

# A forgotten topological index

Boris Furtula · Ivan Gutman

Received: 6 November 2014 / Accepted: 27 January 2015 / Published online: 6 February 2015  
© Springer International Publishing Switzerland 2015

**Abstract** In 1972, within a study of the structure-dependency of total  $\pi$ -electron energy ( $\mathcal{E}$ ), it was shown that  $\mathcal{E}$  depends on the sum of squares of the vertex degrees of the molecular graph (later named first Zagreb index), and thus provides a measure of the branching of the carbon-atom skeleton. In the same paper, also the sum of cubes of degrees of vertices of the molecular graph was shown to influence  $\mathcal{E}$ , but this topological index was never again investigated and was left to oblivion. We now establish a few basic properties of this “forgotten topological index” and show that it can significantly enhance the physico-chemical applicability of the first Zagreb index.

**Keywords** Zagreb index · First Zagreb index · Topological index · Forgotten index

**Mathematics Subject Classification** Primary: 05C07 · Secondary: 05C90

## 1 Introduction

In 1972 a paper [1] was published, in which approximate formulas for the total  $\pi$ -electron energy ( $\mathcal{E}$ ) were derived. One of the terms occurring in these formulas was the sum of squares of vertex degrees of the molecular graph (in [1] denoted by  $\Sigma\sigma_1^2$ ). This quantity was recognized to be a measure of the extent of branching of the carbon-atom skeleton of the underlying molecule [1, 2], and ten years later was named the *first*

---

B. Furtula (✉) · I. Gutman  
Faculty of Science, University of Kragujevac, P.O. Box 60, 34000 Kragujevac, Serbia  
e-mail: furtula@kg.ac.rs

I. Gutman  
State University of Novi Pazar, Novi Pazar, Serbia  
e-mail: gutman@kg.ac.rs

Zagreb index [3]. Eventually, it became one of the most popular and most extensively studied graph-based molecular structure descriptors. For details of the theory and applications of the first Zagreb index see the surveys [4–6] and the references cited therein.

In the formulas for  $\mathcal{E}$  reported in the paper [1], also the sum of cubes of vertex degrees was encountered (in [1] denoted by  $\Sigma\sigma_1^3$ ). This quantity, in an obvious manner, is also a measure of branching. Yet, for reasons not easy to understand, it did not attract any attention, either in [1] or in any of the numerous subsequent publications.

The purpose of the present paper is to shed some light on this “*forgotten topological index*”, to establish some of its basic properties, and—especially—to show that in combination with the first Zagreb index, it possesses an outstanding applicative potential.

## 2 Notation and definitions

Let  $G$  be a molecular graph [7,8] with vertex set  $V(G)$  and edge set  $E(G)$ . If  $u$  and  $v$  are two adjacent vertices of  $G$ , then the edge connecting them will be denoted by  $uv$ . By  $\deg(v)$  we denote the degree (=number of first neighbors) of the vertex  $v$  of the graph  $G$ .

There are two Zagreb indices [2,3]: the first  $M_1$  and the second  $M_2$ , defined as

$$M_1 = M_1(G) = \sum_{v \in V(G)} \deg(v)^2$$

and

$$M_2 = M_2(G) = \sum_{uv \in E(G)} \deg(u) \deg(v)$$

respectively. The first Zagreb index can be rewritten also as [9,10]

$$M_1 = M_1(G) = \sum_{uv \in E(G)} [\deg(u) + \deg(v)]. \quad (1)$$

With this notation, the *forgotten topological index* is defined as

$$F = F(G) = \sum_{v \in V(G)} \deg(v)^3 = \sum_{uv \in E(G)} [\deg(u)^2 + \deg(v)^2]. \quad (2)$$

To our best knowledge, after appearing in the work [1], the quantity  $F$  was never again considered in the chemical and/or mathematical literature. On the other hand, for  $\alpha$  being an arbitrary real number, the generalized version of the first Zagreb index, namely

$$M_\alpha = M_\alpha(G) = \sum_{v \in V(G)} \deg(v)^\alpha = \sum_{uv \in E(G)} [\deg(u)^{\alpha-1} + \deg(v)^{\alpha-1}]$$

was studied in several earlier works [11–14]. However, in these works, the special case  $\alpha = 3$  was by no means singled out and no special attention was given to it.

