© 2002 The Chemical Society of Japan Bull. Chem. Soc. Jpn., 75, 1723–1727 (2002)

The Relation between the Eigenvalue Sum and the Topological Index *Z* Revisited

Ivan Gutman, Dusica Vidović, and Haruo Hosoya[*,†]

Faculty of Science, University of Kragujevac, P. O. Box 60, YU-34000 Kragujevac, Yugoslavia

†Department of Information Sciences, Ochanomizu University, Bunkyo-ku, Tokyo 112-8610

(Received December 26, 2001)

The topological index Z (put forward by one of the present authors, *Bull. Chem. Soc. Jpn.*, **44**, 2332 (1971)) has proved to be a structure-descriptor suitable for modeling physico-chemical properties of both saturated and conjugated unsaturated organic compounds. A linear relation between log Z and the Hückel molecular orbital total π-electron energy E_π was established a long time ago (*Theor. Chim. Acta* **38**, 37 (1975)), but was restricted to π-electron species without non-bonding molecular orbitals ($n_0 = 0$, where n_0 is the number of zero eigenvalues of the molecular graph). Recently a novel structure-descriptor ES was put forward, the "eigenvalue sum" equal to $E_\pi/2$, but applicable to saturated hydrocarbons. In this case n_0 need not be zero. We demonstrate that, for each group of alkane isomers with a fixed value of n_0, there is a linear relation between ES and log Z.

The topological index Z was put forward in the early 1970s by one of the present authors[1] and was eventually demonstrated to be a structure-descriptor applicable for designing mathematical models of various physico-chemical properties of both saturated[2–7] and conjugated unsaturated[8–12] organic substances; for more detail see the review[13] and elsewhere.[14,15]

Among the first problems considered in the theory of the Z index was the existence of a relation between Z and the Hückel molecular orbital (HMO) total π-electron energy E_π (expressed, as usual, in the units of the carbon-carbon resonance integral β). Indeed, such a relation was soon discovered:[9] In the case of acyclic polyenes E_π is (as a good approximation) a linear function of the logarithm of the Z index. In the case of cyclic conjugated systems, the situation was found[9,16] to be somewhat more complicated.

As is well known,[17–19] conjugated π-electron systems possessing non-bonding molecular orbitals are highly unstable and usually do not exist. Therefore it is by no means surprising that the studies of the role of the Z index in HMO theory were almost completely restricted to systems without non-bonding MOs. (One paper[10] represents a noteworthy exception). As a consequence, the regularities that will be reported in this paper remained unnoticed in the researches[8–12,16] concerned with the HMO model and done in the 1970s.

Recently, Randić and Vračko[20] conceived a novel structure-descriptor ES which they named "eigenvalue sum". Initially they used ES in the study of primary DNA sequences.[20] Eventually they applied ES to alkanes.[21] As will be seen in the subsequent section, the eigenvalue sum is, formally, equal to half of the E_π-value of the corresponding molecular graph. Therefore, although the chemical interpretations of E_π and ES are completely different, the underlying mathematical analysis is one and the same.

In what follows we demonstrate that, for each group of alkane isomers with a fixed value of n_0, the number of zero eigenvalues, there is, as a good approximation, a linear relation between ES and log Z and that the (ES, log Z)-points pertaining to different n_0-values lie on different, nearly parallel, lines.

Preparations

Let G be a molecular graph,[14,22] possessing n vertices and m edges. Let $\lambda_1, \lambda_2, ..., \lambda_n$ be the eigenvalues of G. Let n_0 of these eigenvalues be equal to zero.

The number of ways in which k mutually non-touching edges can be selected in G is denoted by $p(G,k)$. In addition, $p(G,0) = 1$ and $p(G,1) = m$. Then the topological index Z is defined as[1]

$$Z = Z(G) = \sum_{k \geq 0} p(G,k).$$

If G represents the carbon-atom skeleton of a conjugated hydrocarbon, then in the HMO approximation, $\alpha + \lambda_i \beta$ is the energy of the i-th π molecular orbital.[17,19,23] Consequently, for a closed-shell system in its ground state the total π-electron energy is

$$E_\pi = E_\pi(G) = n\alpha + 2\beta \sum_{+} \lambda_i$$

where α and β are the standard HMO parameters, and where $\sum_{+}$ denotes summation over positive eigenvalues. In terms of the usual "β-units" (namely $\alpha = 0$, $\beta = 1$), the above expression becomes

$$E_\pi = E_\pi(G) = 2\sum_{+} \lambda_i. \tag{1}$$

If G is the graph-representation of the carbon-atom skeleton of a saturated hydrocarbon, then the Randić–Vračko "eigenvalue sum" is given by

$$ES = ES(G) = \sum_{+} \lambda_i. \tag{2}$$

Comparing Eqs. 1 and 2, we arrive at the equality

$$ES(G) = \frac{1}{2} E_\pi(G). \tag{3}$$

When using relation of Eq. 3 one must always bear in mind that the graph G on its left-hand side has a different chemical interpretation than the graph G on its right-hand side. On the other hand, Eq. 3 enables one to apply to saturated hydrocarbons the numerous mathematical results of the theory of HMO total π-electron energy previously obtained.

