

Molecular Complexity of Certain Homologous Series of Condensed Benzenoid Hydrocarbons: Limiting Values of the Patency Index and the Index of Spanning-Tree Density

Roger B. Mallion,^{1,*} Paul Pollak,² Paweł J. Skrzyński³

¹ School of Physical Sciences, University of Kent, Canterbury CT2 7NH, England, United Kingdom

² The King's School, Canterbury CT1 2ES, England, United Kingdom

³ The Canterbury High School, Canterbury CT2 8QA, England, United Kingdom

* Corresponding author's e-mail address: R.B.Mallion@kent.ac.uk

RECEIVED: August 30, 2017 * REVISED: November 16, 2017 * ACCEPTED: November 16, 2017

Abstract: Two indices of molecular complexity, the Patency Index (2017) and the Spanning-Tree Density (2003), are applied to three homologous series of condensed benzenoid hydrocarbons. Calculation of the Spanning-Tree Density requires finding the number of spanning trees in a given molecular graph, which may be achieved by applying the Cycle Theorem (2004) or, in the case of planar graphs in a planar embedding, the theorem of Gutman, Mallion & Essam (1983). To compute the Patency Index, it is necessary to count the number of *ladders* in the embedded molecular graph. This is done by means of the Dual Cycle Theorem (2017). In the latter, a *ladder* is conceived of as the edge-set relating to faces as the edge-set of a spanning tree relates to vertices. For a planar graph in a planar embedding, the number of ladders is equal to the number of spanning trees. The three homologous series investigated here are the Linear $[n]$ -Acenes (A_n), the $[n]$ -Phenacenes (Ph_n) and the $[n]$ -Helicenes (H_n) (the latter of which are geometrically non-planar but graph-theoretically planar). For these three series, the Spanning-Tree Densities and the Patency Indices may be obtained in closed form and so their behaviour as $n \rightarrow \infty$ is easily examined. Though neither index distinguishes between individual members of the three series (A_n , Ph_n and H_n) for a specified value of n , this does not preclude the possibility that, *within* each series, either index may exhibit correlation with physical or chemical data.

Keywords: molecular complexity, spanning trees, ladders of embedded graphs, patency index, spanning-tree density.

INTRODUCTION

HOW *INTRICATE*^[1] a molecular structure is has, in recent years, frequently been called the *complexity* of the molecule; for reviews, see, for example, Ref. [1] (and the Refs. [1–7,11–13,19,22,24–26] that are cited in Ref. [1]), Ref. [2] (and Refs. [11–24] cited there) and see also Refs. [3–6] of the present paper. This appellation is perhaps unfortunate because much of this work involves the application of Graph Theory and the term ‘complexity’ was already long used by graph-theorists (*e.g.*, Ref. [7]) to denote the number of spanning trees in a graph. (Recall that a spanning tree — necessarily of a one-component graph G — is a connected acyclic sub-graph that contains all

the vertices of G .) Deliberately avoiding the word ‘complexity’ — and introducing instead the word ‘intricacy’ — Trinajstić and one of the present authors (RBM) defined^[1] what they called the *Spanning-Tree Density* of a molecular graph ($STD(G)$) as the ratio

$$STD(G) = \frac{t(G)}{{}^e C_{(v-1)}}, \quad (1)$$

where $t(G)$ is the number of spanning trees in the molecular graph (G), which is taken to have e edges and v vertices. (${}^e C_{(v-1)}$, often also written as

$$\binom{e}{v-1},$$

is the number of ways in which $(v-1)$ distinct objects can be selected from e unlike objects.) $STD(G)$ thus effectively represents the *probability* that, if precisely $(v-1)$ edges are removed at random from the graph, that process results in the creation of a spanning tree of the graph. $t(G)$ may be computed in many ways (for reviews see, for example, Refs. [9–14]) but the present authors^[2] have frequently applied what Kirby *et al.* called the Cycle Theorem^[15,16] for suitably embedded general one-component molecular-graphs or the theorem of Gutman, Mallion & Essam^[17] for the special case of planar graphs in a planar embedding. In an attempt to relate $STD(G)$ for a molecular graph to the ‘complexity’ of that molecule, Mallion & Trinajstić^[1] considered the *reciprocal* of the Spanning-Tree Density and postulated that the greater the ‘Reciprocal Spanning-tree Density’ of a given molecular graph, the more *intricate* (in the sense of being ‘complicated’) that molecular structure should be considered to be.

