Paths and walks in acyclic structures: plerographs *versus* kenographs

Damir Vukičević,*[a] Ante Miličević,[b] Sonja Nikolić,[b] Jelena Sedlar,[c] and Nenad Trinajstić[b]

[a] Department of Mathematics, The University of Split, Teslina 12, HR-21000 Split, Croatia
[b] The Rugjer Bošković Institute, P.O.B. 180, HR-10002 Zagreb, Croatia
[c] Faculty of Civil Engineering, The University of Split, HR-21000 Split, Croatia
E-mail: <u>vukicevi@pmfst.hr</u>

Abstract

The relationship between the plerographs and kenographs (called by Cayley plerograms and kenograms) representing acyclic structures, exemplified by octanes, is studied *via* paths and walks. It is found that the relationship for the net numbers of walks is approximately linear. This result in conjuction with several related reports in the literature supports the exclusive use of kenographs for acyclic structures in modern chemical graph theory.

Keywords: Chemical graphs, kenogram, kenopgraph, path, plerogram, plerograph, tree, walk, octanes

Introduction

Arthur Cayley (1821-1895), a distinguished 19th century English mathematician, considered in his paper on the mathematical theory of isomers (published in 1847) two types of molecular graphs that he named *plerograms* and *kenograms*.[1] In modern chemical graph theory, the plerograms (*P*) are molecular graphs in which *all* atoms are represented by vertices whilst the kenograms (*K*) are referred to as a *hydrogen-suppressed* or *hydrogen-depleted* molecular graphs.[2] Gutman and Polansky in their book *Mathematical Concepts in Organic Chemistry*[3] used the terms *complete molecular graphs* and *skeleton graphs* for plerograms and kenograms, respectively.

We use here the terms *plerographs* and *kenographs* for plerograms and kenograms, respectively. We adopted these terms because plerograms and kenograms are *graphs* rather than types of *diagrams*, and we also wanted to preserve the roots of Cayley's terms. Therefore, we substituted *grams* in Cayley's terms with *graphs*. Cayley could not do this because the name

graph was not yet adopted in 1847. This happened only after the one-page paper by Sylvester[4] appeared in *Nature* (in 1877) in which he introduced the term *graph*, stimulated as he stated by *chemicographs* in *Lectures notes for chemical students* (London, 1866) by Edward Frankland (1825-1899). Frankland used in his *Notes* the term *graphic-like symbolic formulae*.

The idea to of finding invariants for plerographs from the corresponding kenographs was investigated recently by several authors. In particular, Gutman *et al.*[5] has expressed the Wiener index[6] $W(P)$ in terms of $W(K)$ and the Harary index[7,8] $H(P)$ in terms of $H(K)$ for alkane isomers. They found the relationships between $W(P)$ and $W(K)$, and $H(P)$ and $H(K)$ to be linear. Bonchev *et al.*[9,10] derived Wiener indices for plerographs that they called *thorny graphs*. The name thorny graphs or *t-thorn graphs* for plerographs was introduced by Gutman.[11]

Here we report the computation of paths and walks for plerographs and kenographs representing alkanes. In Figure 1, we give as examples these two types of molecular graphs for 2,3,3-trimethylpentane. Molecular graphs will be depicted in a usual way– atoms will be replaced by vertices and bonds by edges.[12] In alkane graphs, called also molecular trees, the maximal vertex-degree is four.

