taken from the literature, and the third one is the graph of the molecule [1.1.1]propellane, see e.g. ref 63.

6. ESSR: EXTENDED SET OF SMALLEST RINGS

The **Extended Set of Smallest Rings** was introduced by Downs et al.¹⁷ as an approach to design an optimal ring set for retrieval purposes. ESSR by definition is limited to planar graphs. Paraphrasing the original definition, a cycle C is in $\text{ESSR}(G)$ provided it satisfies at least one of the following conditions:

- (i) There is a planar embedding of G such that C is a chordless face. ("Simple faces")
- (ii) C is a shortest cycle through at least one of its edges. ("Class I cut face")
- (iii) There is a planar embedding of G such that $|C| \geq |C'|$ for all faces C' adjacent to C , and there is at least one adjacent face C'' for which $|C| = |C''|$. ("Class II cut face")

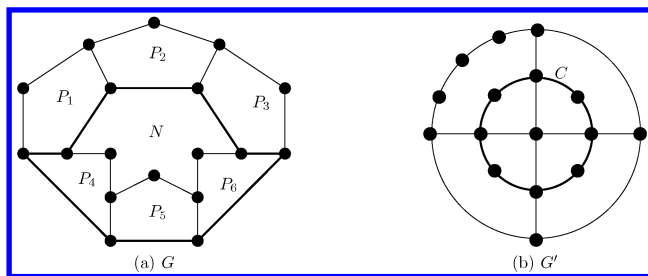


Figure 8. (a) Not all relevant cycles are in ESSR. Graph G has $|E| = 24$, $|V| = 18$, i.e., $\mu(G) = 7$. The relevant cycles are the six pentagons P_i , $i = 1, \dots, 6$, which form $\mathcal{J}(G)$, and the octagon O shown in bold. O is obviously linearly independent from the pentagons and it is the next-shortest cycle in G . On the other hand, there is no planar embedding in which O is a face and O does not have an adjacent octagon. Thus $O \notin \text{ESSR}(G)$. (b) Cycles in $\text{ESSR}(G')$ need not be relevant. The cycle C is not relevant because it is the sum of the (smaller) faces in its interior. One easily verifies that there is no planar embedding that has C as a face. $|C| = 8$ while the shortest cycle through each of its edges is contained in a quadrangle. On the other hand, C fulfills (iii) as the upper left face adjacent to C also has length 8.

An algorithm for computing the ESSR is described in ref 15. Downs et al. [see ref 17, p 192] claim that $\text{ESSR}(G) = \text{Faces}(G) \cup \mathcal{R}(G)$. The graph G in Figure 8(a), which is used to give an example of a “class III cut face” (a set of cycles not contained in ESSR) in ref 17, Figure 16 and ref 13, Figure 7, however, provides an example of a graph for which $\mathcal{R}(G) \not\subseteq \text{ESSR}(G)$.

It is also not true that $\text{ESSR}(G) \subseteq \text{Faces}(G) \cup \mathcal{R}(G)$. Figure 8(b) shows a graph containing a cycle C that satisfies condition (iii) but is neither relevant nor the face of any planar embedding.

The problem with the definition in ref 17 is that there are cycles, such as C in Figure 8(b) that are both “class II” and “class III cut faces” as defined in ref 17. It appears that $\text{ESSR}'(G) = \text{Faces}(G) \cup \mathcal{R}(G)$ is indeed a useful definition for a chemically relevant “extended set of smallest rings” that could be generated reasonably efficiently by computing $\mathcal{R}(G)$ using Vismara’s algorithm⁵⁶ and $\text{Faces}(G)$ e.g. with the aid of SPQR-trees.⁴¹

7. SET OF SMALLEST CYCLES AT EDGES: SSCE

Let $G(V, E)$ be a 2-connected graph. Set $G_0 := G(V, E_0)$ with $E_0 := E$ and define $G_k(V, E_k)$ recursively as the graph with vertex set V and edge set

$$E_k = \{e \in E \mid e \text{ is contained in exactly one cycle } C \in \mathcal{J}(G_{k-1})\} \quad (6)$$

In other words, in each step we remove all edges that are covered at least twice by $\mathcal{J}(G_{k-1})$. There is a smallest number k such that $E_k = E_{k+1}$ which we denote by $\xi(G)$. Clearly, $\xi(G) = 0$ if G is a cycle.