Very recently, Kirby and the present authors have offered (in this journal)^[2] a Dual of the Cycle Theorem, for a graph embedded on a closed surface, in which the entities counted are not spanning trees^[8] but, rather, what might be regarded as ‘duals’ of the edge-sets of spanning-trees. These we define in Ref. [2] as *ladders*. An index for a molecular graph G , called by Kirby *et al.*^[2] the *Patency Index* ($P(G)$) of G , was there defined as follows:

$$P(G) = \frac{l(G)}{{}^e C_{(f-1)}}, \quad (2)$$

where $l(G)$ is the number of ladders^[2] in G , which has e edges, v vertices and f faces. The method of calculating $l(G)$ is described in Ref. [2]. It was shown there^[2] that, for a planar graph on a ‘planar’ embedding (*i.e.*, one not using *generic* circuits, defined in Ref. [15], even if the surface is capable of sustaining such),

$$P(G) = STD(G). \quad (3)$$

In this note we apply these two indices to some control sets of structures which constitute well-known homologous series of condensed benzenoid hydrocarbons^[18,19] — in the first instance, we consider the Linear $[n]$ -Acenes — and we show that, for these series, it is possible to solve for $P(G)$ and $STD(G)$ in *closed form* and we are, thereby, able to examine the limiting behaviour of these two indices as $n \rightarrow \infty$. We label the molecular graphs of this first series of conjugated systems, the Linear Acenes, as A_n (with A_1 standing for Benzene, A_2 for Naphthalene, *etc.*) Figure 1 shows A_4 (Tetracene) embedded on the surface of a sphere (denoted by the red circle — please see Ref. [2] for details of the implications of this notation.) Consideration is then extended to two analogous homologous series (the $[n]$ -Phenacenes and the $[n]$ -Helicenes — illustrated later),

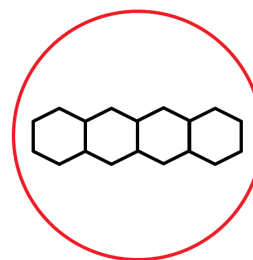


Figure 1. A_n embedded on the surface of a sphere (with, here, $n = 4$). The red circle has the significance that was explained in Ref. [2] — effectively, it makes this representation of the sphere topologically faithful.

as well as to some more-compact tessellations of n hexagons that constitute the molecular graphs of other condensed, benzenoid hydrocarbons.

ALGEBRAICAL ARGUMENTS TO OBTAIN EXPRESSIONS FOR $STD(A_n)$ AND $P(A_n)$ IN CLOSED FORM

The sub-determinant within the face-to-face incidence-matrix^[2,20] for A_4 (say, $f(4)$), obtained by retaining only the entries relating to the hexagonal faces — or, alternatively, the corresponding determinant obtained by an application of the theorem of Gutman *et al.*^[17] — is

$$f(4) = \det \begin{pmatrix} 6 & -1 & 0 & 0 \\ -1 & 6 & -1 & 0 \\ 0 & -1 & 6 & -1 \\ 0 & 0 & -1 & 6 \end{pmatrix} = 1189,$$

and equals the number of spanning trees (and also the number of ladders^[2]) in A_4 . It is not difficult to see that

$$f(n) = 6f(n-1) - f(n-2), \quad (4)$$

and also that $f(1) = 6$ and $f(2) = 35$. The recurrence relation (4), with these initial conditions, can be solved in the usual way to give

$$f(n) = \frac{(3 + \sqrt{8})^{n+1}}{\sqrt{32}} - \frac{(3 - \sqrt{8})^{n+1}}{\sqrt{32}} \quad (5)$$