From Figure 1 we see that the kenograph is much simpler structure than the plerograph. This is a probable reason why the practitioners of chemical graph theory opted for kenographs in their work. Nevertheless, it is of interest to investigate the relationship between plerographs and kenographs.
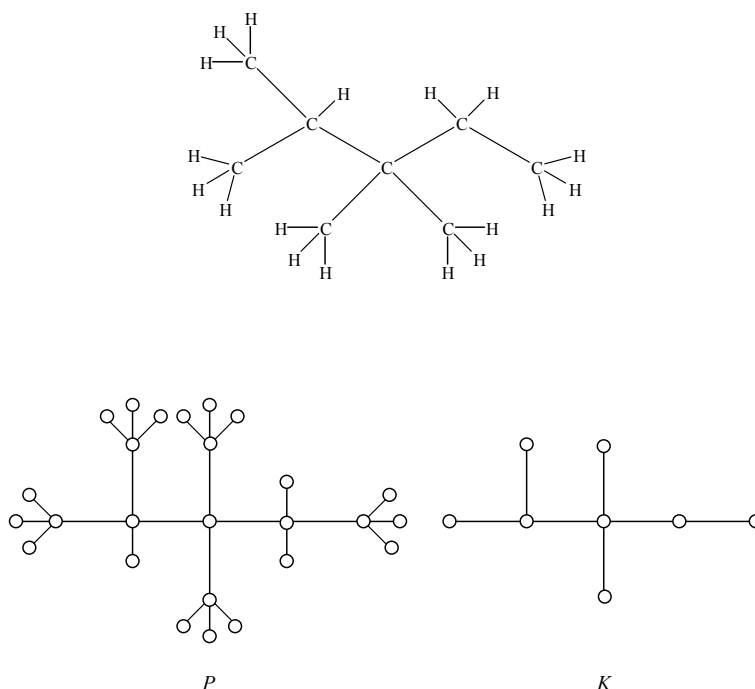

Figure 1. Two graph-theoretical representations of 2,3,3-trimethylpentane: plerograph *P* and kenograph *K*. The plerograph has 3 *N* + 2 vertices while the kenograph has *N* vertices. In this case N is 8.

Basic Graph-Theoretical Concepts[13,14]

An *acyclic graph* is a graph that has no cycles. A *tree* is a connected acyclic graph. A *leaf* is the end-vertex. A *path p* is any subtree in which all interior vertices have valence 2. *Length of a path* is the number of edges in the path. A *walk w* is an alternating sequence of vertices and edges, with each edge being incident to the vertices immediately preceding and succeeding it in the sequence. *Length of a walk* is the number of edges in the walk. For the sake of simplicity, paths (walks) of the length i shall be called *i-paths* (*i-walks*).

For a set of vertices *X,* we use *G(X)* to denote the *induced subgraph* of *G* whose vertex set is *X* and whose edge set is the subset of *E(G)* consisting of those edges with both ends in *X*. A subgraph is *induced by a walk* if it is induced by set of vertices of the walk. A walk *W* in graph *G* is *complete* if *G(W)=G*.

Let $p_{i,P}$ ($p_{i,K}$) be the number of different *i*-paths in a plerograph (kenograph), and $w_{i,P}$ ($w_{i,K}$) the number of different *i*-walks in a plerograph (kenograph). The dependency of $p_{i,P}$ and $w_{i,P}$ ($1 \le i \le 5$) on the number and valency of vertices in kenograph is determined in this paper.

Paths

In the case of alkanes, all vertices in the related kenographs correspond to carbon atoms, while in plerographs the mono-valent vertices correspond to hydrogen atoms while four-valent vertices correspond to carbon atoms. So, let h be the number of vertices of degree one in the plerograph and c the number of four-valent vertices in the plerograph. Obviously, the total number of vertices in plerograph is