### 3 On chemical applicability of the $F$ -index

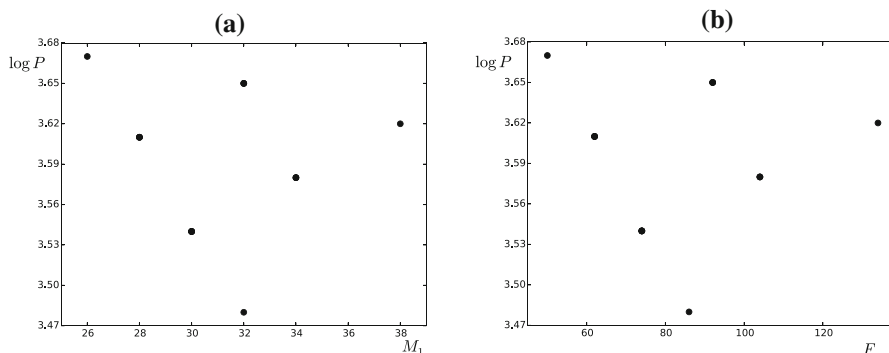
Concurring with a suggestion of the *International Academy of Mathematical Chemistry*, the predictive ability of the  $F$ -index was tested using a dataset of octane isomers, found at <http://www.molecularDescriptors.eu/dataset/dataset.htm>. This dataset has been chosen because the  $F$ -index, in its simplest form, does not recognize heteroatoms and multiple bonds. The octane dataset consists of the following data: boiling point, melting point, heat capacities, entropy, density, heat of vaporization, enthalpy of formation, motor octane number, molar refraction, acentric factor, total surface area, octanol-water partition coefficient, and molar volume. The  $F$ -index was correlated with each of these properties and the results were compared with those obtained by using the first Zagreb index. It was found that the predictive ability of the  $F$ -index is quite similar to that of  $M_1$ . In the case of entropy and acentric factor, both  $M_1$  and  $F$  yield correlation coefficients greater than 0.95. On the other hand, for other physico-chemical properties, neither  $M_1$  nor  $F$  are satisfactorily correlated. An example is shown in Fig. 1.

In order to improve the predictive ability of these indices, a simple linear model is devised:

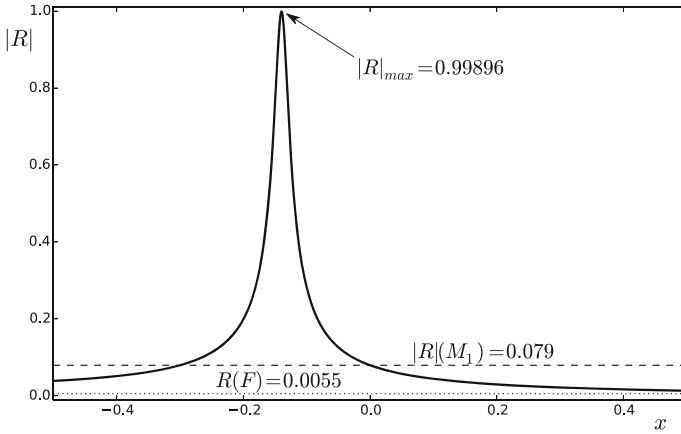
$$M_1 + \lambda F \quad (3)$$

where  $\lambda$  is fitting parameter. In order to achieve the best correlation, its value was varied from  $-20$  to  $20$ .