APPROXIMATIONS AND NUMERICAL CALCULATIONS

As $f(n)$ is an integer the expression in equation (5) can be replaced by

$$f(n) = \text{the integral part of } \frac{(3 + \sqrt{8})^{n+1}}{\sqrt{32}},$$

i.e.,

$$f(n) = \left[\frac{(3 + \sqrt{8})^{n+1}}{\sqrt{32}} \right]. \quad (6)$$

As a check on this we may note that

$$f(3) = \det \begin{pmatrix} 6 & -1 & 0 \\ -1 & 6 & -1 \\ 0 & -1 & 6 \end{pmatrix} = 204,$$

and, using equation (6), we find that

$$f(3) = \text{the integral part of } \frac{(3 + \sqrt{8})^4}{\sqrt{32}} = [204.00015] = 204.$$

It is noteworthy that, even for a small value of n , the decimal part of

$$\frac{(3 + \sqrt{8})^{n+1}}{\sqrt{32}}$$

is very small. If the Patency Index for A_n is evaluated from

Table 1. Approximate values of $P(A_n)$ (to 4 decimal places), calculated from equation (7), for Linear Acenes A_1 – A_{20}

n	$P(A_n)$ (to 4 decimal places)
1	1.0000
2	0.6364
3	0.3643
4	0.1987
5	0.1054
6	0.0549
7	0.0282
8	0.0144
9	0.0073
10	0.0036
11	0.0018
12	0.0009
13	0.0005
14	0.0002
15	0.0001
16	0.0001
17	0.0000
18	0.0000
19	0.0000
20	0.0000

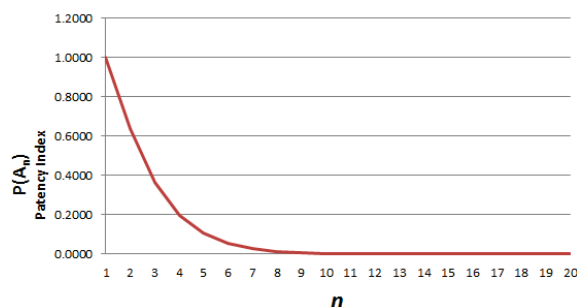


Figure 2. Plot of $P(A_n)$ (along the ordinate) vs. n (along the abscissa) for Linear Acenes A_1 – A_{20} .

$$P(A_n) = \frac{f(n)}{{}_{5n+1}C_n}, \quad (7)$$

then sufficient accuracy for $P(A_n)$ is obtainable simply by using

$$\frac{(3 + \sqrt{8})^{n+1}}{\sqrt{32}}$$

for $f(n)$, without the need for the second (subtractive) term in equation (5), or for the process of extracting the ‘integral-part-of’, as is done in equation (6), above. Table 1 shows (correct to 4 decimal places) values of A_n , so calculated, for several n in the range $1 \leq n \leq 20$. It will be seen from the Table that, by about $n = 17$, $P(A_n)$ is effectively zero. (In a chemical context, however, this is in fact irrelevant because stable Linear Acenes do not exist beyond $n = 5$ (Pentacene)^[21,22].) It will be observed that $P(A_n)$ is roughly halved as one goes along the series — that is, each time n is increased by 1. Between $n = 1$ and about $n = 15$, $P(A_n)$ thus can be said approximately to exhibit an *exponential decay*, as is illustrated in Figure 2.