Dury et al.¹⁸ define the Set of Smallest Cycles at Edges (SSCE) as the union

$$\text{SSCE}(G) = \bigcup_{k=0}^{\xi(G)} \mathcal{J}(G_k) \quad (7)$$

In addition, ref 18 contains provisions to remove certain cycles from $\mathcal{J}(G_k)$ for $k \geq 1$ which explicitly depend on finding planar embeddings with the shortest exterior cycle. The discussion in ref 18 in fact refers to planar graphs only. Here, we disregard this complication, which feels unnatural from a mathematical point of view and consider $\text{SSCE}(G)$ for arbitrary graphs.

The example in Figure 9 shows that $\text{SSCE}(G)$ in general does not contain a cycle basis of G . If $\text{SSCE}(G)$ contains a cycle basis of G , it is not always minimal. To see this consider the graph in Figure 8(a). We set $G_0 := G$ and obtain $\mathcal{J}(G_0) = \{P_1, P_2, P_3, P_4, P_5\}$. The graph G_1 then consists of the 9-cycle N and the outline N' , which also has length 9. Of course $\mathcal{J}(G_1) = \{N, N'\}$, and $G_2 = G_1$, i.e., $\xi(G) = 1$. A minimum cycle basis of G , on the other hand, contains the octagon O . Figure 10 shows another, rather spectacular molecular example.

As a further remark, it is by no means clear how to compute $\mathcal{J}(G)$ efficiently for a general graph because there may be exponentially many different shortest cycles passing through a given edge.

8. THE SET OF B-RINGS

The set of β -rings⁴² is one of the earliest attempts to extend the SSSR to include the hexagon of norbornane, depicted in Figure 11(a), but without including much larger “envelope” rings in other molecules such as naphthalene in Figure 11-(b). A β -set is obtained from the length-sorted list of all chordless faces of a plane embedding by the same greedy procedure that is used to compute the relevant cycles $\mathcal{R}(G)$: The algorithm processes the cycles by increasing length and includes all cycles of a given length that are linearly independent from all *shorter* cycles already contained in the set.

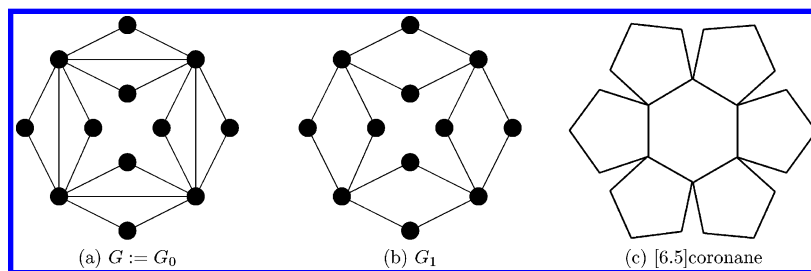


Figure 9. $\text{SSCE}(G)$ does not contain a cycle basis in general. The shortest cycles $\mathcal{J}(G_0)$ of (a) are the eight triangles; removing their common edges erases the central square (which is contained in the unique minimum cycle basis of G). The 4-cycles in G_1 are edge-disjoint, hence $\xi(G) = 1$ and $\text{SSCE}(G)$ consists of the eight triangles and the four 4-cycles that are the \oplus -sum of adjacent triangles. Thus $\text{SSCE}(G)$ contains only 8 linearly independent cycles, while $\mu(G) = 20 - 12 + 1 = 9$. (c) The graph of 6.5.coronane (heptacyclo-[19.3.0.0.1.50.5.90.9.130.13.170]tetracosan)⁵⁹ serves as a chemical counterexample: here $\text{SSCE}(G)$ consists only of the six pentagons.