$$v_P = c + h \tag{1}$$

In the plerograph, 1-paths correspond to edges, so we have

$$p_{1,P} = e_P = v_P - 1 \tag{2}$$

since the plerograph of any alkane is a tree. Note that

$$e_P = (4c + h)/2 \tag{3}$$

and, since the plerograph is also a tree, it follows that

$$e_P = v_P - 1 = c + h - 1 \tag{4}$$

Thus, we have

$$(4c + h)/2 = c + h - 1 \tag{5}$$

$$h = 2\,c + 2 \tag{6}$$

and, therefore,

$$v_P = c + h = 3\,c + 2 = 3\,v_K + 2 \tag{7}$$

where v_K is the total number of vertices in kenograph. Finally, we get

$$p_{1,P} = 3\,v_K + 1 \tag{8}$$

If we consider i-paths in the plerograph, $i \geq 2$, we see that leaves of any such path can correspond to both, hydrogen and carbon atoms, but interior vertices have to correspond to carbon atoms since they are not mono-valent. For any path p let the *interior of the path,* denoted by $int(p)$, be a subgraph of the path induced in a graph by the interior vertices of the path. Thus, we can conclude that the interior of any path in the plerograph is a subgraph of the kenograph. If the interior of the path (of length ≥ 3) in the kenograph is fixed, each of the terminal leaves of p can be chosen in 3 different ways, so there are 9 different paths in the plerograph with that interior. The number $p_{i,P}$ will be determined if we count how many subgraphs of the kenograph can be the interior of an i-path in the plerograph.

The interior of any 2-path in the plerograph is a single vertex in the kenograph, so for the given vertex the path in the plerograph can be chosen in $\binom{4}{2} = 6$ different ways. Therefore,

$$p_{2,P} = 6\,v_K \tag{9}$$

The interior of any 3-path in the plerograph is an edge in the kenograph and can be chosen in e_K different ways (*i.e.*, v_K - 1 since the kenograph is a tree). We get

$$p_{3,P} = 9\,(v_K - 1) = 9\,v_K - 9 \tag{10}$$

The interior of any 4-path in the plerograph is a 2-path in the kenograph, therefore,

$$p_{4,P} = 9\,p_{2,K} \tag{11}$$

Now, consider any 2-path in the kenograph. There are v_K possible ways to choose u (where u is the interior vertex of the path), and for u chosen there are $\binom{d_K(u)}{2}$ different ways to choose leaves. Thus, we have

$$p_{2,K} = \sum_{u \in V(K)} \binom{d_K(u)}{2} = \sum_{u \in V(K)} \frac{d_K(u)^2 - d_K(u)}{2} = \frac{1}{2} M_1(G) - e_K \qquad (12)$$

where M_1 is the first Zagreb index, defined as[15-18]

$$M_1(G) = \sum_{u \in V(G)} [d_G(u)]^2 \qquad (13)$$

where $d_G(u)$ is the degree (=the number of the first neighbors) of a vertex u. Finally,

$$p_{4,P} = 9\, p_{2,K} = 9\, M_1(K)/2 - 9\, v_K + 9 \qquad (14)$$

The interior of any 5-path in the plerograph is a 3-path in the kenograph, so

$$p_{5,P} = 9\, p_{3,K} \qquad (15)$$

Considering 3-paths in the kenograph, let u and v be the interior vertices of any such path. Any member of $E(K)$ can be uv and if we fix that edge, the leaf adjacent to u can be chosen in $d_K(u) - 1$ different ways and the leaf adjacent to v can be chosen in $d_K(v) - 1$ different ways. Therefore,

$$
\begin{aligned}
p_{3,K} &= \sum_{uv \in E(K)} \big(d_K(u) - 1\big)\big(d_K(v) - 1\big) = \\
&= \sum_{uv \in E(K)} \big(d_K(u)d_K(v) - d_K(u) - d_K(v) + 1\big) = \\
&= \sum_{uv \in E(K)} d_K(u)d_K(v) - \sum_{uv \in E(K)} \big(d_K(u) + d_K(v)\big) + \sum_{uv \in E(K)} 1 = \\
&= M_2(K) - \sum_{uv \in E(K)} \big(d_K(u) + d_K(v)\big) + e_K = \\
&= M_2(K) - \sum_{u \in V(K)} d_K(u)^2 + v_K - 1 = \\
&= M_2(K) - M_1(K) + v_K - 1.
\end{aligned} \qquad (16)
$$

where M_2 is the first Zagreb index, defined as[15,16,18,19]

$$M_2(G) = \sum_{u,v \in E(G)} d_G(u)\, d_G(v) \qquad (17)$$

Finally, we get

$$p_{5,P} = 9\, p_{3,K} = 9\, M_2(K) - 9\, M_1(K) + 9\, v_K - 9 \qquad (18)$$

Walks

Every 1-walk in the plerograph induces an edge in the plerograph. For each edge in the plerograph there are two different 1-walks that induce that edge. So, we have

$$w_{1,P} = 2\, e_P = 2\,(v_P - 1) = 2\,(3\, v_K + 2 - 1) = 6\, v_K + 2 \tag{19}$$

Let us consider 2-walks. We shall divide all such walks in two groups: Walks that induce 1-paths in the plerograph and walks that induce 2-paths in the plerograph.