EXTENSION TO OTHER HOMOLOGOUS BENZENOID SERIES

Such an exponential decrease is not observed in the case of some physical or chemical properties,^[23] such as melting points and boiling points, in the first five (chemically stable) members of the Linear-Acene series, A_n . (Inverse) correlation with such properties might possibly have been hoped for if the indices of Patency and Spanning-Tree Density are in some way quantifying ‘molecular complexity’. However, this lack of correlation in the Linear Acene series just referred to is, perhaps, only to be expected when it is realised that successive annellation to form the Linear-Acene series (A_n) is not the only way in which benzenoid rings may be linearly tessellated: in addition to the [n]-Acenes, there are also

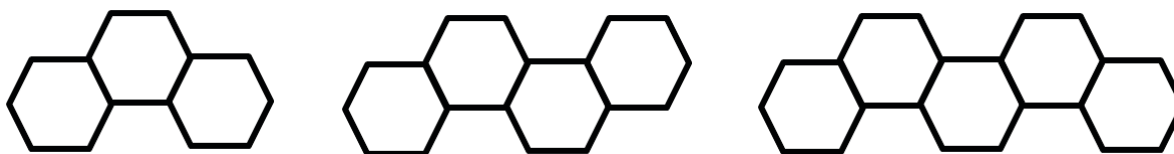


Figure 3. The molecular graphs of the first three members of the $[n]$ -Phenacene series (Ph_n , with $n = 3, 4$ and 5).

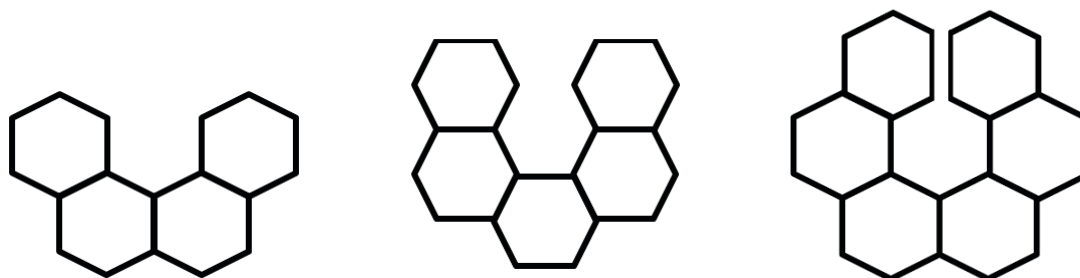


Figure 4. The molecular graphs of the first three members of the $[n]$ -Helicene series (H_n , with $n = 4, 5$ and 6).

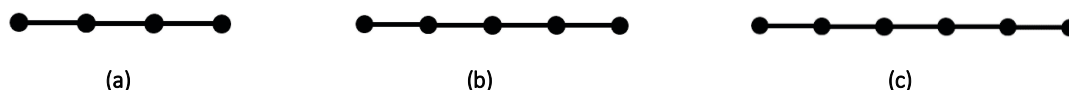


Figure 5. The inner duals [17, 29] of the molecular graphs (when embedded on a sphere or a plane) of (a) A_4 , Ph_4 and H_4 ; (b) A_5 , Ph_5 and H_5 ; (c) A_6 , Ph_6 and H_6 .

- (a) the $[n]$ -Phenacene series (Ph_n) — $n > 2$ — in which successive addition of benzenoid rings is effected in a zig-zag fashion to yield the structures of Phenanthrene ([3]-Phenacene), Chrysene ([4]-Phenacene), Picene ([5]-Phenacene), *etc.*, as in Figure 3, and
- (b) the $[n]$ -Helicene series (H_n) — $n > 3$ — in which rings overlap physically in space and, helped also by their attached protons, crowd each other out so that the molecule itself is geometrically (though *not* graph-theoretically) non-planar.^[24–28] This results in structures such as [4]-Helicene (also known as 3,4 Benzophenanthrene and Tetrahelicene), [5]-Helicene (Pentahelicene), [6]-Helicene (Hexahelicene), *etc.* Some of these are illustrated in Figure 4.