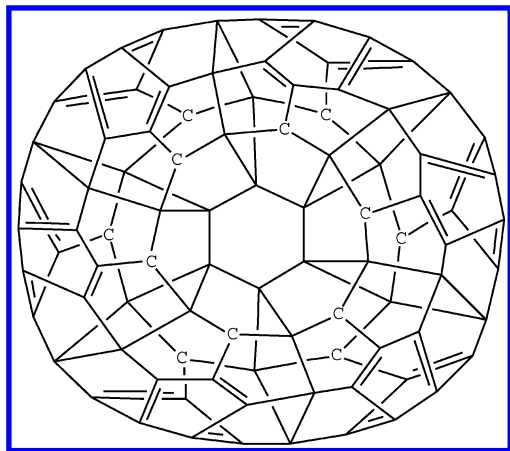


Figure 10. The set of relevant cycles of the toroidal molecule [5,6]fullerene-C90-D6d-8,10,12,14,16,18,79,81,83,85,87,89-dodecayl²⁴ consists of all 66 pentagons and the hexagon in the center. There are 90 vertices and 150 edges and thus $\mu(G) = 61$. A minimum cycle basis consists of 60 pentagons and the hexagon and hence has length $\ell = 306 > 2|E|$. Thus there are edges that are covered by at least three cycles in every cycle basis. MacLane's theorem⁴⁰ thus implies that G is nonplanar. The set $\mathcal{J}(G_0)$ consists of all pentagons, and every edge is covered by at least two of them. Hence $G_1 = \emptyset$ and $\xi(G) = 1$. Thus $\text{SSCE}(G)$ does not contain the essential hexagon and consequently no cycle basis.

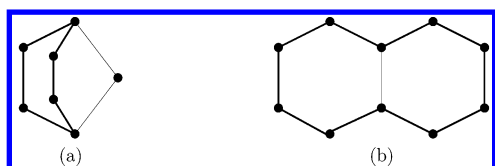


Figure 11. (a) The graph of the norbornane molecule and (b) naphthalene molecule.

In case of the β -rings, this method contains two modifications: (i) all three and four-cycles are in β irrespective of linear dependence, and (ii) instead of testing linear independence from all shorter cycles one checks only that a cycle is linearly independent from all subsets containing at most three shorter β -rings.

In general, the set of β -rings does not contain a minimum cycle basis because not all planar graphs have a minimum cycle basis consisting of faces. Furthermore, the definition is not unique since it depends on a particular planar embedding.

We propose to use a variation β^* of the definition of β -rings that contains all chordless cycles instead of the faces of a particular plane embedding as the starting set. Clearly, β^* is uniquely defined given the graph G and one easily verifies $\mathcal{R} \subseteq \beta^*$.

9. THE SET OF ALL ELEMENTARY CYCLES

The generation of the set $\mathcal{C}^*(G)$ of all elementary cycles of G is an alternative to the restricted sets of "short" cycles considered so far. So-called "cycle space algorithms" attempt to construct $\mathcal{C}^*(G)$ from a basis \mathcal{B} by iteratively combining two elementary cycles using the \oplus operation and retaining the result if and only if it is again an elementary cycle. Let us denote the set of elementary cycles that can be obtained in this way by $\langle \mathcal{B} \rangle$. In this language, a basis is cyclically robust if and only if $\langle \mathcal{B} \rangle = \mathcal{C}^*(G)$.

Dixon and Goodman¹¹ conjectured that every strictly fundamental cycle basis is cyclically robust. A counter-

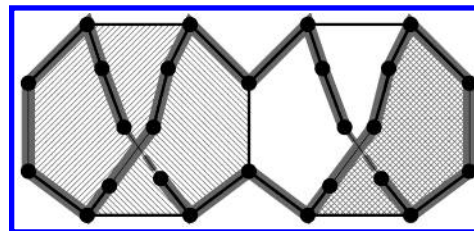


Figure 12. The bold edges indicate a spanning tree T . The elementary cycle $C = \oplus_{i=1}^6 C_i$ (highlighted) has the property that $C \oplus C_i$ is not elementary for every basis cycle $i = 1, \dots, 6$. In this case the Kirchhoff-basis w.r.t. T is a minimum cycle basis since $|C_i| = 8$ for all $i = 1, \dots, 6$ and there is no shorter cycle.