For each 1-path (*i.e.*, edge uv) in the plerograph there are two complete 2-walks on that path: One starting at vertex u and other starting at vertex v. Therefore, there are $2\, p_{1,P}$ 2-walks in the plerograph that induce 1-paths in the plerograph. For each 2-path in the plerograph, there are two complete 2-walks on that path: One starting at one leaf and the other walk starting at another leaf. Therefore there are $2\, p_{2,P}$ 2-walks in the plerograph that induce 2-paths in the plerograph.

Since every 2-walk induces a 1-path or 2-path and no 2-walk induces both, *i.e.*, the above division in groups partitions the set of all 2-walks in the plerograph, we can conclude

$$w_{2,P} = 2\, p_{1,P} + 2\, p_{2,P} \tag{20}$$

All 3-walks of any plerograph can be divided into three disjoint sets of walks:
1. walks that induce 1-paths in the plerograph
2. walks that induce 2-paths in the plerograph
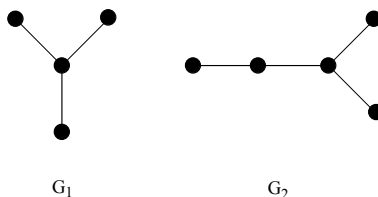3. walks that induce 3-paths in the plerograph.

Analogous to 2-walks we can conclude that there are $2p_{1,P}$ 3-walks in the plerograph that induce 1-paths in the plerograph and $2p_{3,P}$ 3-walks in the plerograph that induce 3-paths in the plerograph.

For each 2-path in the plerograph there are 4 different complete 3-walks on that path: One starting at each leaf and two starting at each interior vertex. Therefore there are $4\, p_{2,P}$ 3-walks in the plerograph that induce 2-paths in the plerograph.

Since, the above division partitions the set of all 3-walks in plerograph (*i.e.*, every 3-walk is in one group and no 3-walk is in two groups) we can conclude

$$w_{3,P} = 2\, p_{1,P} + 4\, p_{2,P} + 2\, p_{3,P} \tag{21}$$

For 4-walks and 5-walks we need to consider the following graphs:

G_1 G_2

All 4-walks of the plerograph can be divided into 5 disjoint sets of walks:

1. walks that induce 1-paths in the plerograph
2. walks that induce 2-paths in the plerograph
3. walks that induce 3-paths in the plerograph
4. walks that induce 4-paths in the plerograph
5. walks that induce subgraphs of the plerograph isomorphic to G_1.

Analogous to 2-walks we can conclude that there are $2p_{1,P}$ 4-walks in the plerograph that induce 1-paths in the plerograph and $2p_{4,P}$ 4-walks in the plerograph that induce 4-paths in the plerograph.

For each 2-path in the plerograph there are 8 different complete 4-walks on that path: Three starting at each leaf and two starting at each interior vertex. Therefore there are $8\,p_{2,P}$ 4-walks in the plerograph that induce 2-paths in the plerograph.

For each 3-path in the plerograph there are 4 different complete 4-walks on that path: One starting at each leaf and one starting at each of the interior vertices. Therefore there are $4\,p_{3,P}$ 4-walks in the plerograph that induce 3-paths in the plerograph.

For each subgraph of the plerograph that is isomorphic to G_1, there are 6 different complete 4-walks on that subgraph: Two starting at each leaf. Here, we also have to determine how many such subgraphs of the plerograph there are. An interior vertex of G_1 can be any vertex in kenograph and for the interior vertex chosen, we choose leaves in the plerograph, which can be done in $\binom{4}{3} = 4$ different ways. So, there are $4v_K$ subgraphs of the plerograph isomorphic to G_1.