However, the graphs that are the *inner-duals*[†][17,29] arising in the spanning-tree counting theorem of Gutman

et al.^[17] for the molecular graphs of the series A_n , Ph_n , and H_n (when embedded on a sphere or a plane) are in fact *the same* (for the identical values of n) — these inner duals are shown in Figure 5 for $n = 4, 5$ and 6 — and, in addition, because the number of edges and vertices in the molecular graphs A_n , Ph_n and H_n are likewise all the same (for a given value of n), it thus follows that (again, for a specified value of n)

$$STD(A_n) = STD(Ph_n) = STD(H_n).$$

Furthermore, because of the result presented in Ref. [2] ensuring that $P(G) = STD(G)$ for (graph-theoretically) planar G in a planar embedding, and, in addition, because the number of edges and faces (rings) in the molecular graphs A_n , Ph_n and H_n are again all the same (for a given value of n), it thus also follows that (once more, for a specified value of n):

$$P(A_n) = P(Ph_n) = P(H_n).$$

[†] It should be noted in passing that although (for a common value of n) the graphs that are the inner duals of the three series A_n , Ph_n and H_n (when the molecular graphs of these homologous series are embedded on a sphere or a plane) are *identical*, what Balaban & Harary^[30] called the ‘characteristic graphs’ of A_n , Ph_n and H_n are *different*. This is because^[17] the so called ‘characteristic graphs’ of benzenoid systems, as defined by Balaban & Harary in Ref. [30], are in fact *not* actually graphs at all, because they contain *geometrical* information, in addition to information about the connectivity of the benzenoid rings — which is solely what is conveyed by the inner duals of the molecular graphs of these three series, A_n , Ph_n and H_n (illustrated in Figure 5 for $n = 4, 5$ and 6).

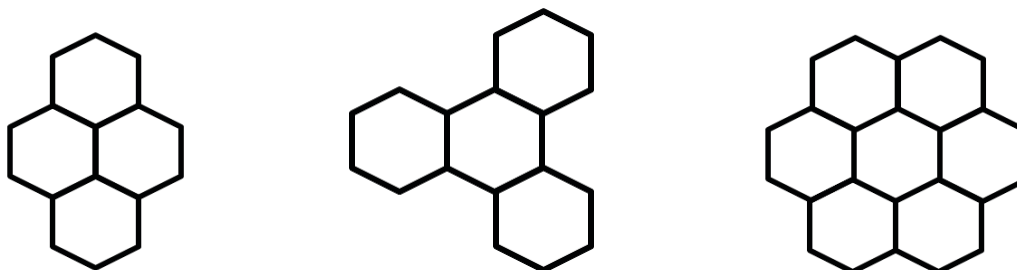


Figure 6. The molecular graphs of Pyrene (left), Triphenylene (middle) and Coronene (right).

APPLICATION TO OTHER TYPES OF CONDENSED BENZENOID SYSTEMS

Because of the identity of the inner duals^[17,29] for corresponding members of the three series A_n , Ph_n , and H_n (that is, members of these series for a specified value of n), and because, furthermore, analogous members of A_n , Ph_n , and H_n contain exactly the same number of edges, vertices and rings, the Patency Index^[2] and the Index of Spanning-Tree Density^[1] are not able to distinguish between corresponding members of the three series. This is clearly a shortcoming and it is possible that the molecular graphs of non-linear benzenoids, variously condensed — such as Pyrene, Triphenylene and Coronene (Figure 6) — would be more fruitful benzenoid systems for exploring the utility of the novel index $P(G)$ and the longer-established one, $STD(G)$, though it would be more difficult to obtain results for them in closed form.

The three homologues series A_n , Ph_n , and H_n — with identical inner duals^[17] (see Figure 5), when the molecular graphs of the series are embedded on a sphere or in a plane — do, though, have the advantage that, as shown in this note, their indices whose reciprocals are claimed to reflect the *intricacy*^[1,2] of the individual members of these three series — namely, $STD(G)$ and $P(G)$ — may be expressed *in closed form* whereas, for more randomly condensed systems such as the four-ringed benzenoids Pyrene and Triphenylene (left-hand and middle of Figure 6, respectively) or even symmetrically condensed systems such as the seven-ringed Coronene (right-hand of Figure 6), each structure would have to be considered individually. For example, for the above three systems,

$$I(\text{Pyrene}) = \det \begin{pmatrix} 6 & -1 & 0 & -1 \\ -1 & 6 & -1 & -1 \\ 0 & -1 & 6 & -1 \\ -1 & -1 & -1 & 6 \end{pmatrix} = 1092,$$