example, however, was given in ref 50 consisting of a simple cycle $C = \oplus_{i=1, \dots, \mu} C_i$, with the property that for each $i = 1, \dots, \mu$, $C \oplus C_i$ is the union of edge-disjoint cycles, i.e., a nonelementary cycle. This example—slightly modified—is drawn in Figure 12. Later it was shown that every planar cycle basis is cyclically robust.¹² Iterative procedures thus have a chance to work correctly when they start from a planar basis. Examples are given in refs 51 and 12 where a parallel implementation is described. Furthermore, complete graphs and complete bipartite graphs have cyclically robust bases.³⁴ Whether all graphs have cyclically robust bases is still unknown, however. The example in Figure 12 shows, furthermore, that minimum cycle bases are not always robust.

Recently, a modification of the iterative procedure called *cyclical conjunction* was proposed as "an efficient operator for the extraction of cycles from a graph".^{5,6} We paraphrase here the original rather complicated definition. Let C and C' be two elementary cycles (including the empty cycle) and let us denote their cyclical conjunction [The papers^{5,6} use \diamond .] by $C \diamond C'$. Then $C \diamond C' = \emptyset$ if $C \cap C' = \emptyset$. If $C \cap C'$ is a path, then $C \diamond C' = C \oplus C'$. In general, however, $C \cap C'$ is a collection of disjoint paths. In the general case, the procedure outlined in ref 6 would produce one of the elementary cycles which constitute $C \oplus C'$. Cyclical conjunction thus differs from the simpler iterative procedure in that it retains one of the elementary components whenever the \oplus -sum of two elementary cycles is nonelementary.

The claim is that, starting from a Kirchhoff basis \mathcal{B} of G , repeated application of \diamond produces all elementary cycles of G . Syslo's counterexample in Figure 12, however, is also applicable to cyclical conjunction. The cycle C would not be obtained, as a nonelementary cycle is not kept in the set of cycles produced by the \diamond operation; by choice of C , this occurs for any order in which the \diamond procedure is carried out. By the uniqueness of basis representations, C cannot be obtained by a \diamond -operation of two other elementary cycles.

We remark that fast algorithms for extracting $\mathcal{C}^*(G)$ that are based on backtracking rather than cycle space operations are known, see e.g. ref 29 and the references therein.

The Set of Elementary Rings (SER) defined by Takahashi⁵² is, in our notation, given by $\text{SER} = \langle \mathcal{M} \rangle$, where \mathcal{M} is a minimum cycle basis. The example above shows that SER is in general only a subset of the set $\mathcal{C}^*(G)$ of elementary cycles. Furthermore, it is not uniquely defined: If a different minimum cycle basis \mathcal{M} than \mathcal{B} is chosen in Figure 12, for instance the one that consists out of the three cycles of length eight indicated by the different shades in the background and the three cycles obtained as their mirror

images along the central vertical edge, then one readily verifies $C \in \langle \mathcal{M} \rangle$, whereas we have shown that $C \notin \langle \mathcal{B} \rangle$.

10. DISCUSSION

A common problem of at least some of the chemically motivated ring sets is a lack of abstract mathematical foundation since they arose from intuitions about a limited set of rather special examples. The definitions, therefore, besides feeling artificial and unnecessarily complicated, often do not transfer to more general molecular graphs. Consequently, many proposed algorithms bear heuristic features, and their results have to be carefully examined before attempting further deductions. Based on this observation we propose here that the definition of a *good chemical ring set* should possess four fundamental properties:

(1) *Uniqueness*. The cycle set \mathcal{X} should be mathematically well-defined in the sense that different implementations of the algorithms invariably produce the same set of cycles, independent of permutations of the edge and vertex labels. This guarantees that molecular descriptors for comparison and substructure search derived from the cycle set are uniquely defined.

(2) *Completeness* of description. \mathcal{X} must be sufficient to describe the complete ring structure, i.e., \mathcal{X} can generate all cycles by means of one or more well-defined operations. This way, it is guaranteed that all further cycles can be computed from \mathcal{X} if desired.

(3) *Discrimination*. There should be only very few graphs/molecules with the same set of cycles. This is a prerequisite for efficient databases searches. On the other hand, cycle sets should not be too large, for although all cycles of a graph of course yield complete information, an interpretation becomes more difficult in addition to the computational issues.