Therefore, there are $24v_K$ different 4-walks in the plerograph that induce subgraphs of the plerograph isomorphic to G_1.

Since the above division partitions the set of all 4-walks in the plerograph, we have

$$w_{4,P} = 2\,p_{1,P} + 8\,p_{2,P} + 4\,p_{3,P} + 2\,p_{4,P} + 24\,v_K \qquad (22)$$

Finally, all 5-walks of the plerograph can be divided into 7 disjoint sets of walks:

1. walks that induce 1-paths in the plerograph
2. walks that induce 2-paths in the plerograph
3. walks that induce 3-paths in the plerograph
4. walks that induce 4-paths in the plerograph
5. walks that induce 5-paths in the plerograph
6. walks that induce subgraphs of the plerograph isomorphic to G_1
7. walks that induce subgraphs of the plerograph isomorphic to G_2.

Analogous to 2-walks we can conclude that there are $2p_{1,P}$ 5-walks in the plerograph that induce 1-paths in the plerograph and $2p_{5,P}$ 5-walks in the plerograph that induce 5-paths in the plerograph.

For each 2-path in plerograph there are 12 different complete 5-walks on that path: Three starting at each leaf and six starting at interior vertex. Therefore, there are $12p_{2,P}$ 5-walks in the plerograph that induce 2-paths in the plerograph.

For each 3-path in the plerograph there are 12 different complete 5-walks on that path: Four starting at each leaf and two starting at each interior vertex. Therefore, there are $12p_{3,P}$ 5-walks in plerograph that induce 3-paths in the plerograph.

For each 4-path in the plerograph there are 4 different complete 5-walks on that path: One starting at each leaf and one starting at each interior vertex adjacent to a leaf. Therefore, there are $4p_{4,P}$ 5-walks in the plerograph that induce 4-paths in the plerograph.

For each subgraph of the plerograph that is isomorphic to G_1, there are 12 different complete 5-walks on that subgraph: Two starting at each leaf and 6 starting at an interior vertex. We already have determined that there are $4v_K$ subgraphs of the plerograph isomorphic to G_1. Therefore, there are $12 \times 4v_K = 48v_K$ different 5-walks in the plerograph that induce subgraphs of the plerograph isomorphic to G_1.

For each subgraph of the plerograph that is isomorphic to G_2, there are 4 different complete 5-walks on that subgraph. Let vertices of G_2 be labeled as on the following picture.

G_2
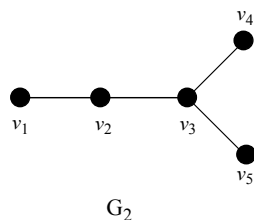
Then two of those walks start at v_1, one at v_4 and one at v_5. Now, as to the number of subgraphs of the plerograph isomorphic to G_2, any edge of the kenograph can be chosen as v_2v_3 and for that edge chosen leaf v_1 can be chosen in 3 different ways and pair of leaves v_4 and v_5 can be chosen in $\binom{4}{2}=6$ different ways. Therefore, there are $18e_K = 18(v_K - 1) = 18v_K - 18$ subgraphs of the plerograph isomorphic to G_2. Finally, we get that number of 5-walks that induce subgraphs of the plerograph isomorphic to G_2 is equal to $4(18v_K - 18) = 72v_K - 72$.

Since the above division partitions the set of all 5-walks in the plerograph, we have

$$w_{5,P} = 2\,p_{1,P} + 12\,p_{2,P} + 12\,p_{3,P} + 4\,p_{4,P} + 2\,p_{5,P} + 48\,v_K + 72\,v_K - 72 \qquad (23)$$

Correlation between walks in plerographs and kenographs
We correlated the number of walks in plerographs against the number of walks in kenographs of 18 octanes. Octanes were selected following Randić's advice:[20,21] He recommended octanes as a test set because it consists of only 18 isomers that possess structural properties that are also present in other alkanes.