so that

$$P(\text{Pyrene}) = \frac{1092}{{}^{19}C_4} \approx 0.2817,$$

and

$$I(\text{Triphenylene}) = \det \begin{pmatrix} 6 & -1 & -1 & -1 \\ -1 & 6 & 0 & 0 \\ -1 & 0 & 6 & 0 \\ -1 & 0 & 0 & 6 \end{pmatrix} = 1188,$$

so that

$$P(\text{Triphenylene}) = \frac{1188}{{}^{21}C_4} \approx 0.1985,$$

indicating, according to the criteria established in Refs. [1] & [2], that the less-condensed Triphenylene is more *intricate* than the more-condensed Pyrene, which likewise comprises four benzenoid rings. Furthermore:

$$I(\text{Coronene}) = \det \begin{pmatrix} 6 & -1 & 0 & 0 & 0 & -1 & -1 \\ -1 & 6 & -1 & 0 & 0 & 0 & -1 \\ 0 & -1 & 6 & -1 & 0 & 0 & -1 \\ 0 & 0 & -1 & 6 & -1 & 0 & -1 \\ 0 & 0 & 0 & -1 & 6 & -1 & -1 \\ -1 & 0 & 0 & 0 & -1 & 6 & -1 \\ -1 & -1 & -1 & -1 & -1 & -1 & 6 \end{pmatrix} = 176400,$$

so that

$$P(\text{Coronene}) = \frac{176400}{{}^{30}C_7} \approx 0.0866,$$

indicating that, of the three structures in Figure 6, symmetrical, seven-ringed Coronene has the greatest intricacy.

CONCLUDING REMARKS

As already pointed out, the advantage of considering any or all of the three homologous series studied here is that, because both their Spanning-Tree Densities and Patency Indices may be expressed in closed form, we have been able to

show that both indices tend to zero (indicating increasing intricacy) as the number of benzenoid rings in each of the three series becomes infinitely large. It is clear that, for a given n , the value of both the Patency Index and of the Index of Spanning-Tree Density is the same for A_n , Ph_n and H_n . It thus follows that the value of neither index is likely to be directly related to some physical or chemical datum for those series of compounds. This, however, does not preclude the possibility that, *within* each series, either index may exhibit correlation with some such datum. This is a matter for practical investigation.