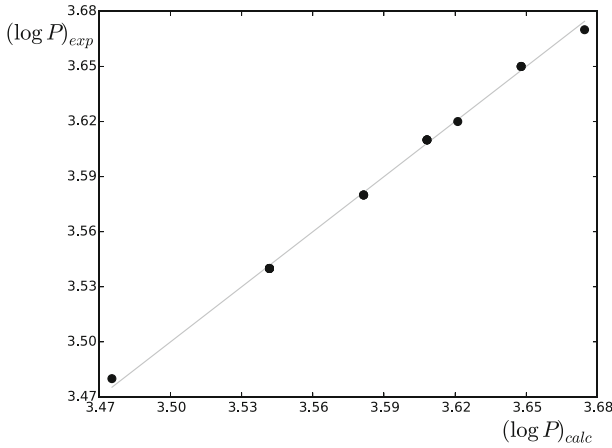
The above model was applied to each of the physico-chemical properties given in the octane database. Unfortunately, for all but a single physico-chemical property, the improvement achieved by the model (3) was insignificant. Surprisingly, however, in the case of the octanol-water partition coefficient, a major improvement could be gained.



**Fig. 1** Logarithm of the octanol-water partition coefficient ( $P$ ) plotted versus the first Zagreb and the forgotten indices, **a**  $\log P$  versus  $M_1$ ; correlation coefficient is 0.079, **b**  $\log P$  versus  $F$ ; correlation coefficient is 0.0055



**Fig. 2** Solid line shows the variation of the absolute value of the correlation coefficient with  $\lambda$ ; the dashed and dotted lines pertain to the correlation coefficients of  $\log P$  versus  $M_1$  and  $F$ , respectively



**Fig. 3** Experimental versus calculated  $\log P$  of octanes

From Fig. 1 is seen that the correlation between  $\log P$  and  $M_1$  and  $F$  is practically nil. Yet, by varying  $\lambda$  in formula (3), it was found that the absolute value of the correlation coefficient reaches a sharp maximum at  $\lambda = -0.140$  (see Fig. 2).

Thus we arrive at the following expression for the octanol-water partition coefficient of octanes:

$$(\log P)_{calc} = -0.2058(M_1 - 0.14 F) + 7.5864$$

Its quality is seen in Fig. 3. The correlation coefficient is 0.99896 and the mean absolute percentage error (MAPE) 0.06 %. Such a MAPE-value indicates a high accuracy of the prediction of  $\log P$ -values.

#### 4 Estimating the $F$ -index

Of several bounds for the  $F$ -index that could be deduced, we state here only the following three.

**Proposition 1** *Let  $G$  be a graph with  $m$  edges, whose first Zagreb index is  $M_1(G)$ . Then*

$$F(G) \geq \frac{M_1(G)^2}{2m}. \quad (4)$$

**Proposition 2** *Let  $G$  be a graph with  $m$  edges, whose first and second Zagreb indices are  $M_1(G)$  and  $M_2(G)$ , respectively. Then*

$$F(G) \geq \frac{M_1(G)^2}{m} - 2M_2(G). \quad (5)$$

In order to obtain (4), associate a weight  $w(u)$  with the vertex  $u$  of  $G$ . Then

$$\langle d \rangle_w = \frac{\sum_{u \in V(G)} w(u) \deg(u)}{\sum_{u \in V(G)} w(u)} \quad \text{and} \quad \langle d^2 \rangle_w = \frac{\sum_{u \in V(G)} w(u) \deg(u)^2}{\sum_{u \in V(G)} w(u)}$$

are, respectively, the weighted averages of vertex degrees and of squares of vertex degrees. For any non-negative weight,  $\langle d^2 \rangle_w \geq (\langle d \rangle_w)^2$ . Choosing  $w(u) = \deg(u)$ , and recalling that  $\sum_{u \in V(G)} \deg(u) = 2m$ , we straightforwardly arrive at the bound (4).

Using the right-hand side expression in (2), we get

$$F(G) = \sum_{uv \in E(G)} [\deg(u)^2 + \deg(v)^2 + 2 \deg(u) \deg(v)] - 2 \sum_{uv \in E(G)} \deg(u) \deg(v)$$

resulting in the identity

$$F(G) = \sum_{uv \in E(G)} [\deg(u) + \deg(v)]^2 - 2M_2(G). \quad (6)$$

Now,

$$\sum_{uv \in E(G)} [\deg(u) + \deg(v)]^2 = \frac{1}{m} \sum_{uv \in E(G)} [\deg(u) + \deg(v)]^2 \sum_{uv \in E(G)} 1^2$$

and by the Cauchy–Schwarz inequality

$$\sum_{uv \in E(G)} [\deg(u) + \deg(v)]^2 \sum_{uv \in E(G)} 1^2 \geq \left( \sum_{uv \in E(G)} [\deg(u) + \deg(v)] \times 1 \right)^2.$$

Bearing in mind relation (1), we thus get

$$\sum_{uv \in E(G)} [\deg(u) + \deg(v)]^2 \geq \frac{M_1(G)^2}{m}$$

which combined with (6) results in the lower bound (5).