(4) *Efficiently computable*. This does not necessarily mean polynomial time algorithms for general graphs, because chemical graphs are (relatively) small and close to planar.

We submit that it is much less important that the cycle set is

(*) *Chemically relevant*.

The reason is that “chemical relevance” is not a well-defined and context independent property. As we have seen above, attempts to create sets of “short cycles” purely based on chemical intuition, which is naturally inspired by “normal” organic molecules, leads to undesired artifacts and ad hoc modifications of ring sets. Furthermore, the “typical” examples may be different from a substantial fraction of molecules that are known today and that are stored in databases. This is shown by the fact that at least some of our counterexamples are real molecular graphs.

In Table 2 we summarize the properties of the cycle sets studied in the previous sections with respect to the desirable characteristics listed above. To ensure point (2), it is important that a cycle basis is contained in the set, further, from a chemically inspired point of view and with respect to the second part of (3), a minimum cycle basis would be best. Of course, cycle sets only describe two-connected components of molecular graphs, and graph descriptors derived from them cannot distinguish between isomers or tautomers of a molecule. On the other hand, by Whitney’s theorem,⁶¹ if the cycle bases of two-connected graphs can be mapped to each other, the graphs are already isomorphic

Table 2. Summary of the Properties of the Ring Sets Considered in This Contribution^a

ring set	type	unique	contains basis	contains MCB	contains $\mathcal{R}(G)$	size
SSSR	<i>G</i>	no	yes	MCB	no	$O(m)$
ESER/DESER	<i>G</i>	yes	no	no	no	exp
Faces(<i>G</i>)	<i>P</i>	yes	yes	no	no	exp
ESSR	<i>P</i>	yes	yes	no	no	exp
SSCE	<i>G</i>	yes	no	no	no	exp
β -rings	<i>P</i>	no	yes	no	no	$O(m + n^4)$
SER	<i>G</i>	no	yes	yes	no	exp
elementary cycles	<i>G</i>	yes	yes	yes	yes	exp
\mathcal{H} -rings						
relevant cycles	<i>G</i>	yes	yes	yes	yes	exp

^a MCB is used as the abbreviation for “minimum cycle basis”. Ring sets of type *G* are defined for general graphs, and those of type *P* for planar graphs only. A worst case size that increases exponentially with the number of vertices is indicated by “exp”.

Table 3. Properties of Three “Well-Behaved” Ring Sets

ring set	type	unique	contains			size
			basis	MCB	$\mathcal{R}(G)$	
ESSR' = Faces(<i>G</i>) \cup $\mathcal{R}(G)$	<i>P</i>	yes	yes	yes	yes	exp
β^* -rings	<i>G</i>	yes	yes	yes	yes	exp
$\mathcal{R}(G)$	<i>G</i>	yes	yes	yes	$\mathcal{R}(G)$	exp

up to small modifications. For this reason, graph descriptors derived from a cycle set containing a cycle basis are good candidates for fulfilling (3). Among the studied cycle sets, only the set of elementary cycles fulfills all requirements. Its drawback is that it is generally very large.

Hence, we propose the three sets of Table 3 as more convenient alternatives. Among them, the set of relevant cycles $\mathcal{R}(G)$, is of particular interest because it can be readily computed using Vismara’s algorithm.⁵⁶ While all good cycle sets have an exponential size in the worst case, they nevertheless seem to be well within computational reach for molecular graphs.

ACKNOWLEDGMENT

This work is supported by the Austrian *Fonds zur Förderung der Wissenschaftlichen Forschung*, Project Nos. P-13565 (P.M.G) and P14094 (P.M.G, J.L), by the Schwerpunktprogramm Nr. 1126 of the Deutsche Forschungsgemeinschaft (F.B), and the DFG Bioinformatics Initiative (P.F.S).