In Table 1 we report the number of walks with up to the length 7 and in Figure 2 we give the scatter-plot between the number of said walks in octane kenographs and plerographs.

The number of walks was obtained by the matrix multiplication since it is well-established that the number of walks of length λ beginning at vertex i and ending at vertex j is given by the element $(\mathbf{A}^{\lambda})_{ij}$, the ij-element, in the λ-th power of the adjacency matrix $\mathbf{A}$.[22-24] The computation was carried out by means of a $C++$ program.

Table 1. The number of walks with up to the length 7 for kenographs and plerographs representing isomeric octanes

| Octane	Number of walks of length at most 7 in the kenograph	Number of walks of length at most 7 in the plerograph
n-Octane	1254	29006
2-Methylheptane	1528	29834
3-Methylheptane	1676	30176
4-Methylheptane	1712	30212
3-Ethylhexane	1856	30554
2,2-Dimethylhexane	2284	31490
2,3-Dimethylhexane	2136	31328
2,4-Dimethylhexane	1994	31040
2,5-Dimethylhexane	1822	30680
3,3-Dimethylhexane	2602	32138
3,4-Dimethylhexane	2272	31652
3-Ethyl-2-methylpentane	2304	31688
3-Ethyl-3-methylpentane	2882	32750
2,2,3-Trimethylpentane	3072	33236
2,2,4-Trimethylpentane	2634	32372
2,3,3-Trimethylpentane	3218	33524
2,3,4-Trimethylpentane	2592	32462
2,2,3,3-Tetramethylbutane	4094	35072
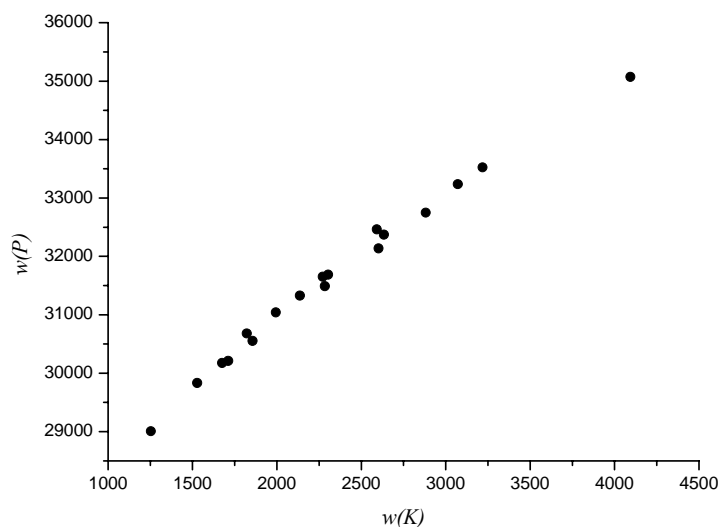
Figure 2. The scatter-plot between the number of walks in the plerographs $w(P)$ and the number of walks in the kenographs $w(K)$ representing 18 octanes.

The obtained relationship between the number of walks in octane-kenographs $w(K)$ and the number of walks in octane-plerographs $w(P)$ is approximately linear:

$$w(P) = 26678.9 + 2.1\ w(K) \tag{24}$$

with the correlation coefficient $r=0.995$ and the standard deviation of the linear correlation $s=160$. Therefore, if the walks are used, for example, in the QSPR or QSAR[25,26] modeling of acyclic structures, it appears that it is sufficient to use kenographs for representing molecules under study. The use of plerographs may possibly bring slight new insights, but the computation will be more time-consuming since plerographs are more complex structures than kenographs as shown by the computed walks for octanes in Table 1. However, it should be pointed out that the plerographs correspond more closely to molecular structures than the kenographs.