REFERENCES

- [1] R. B. Mallion, N. Trinajstić, *MATCH Commun. Math. Comput. Chem.* **2003**, *48*, 97.
- [2] E. C. Kirby, R. B. Mallion, P. Pollak, P. J. Skrzyński, *Croat. Chem. Acta*, **2017**, *90*, 75.
- [3] D. Bonchev, 'The problems of computing molecular complexity' in *Computational Chemical Graph Theory*, (Ed D. H. Rouvray), Nova Science Publishers, New York, **1990**, Chapter 2, pp. 34–63.
- [4] M. M. Waldrop, *Complexity*, Touchstone/Simon & Schuster, New York, **1992**.
- [5] B. Kaye, *Chaos and Complexity*, VCH, Weinheim, Germany, **1993**.
- [6] D. Bonchev, D. H. Rouvray, *Complexity in Chemistry, Biology and Ecology*, Mathematical and Computational Chemistry Series, (P. G. Mezey (Series Ed.)), Springer, New York, **2005**.
- [7] W. T. Tutte, *Graph Theory As I Have Known It*, Oxford Lecture Series in Mathematics and its Applications, 11. (Series Eds.: G. Ball, D. Welsh), Clarendon Press, Oxford, UK, **1998**, especially pp. 7, 26, 27, 99 & 112.
- [8] R. J. Wilson, *Introduction to Graph Theory*, First Edition, Oliver & Boyd, Edinburgh, **1972**, pp. 46–48; Fifth Edition, Prentice Hall, Upper Saddle River, New Jersey, USA, **2010**, p. 63.
- [9] J. W. Moon, *Counting Labelled Trees*, Canadian Mathematical Monographs, Ottawa, **1970**, Chapter 5, especially pp. 64–65.
- [10] D. M. Cvetković, M. Doob, H. Sachs, *Spectra of Graphs — Theory and Applications*, First Edition, Deutscher Verlag der Wissenschaften, Berlin (East), **1979**; Second Edition, Academic Press, New York, **1980**; Third Edition, Johann Ambrosius Barth, Heidelberg & Leipzig, **1995**, especially pp. 38–41 and p. 220.
- [11] H. Fleischner, *Eulerian Graphs and Related Topics Part 1*, Volume 2, North Holland, Amsterdam, **1991**, pp. IX 73–IX 78.
- [12] T. K. Dickens, R. B. Mallion, *Croat. Chem. Acta* **2013**, *86*, 387.
- [13] T. K. Dickens, R. B. Mallion, *MATCH Commun. Math. Comput. Chem.* **2016**, *76*, 297.
- [14] E. C. Kirby, R. B. Mallion, P. Pollak, P. J. Skrzyński, *Croat. Chem. Acta* **2016**, *89*, 403.
- [15] E. C. Kirby, D. J. Klein, R. B. Mallion, P. Pollak, H. Sachs, *Croat. Chem. Acta* **2004**, *77*, 263.
- [16] E. C. Kirby, R. B. Mallion, P. Pollak, 'Counting spanning trees in toroidal fullerenes' in *The Mathematics and Topology of Fullerenes*, (Series Eds.: F. Cataldo, P. Milani), (Volume Eds.: F. Cataldo, A. Graovac, O. Ori), (Foreword: Sir Harold Kroto), *Carbon Materials: Chemistry and Physics* Volume 4: Springer, Dordrecht, Heidelberg, London & New York, **2010**, Chapter 10, pp. 187–204.
- [17] I. Gutman, R. B. Mallion, J. W. Essam, *Mol. Phys.* **1983**, *50*, 859.
- [18] S. J. Cyvin, I. Gutman, *Kekulé Structures in Benzenoid Hydrocarbons*, Lecture Notes on Chemistry Series, Springer-Verlag, Berlin (West), **1988**.
- [19] S. J. Cyvin, I. Gutman, *Introduction to the Theory of Benzenoid Hydrocarbons*, Springer-Verlag, Berlin (West), **1989**.
- [20] D. Janežič, A. Miličević, S. Nikolić, N. Trinajstić, *Graph Theoretical Matrices in Chemistry* (Eds.: I. Gutman, B. Furtula), Mathematical Chemistry Monographs, Faculty of Science, University of Kragujevac, Serbia, **2007**, Chapter 2, pp. 5–50, especially pp. 46–49.
- [21] J. E. Anthony, *Angewandte Chem. Int. Edn.* **2008**, *47*, 452.
- [22] S. S. Zade, M. Bendikov, *Angewandte Chem. Int. Edn.* **2010**, *49*, 4012.
- [23] W. M. Hayes (Editor-in-Chief) *CRC Handbook of Chemistry and Physics*, 97th Edition, 2016–2017, CRC Press, Boca Raton, Florida, USA, **2016**.
- [24] M. S. Newman, D. Lednicer, *J. Am. Chem. Soc.* **1956**, *78*, 4765.
- [25] C. A. Coulson, C. W. Haigh, *Tetrahedron* **1963**, *19*, 527.
- [26] R. H. Martin, *Tetrahedron* **1964**, *20*, 897.
- [27] C. W. Haigh, R. B. Mallion, *Mol. Phys.* **1971**, *22*, 945.
- [28] R. H. Martin, *Angewandte Chem. Int. Edn.* **1974**, *13*, 649.
- [29] D. M. Cvetković, I. Gutman, *Publ. Inst. Math. (Beograd)* **1981**, *29*, 49.
- [30] A. T. Balaban, F. Harary, *Tetrahedron* **1968**, *24*, 2505.