REFERENCES AND NOTES

- (1) Balducci, R.; Pearlman, R. S. Efficient exact solution of the ring perception problem. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 822–831.
- (2) Benner, S. A.; Maggio, J. E.; Simmons III, H. E. Rearrangement of a geometrically restricted triepoxide to the first topologically nonplanar molecule. A reaction path elucidated by using oxygen isotope effects on carbon-13 chemical shifts. *J. Am. Chem. Soc.* **1981**, *103*, 1581–1582.
- (3) Berger, F.; Gritzmann, P.; de Vries, S. Minimum cycle bases for network graphs. *Algorithmica* In press.
- (4) Cai, J. Counting embeddings of planar graphs using DFS trees. *SIAM J. Discrete Math.* **1993**, *6*, 335–352.
- (5) Cerruela García, G.; Luque Ruiz, I.; Gómez-Nieto, M. A. Parallel algorithms for graph cycle extraction using the cyclical conjunction operator. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1398–1406.
- (6) Cerruela García, G.; Luque Ruiz, I.; Gómez-Nieto, M. A. Cyclical conjunction: An efficient operator for the extraction of cycles from a graph. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1415–1424.

- (7) Champetier, C. On the null-homotopy of graphs. *Discr. Math.* **1987**, *64*, 97–98.
- (8) Chen, C.-T.; Gantzel, P.; Siegel, J. S.; Baldridge, K.; English, R. B.; Ho, D. M. Synthesis and structure of the nanomolecular multicyclophane “Kuratowskiphane”, an achiral nonplanar $K_{3,3}$ topological stereoisomer. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 2657–2660.
- (9) Chiba, N.; Nishizeki, T.; Abe, A.; Ozawa, T. A linear algorithm for embedding planar graphs using PQ-trees. *J. Comput. Sys. Sci.* **1985**, *30*, 54–76.
- (10) Deo, N.; Prabhu, G. M.; Krishnamoorthy, M. S. Algorithms for generating fundamental cycles in a graph. *ACM Trans. Math. Software* **1982**, *8*, 26–42.
- (11) Dixon, E. T.; Goodman, S. E. An algorithm for the longest cycle problem. *Networks* **1976**, *6*, 139–146.
- (12) Doğrusöz, U.; Krishnamoorthy, M. Enumerating all cycles of a planar graph. *J. Parallel Algorithms Appl.* **1996**, *10*, 21–36.
- (13) Down, G. M. Ring perception. In *Encyclopedia of Computational Chemistry*; von Ragué Schleyer, P., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., Schreiner, P. R., Eds.; John Wiley & Sons Ltd.: Chichester, UK, 1998; Vol. 4; pp 2509–2515.
- (14) Downs, G. M. Rings – the importance of being perceived. In *Chemical Structures 2*; Warr, W. A., Ed.; Springer-Verlag: Berlin, 1993; pp 207–219.
- (15) Downs, G. M.; Gillet, V. J.; Holliday, J. D.; Lynch, M. F. Computer storage and retrieval of generic chemical structures in patents. 9. An algorithm to find the Extended Set of Smallest Rings in structurally explicit generics. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 207–214.
- (16) Downs, G. M.; Gillet, V. J.; Holliday, J. D.; Lynch, M. F. Review of ring perception algorithms for chemical graphs. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 172–187.
- (17) Downs, G. M.; Gillet, V. J.; Holliday, J. D.; Lynch, M. F. Theoretical aspects of ring perception and development of the Extended Set of Smallest Rings concept. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 187–206.
- (18) Dury, L.; Latour, T.; Leherde, L.; Barberis, F.; Vercauteren, D. P. New graph descriptor for molecules containing cycles. Application as screening criterion for searching molecular structures within large databases of organic compounds. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1437–1445.
- (19) Elk, S. Derivation of the principle of Smallest Set of Smallest Rings from Euler's polyhedron equation and a simplified technique for finding this set. *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 203–206.