Equality in both (4) and (5) is attained in the case of regular graphs (e.g. molecular graphs of fullerenes).

The lower bounds (4) and (5) are incomparable. Namely, there exist molecular graphs for which (4) is better than (5), and there exist molecular graphs for which (5) is better than (4). An example for  $M_1^2/(2m) > M_1^2/m - 2M_2$  is 1,2-diethylcyclobutane ( $m = 8$ ,  $M_1 = 36$ ,  $M_2 = 41$ ) whereas an example for  $M_1^2/(2m) < M_1^2/m - 2M_2$  is 1,3-diethylcyclobutane ( $m = 8$ ,  $M_1 = 36$ ,  $M_2 = 40$ ).

In the same way as identity (6) is obtained, we get

$$F(G) = \sum_{uv \in E(G)} [\deg(u) - \deg(v)]^2 + 2M_2(G). \quad (7)$$

If the graph  $G$  is connected (which necessarily is the case for molecular graphs), then the term  $[\deg(u) - \deg(v)]^2$  will assume its greatest possible value if  $\deg(u) = n - 1$  and  $\deg(v) = 1$  (or vice versa), where  $n$  is the number of vertices of  $G$ . Therefore,

$$\sum_{uv \in E(G)} [\deg(u) - \deg(v)]^2 \leq m(n - 2)$$

which substituted back into Eq. (7) yields:

**Proposition 3** *Let  $G$  be a connected graph with  $n$  vertices and  $m$  edges, whose second Zagreb index is  $M_2(G)$ . Then*

$$F(G) \leq 2M_2(G) + m(n - 2). \quad (8)$$

*Equality in (8) is attained if and only if  $G$  is the star graph.*

**Acknowledgments** Boris Furtula thanks the Serbian Ministry of Education and Science for partial support of this work through Grant No. 174033.

## References

1. I. Gutman, N. Trinajstić, Chem. Phys. Lett. **17**, 535 (1972)
2. I. Gutman, B. Ružičić, N. Trinajstić, C.F. Wilcox, J. Chem. Phys. **62**, 3399 (1975)
3. A.T. Balaban, I. Motoc, D. Bonchev, O. Mekenyan, Topics Curr. Chem. **114**, 21 (1983)
4. S. Nikolić, G. Kovačević, A. Miličević, N. Trinajstić, Croat. Chem. Acta **76**, 113 (2003)
5. I. Gutman, K.C. Das, MATCH Commun. Math. Comput. Chem. **50**, 83 (2004)
6. I. Gutman, Croat. Chem. Acta **86**, 351 (2013)
7. N. Trinajstić, *Chemical Graph Theory* (CRC Press, Boca Raton, 1983)
8. I. Gutman, O.E. Polansky, *Mathematical Concepts in Organic Chemistry* (Springer, Berlin, 1986)

9. T. Došlić, B. Furtula, A. Graovac, I. Gutman, S. Moradi, Z. Yarahmadi, *MATCH Commun. Math. Comput. Chem.* **66**, 613 (2011)
10. T. Došlić, T. Réti, D. Vukičević, *Chem. Phys. Lett.* **512**, 283 (2011)
11. X. Li, H. Zhao, *MATCH Commun. Math. Comput. Chem.* **50**, 57 (2004)
12. X. Li, J. Zheng, *MATCH Commun. Math. Comput. Chem.* **54**, 195 (2005)
13. G. Su, L. Xiong, L. Xu, *Appl. Math. Lett.* **25**, 1701 (2012)
14. I. Gutman, *MATCH Commun. Math. Comput. Chem.* **72**, 733 (2014)