Conclusions

In the present work, we reported the relationship between paths and walks in kenographs and plerographs of acyclic molecules, exemplified in this case by octanes. It appears that the relationship between these two representations of octanes is nearly linear. Thus, the kenographs appear as a satisfactory graph-theoretical representation for computing paths and walks of alkanes that can be of use, for example, in QSAR or QSPR. Therefore, until it is found the case when the kenographs and plerographs deviate considerably from the linear correlation, the use of

simpler kenographs is justified. So let the search for such a case continue. It should also be especially interesting to study the comparison between plerographs and kenographs for hetero-systems.

Acknowledgements

References

1. Cayley, A. *Phil. Mag.* **1847**, *47*, 444.
2. Balaban, A. T.; Ivanciuc, O. In: Devillers, J.; Balaban, A. T. Eds; *Topological Indices and Related Descriptors in QSAR and QSPR*, Gordon and Breach, Amsterdam, 1999, pp 21-57.
3. Gutman, I.; Polansky, O. E. *Mathematical Concepts in Organic Chemistry*, Springer-Verlag: Berlin, 1986, p 19.
4. Sylvester, J. J. *Nature* **1877**, *17*, 284.
5. Gutman, I.; Vidović, D.; Petrović, L. *J. Chem. Soc. Faraday Trans*.**1997**, *94*, 857.
6. Wiener, H. *J. Am. Chem. Soc.* **1947**, *69*, 17.
7. Plavšić, D.; Nikolić, S.; Trinajstić, N.; Mihalić, Z. *J. Math. Chem.* **1993**, *12*, 235.
8. Ivanciuc, O.; Balaban, T.-S.; Balaban, A.T. *J. Math. Chem.* **1993**, *12*, 309.
9. Baytautas, L.; Bonchev, D.; Klein, D. J. *MATCH – Commun. Math. Comput. Chem.* **2001**, *44,* 31.
10. Bonchev, D.; Klein, D. J. *Croat. Chem. Acta* **2002**, *75*, 613.
11. Gutman, I. *Publ. Inst. Math.* **1998**, *63*, 31.
12. Trinajstić, N. *Chemical Graph Theory*, 2nd revised Edn; CRC Press: Boca Raton, FL, 1992.
13. Harary, F. *Graph Theory*, 2nd printing, Addison-Wesley: Reading, MA, 1971.
14. Veljan, D. *Combinatorial and Discrete Mathematics*, Algorithm, Zagreb, 2001.
15. Gutman, I.; Trinajstić, N. *Chem. Phys. Lett.* **1972**, *17*, 535.
16. Nikolić, S.; Kovačević, G.; Miličević, A.; Trinajstić, N. *Croat. Chem. Acta* **2003**, *76*, 113.
17. Gutman, I.; Das, K. C. *MATCH – Commun. Math. Comput. Chem.* **2004**, *50*, 83.
18. Zhou, B. *MATCH – Commun. Math. Comput. Chem.* **2004**, *52*,113.
19. Das, K. C.; Gutman, I. *MATCH – Commun. Math. Comput. Chem.* **2004**, *52*,103.
20. Randić, M. *Croat. Chem. Acta* **1993**, *66*, 289.
21. Randić, M. *J. Math. Chem.* **1990**, *4*, 157.
22. Rücker, G.; Rücker, C. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 683.

23. Nikolić, S.; Trinajstić, N.; Tolić, I. M.; Rücker, G.; Rücker, C. In: Bonchev, D.; Rouvray, D. H. Eds; *Complexity in Chemistry – Introduction and Fundamentals*, Taylor & Francis, Lon Rücker don: 2003; pp 29-89.

24. Babić, D.; Klein, D. J.; von Knop, J.; Trinajstić, N. In: Hincheliffe, A. Ed; *Chemical Modelling – Applications and Theory*, The Royal Society of Chemistry, Cambridge: 2004, pp 126-170.

25. Devillers, J.; Balaban, A. T., Eds; *Topological Indices and Related Descriptors in QSAR and QSPR*, Gordon and Breach: Amsterdam, 1999.

26. Todeschini, R.; Consonni, V. *Handbook of Molecular Descriptors*, Wiley-VCH: Weinheim, 2000.