- (20) Figueras, J. Ring perception using breadth-first search. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 986–991.
- (21) Fujita, S. A new algorithm for selection of synthetically important rings: the Essential Set of Essential Rings for organic structures. *J. Chem. Inf. Comput. Sci.* **1987**, *28*, 78–82.
- (22) Fujita, S. Logical perception of ring opening, ring closure, and rearrangement reactions based on imaginary transition structures. Selection of the Essential Set of Essential Rings. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 1–9.
- (23) Gaines, D. F.; Beall, H. Hydrogen–deuterium exchange in decaborane-14: Mechanistic studies. *Inorg. Chem.* **2000**, *39*, 1812–1813.
- (24) Galpern, E. G.; Pinyaskin, V. V.; Stankevich, I. V.; Nesmeyanov, A. N. Carbon clusters containing sp^2 - and sp^3 -hybridized atoms. *Izv. Akad. Nauk SSSR, Ser. Khim.* **1996**, *3*, 551–553.
- (25) Garey, M. R.; Johnson, D. S. *Computers and Intractability*; W. H. Freeman and Company: New York, 1979.
- (26) Gleiss, P. M. *Short Cycles*; Ph.D. Thesis, University of Vienna, 2001.
- (27) Gleiss, P. M.; Leydold, J.; Stadler, P. F. Interchangeability of relevant cycles in graphs. *Elec. J. Comb.* **2000**, *7*, R16 [16 pp]. www.combinatorics.org.
- (28) Golynski, A.; Horton, J. D. A polynomial time algorithm to find the minimal cycle basis of a regular matroid. In *Algorithm Theory – SWAT 2002, Turku, Finland*, volume 2368 of *Lecture Notes in Computer Science*; 2002; pp 551–57.
- (29) Hanser, T.; Jauffret, P.; Kaufmann, G. A new algorithm for exhaustive ring perception in a molecular graph. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 1146–1152.
- (30) Hartvigsen, D.; Mardon, R. Cycle bases from orderings and coverings. *Discr. Math.* **1991**, *94*, 81–94.
- (31) Hartvigsen, D.; Mardon, R. The all-pair min-cut problem and the minimum cycle basis problem on planar graphs. *SIAM J. Discr. Math.* **1994**, *7*, 403–418.
- (32) Hartvigsen, D.; Zemel, E. Is every cycle basis fundamental? *J. Graph Theory* **1989**, *13*, 117–137.
- (33) Horton, J. D. A polynomial-time algorithm to find the shortest cycle basis of a graph. *SIAM J. Comput.* **1987**, *16*, 359–366.
- (34) Kainen, P. C. On robust cycle bases. In *Electronic Notes in Discrete Mathematics*; Alavi, Y., Jones, E., Lick, D. R., Liu, J., Eds.; Elsevier Science Publishers: 2002; Vol. 11. Proceedings of the Ninth Quadren-
- nal International Conference on Graph Theory, Combinatorics, Algorithms and Applications; June 4–9 2000, Kalamazoo, MI.
- (35) Kirchhoff, G. Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird. *Poggendorf Ann. Phys. Chem.* **1847**, *72*, 497–508.
- (36) Köbler, J.; Schöning, U.; Torán, J. *The Graph Isomorphism Problem: Its Structural Complexity*; Progress in Theoretical Computer Science. Birkhäuser, Boston, 1993.
- (37) Kuck, D. Centrohexacyclic or K_5 molecules – development of a growing class of topologically nonplanar organic compounds. *Liebigs Ann.* **1997**, 1043–1057.
- (38) Kuck, D.; Schuster, A. Centrohexaindan, der erste Kohlenwasserstoff mit topologisch nicht-planarer Molekülstruktur. *Angew. Chem.* **1988**, *100*, 1222–1224.
- (39) Leydold, J.; Stadler, P. F. Minimal cycle basis of outerplanar graphs. *Elec. J. Comb.* **1998**, *5*, R16. See <http://www.combinatorics.org> and Santa Fe Institute Preprint 98-01-011.
- (40) MacLane, S. A combinatorial condition for planar graphs. *Fundam. Math.* **1937**, *28*, 22–32.
- (41) Mutzel, P.; Weiskircher, R. Optimizing over all combinatorial embeddings of a planar graph. In *Integer Programming and Combinatorial Optimization*; Cornuejols, G., Burkard, R. E., Wöginger, G. J., Eds.; 7th International IPCO Conference, Graz, Austria, June 9–11, volume 1610 of *Lecture Notes in Computer Science*; Springer: Berlin, 1999; pp 361–376.
- (42) Nickelsen, H. Ringbegriffe in der Chemie-Dokumentation. *Nachr. Dok.* **1971**, *3*, 121–123.
- (43) Nishizeki, T.; Chiba, N. *Planar Graphs: Theory and Algorithms*, volume 32 of *Annals of Discrete Mathematics*; North-Holland: New York, 1988.
- (44) Paquette, L. A.; Williams, R. V.; Vazeux, M.; Browne, A. R. Factors conducive to the cascade rearrangement of sterically congested and geometrically restricted three-membered rings. Facile synthesis of a topologically nonplanar heterocycle. *J. Org. Chem.* **1984**, *49*, 2194.
- (45) Plotkin, M. Mathematical basis of ring-finding algorithms in CIDS. *J. Chem. Doc.* **1971**, *11*, 60–63.
- (46) Shiha, W.-K.; Hsub, W.-L. A new planarity test. *Theor. Comput. Sci.* **1999**, *223*, 179–191.
- (47) Sritana-Anant, Y.; Seiders, T. J.; Siegel, J. S. Design of novel aromatics using the Loschmidt replacement on graphs. In *Carbon Rich Compounds I*, volume 196 of *Topics in Current Chemistry*; de Meijere, A., Ed.; Springer-Verlag: Berlin, Heidelberg, 1998; pp 1–43.
- (48) Steinbeck, C.; Han, Y.; Kuhn, S.; Horlacher, O.; Luttmann, E.; Willighagen, E. The chemistry development kit (CDK): An open-source Java library for chemo- and bioinformatics. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 493–500.
- (49) Stepanec, G. F. Basis systems of vector cycles with extremal properties in graphs. *Uspekhi Mater. Nauk.* **1964**, *19*, 171–175 (Russian).
- (50) Syslo, M. M. Characterizations of outerplanar graphs. *Discrete Math.* **1979**, *26*, 47–53.
- (51) Syslo, M. M. An efficient cycle vector space algorithm for listing all cycles of a planar graph. *SIAM J. Comput.* **1981**, *10*, 797–808.
- (52) Takahashi, Y. Automatic extraction of ring substructures from a chemical structure. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 167–170.
- (53) Tellenbröcker, J.; Kuck, D. Extending the chemistry of [5.5.5.5]-fenestranes – 8-fold peripheral functionalization of fenestrandanes. *Eur. J. Org. Chem.* **2001**, 1483–1489.
- (54) Thulasiraman, K.; Swamy, M. N. S. *Graphs: Theory and Algorithms*; J. Wiley & Sons: New York, 1992.
- (55) Vismara, P. *Reconnaissance et représentation d'éléments structuraux pour la description d'objets complexes. Application à l'élaboration de stratégies de synthèse en chimie organique*. Ph.D. Thesis, Université Montpellier II, France, 1995, 95-MON-2-253.
- (56) Vismara, P. Union of all the minimum cycle bases of a graph. *Electron. J. Comb.* **1997**, *4*, 73–87. Paper No. #R9 (15 pp).
- (57) Voss, H.-J. *Cycles and Bridges in Graphs*; Kluwer: Dordrecht, 1991.
- (58) Wagner, K. Bemerkungen zum Vierfarbenproblem. *J. Deut. Math. Ver.* **1936**, *46*, 26–32.
- (59) Wehle, D.; Fitjer, L. Heptacyclo[19.3.0.0.1^{5,0}.5^{9,0}.9^{13,0}.13^{17,0}.17²¹]-tetracosan ([6.5]Coronan). *Angew. Chem.* **1987**, *99*(2), 135–137.
- (60) Whitney, H. Congruent graphs and the connectivity of graphs. *Am. J. Math.* **1932**, *54*, 150–168.
- (61) Whitney, H. 2-isomorphic graphs. *Am. Math. J.* **1933**, *55*, 245–254.
- (62) Whitney, H. On abstract properties of linear dependence. *Am. J. Math.* **1935**, *57*, 509–533.
- (63) Wiberg, K. B.; Waddell, S. T. Reactions of [1.1.1]propellane. *J. Am. Chem. Soc.* **1990**, *112*, 2194–2216.
- (64) Willet, P.; Barnard, G. J. M.; Downs Similarity searching. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 983–996.