In particular, if G is a tree (= a connected graph without cycles), then a Coulson type integral formula for total π-electron energy[24] can be rewritten as

$$ES(G) = \frac{1}{\pi} \int_0^{+\infty} F(x) dx \tag{4}$$

where

$$F(x) = \frac{1}{x^2} \ln \sum_{k \geq 0} p(G,k) x^{2k}. \tag{5}$$

Recall that for trees, $m = n-1$.

Equations 4 and 5 are applicable to alkanes and will serve as the starting point of our analysis.

Relation between *ES* and *Z*: Empirical Approach

In Fig. 1 is shown a plot of the eigenvalue sums of all decanes versus the respective ln Z-values. The points are grouped into three clusters; a relatively good linear correlation between ES and ln Z exists within each cluster. By direct checking we found that all points belonging to the same cluster correspond to molecular graphs with equal numbers (n_0) of zero eigenvalues. Precisely the same regularity was observed also for all other isomeric alkanes with n = 6–12 carbon atoms.

The graphs whose (ES, ln Z)-points are shown in Fig. 1 are “chemical trees”, i.e., trees in which no vertex has degree greater than 4. In Fig. 2 the (ES, ln Z)-points are given for all 10-vertex trees, in which no restriction is imposed on the degree of vertices. This example shows that the regularity mentioned above is not restricted to chemical trees. Precisely the same regularity was observed also for all other complete sets of n-vertex trees with n = 6–12.

More data on the linear correlation between ES and ln Z can be found in Tables 1 and 2. From these data one can see that all the straight lines pertaining to $n_0 > 0$ are nearly parallel, with slopes belonging to a relatively narrow interval (0.8, 0.9). The lines corresponding to $n_0 = 0$ markedly deviate from this pattern, a detail that the model elaborated in the subsequent section cannot account for.

Concluding this section we wish to point out (without going into details) that analogous regularities hold also for molecules containing cycles. As a self-explanatory example, in Fig. 3 we show the correlation between the eigenvalue sums and ln Z for the 12 constitutional isomers of ethyldimethylcyclohexane.

Relation between *ES* and *Z*: Theoretical Approach

If G is a tree with n vertices and n_0 zero eigenvalues, then $m(G,k) > 0$ for k = 0, 1, ..., $(n - n_0)/2$ and $m(G,k) = 0$ for $k > (n - n_0)/2$. Bearing this in mind, the integrand (Eq. 5) can be transformed into

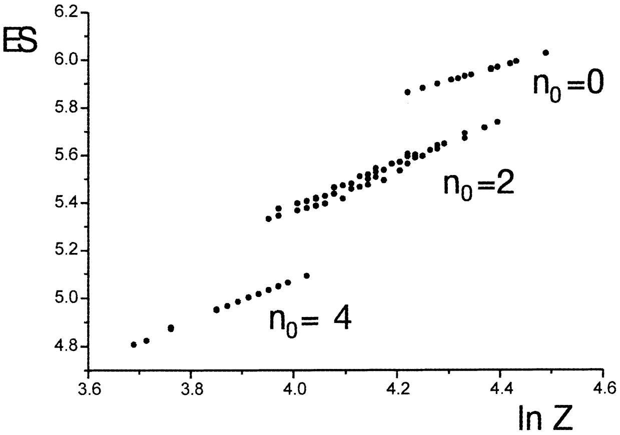

Fig. 1. The dependence of the eigenvalue sums of decanes on the logarithm of the Z index; n_0 is the number of zero eigenvalues in the spectrum of the respective molecular graphs.

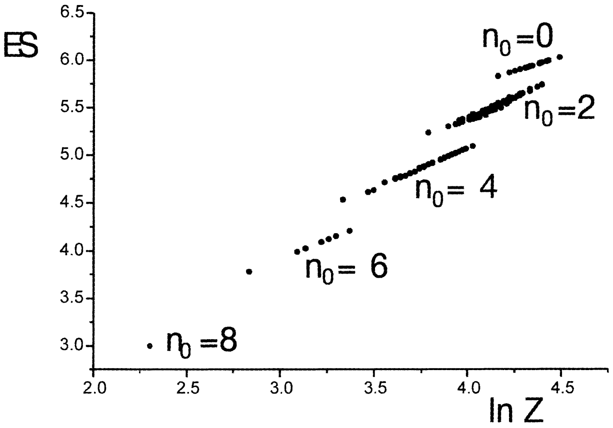

Fig. 2. Same data as in Fig. 1 for all 10-vertex trees.

$$F(x) = \frac{1}{x^2} \ln\big[1 + p(G,1)x^2 + p(G,2)x^4 + \dots p(G,(n-n_0)/2)x^{n-n_0}\big]. \tag{6}$$

It is now easily verified that the function $F(x)$ has the following properties:

$$\lim_{x \to 0} F(x) = m \tag{7}$$

$$F(1) = \ln Z \tag{8}$$

$$\lim_{x \to +\infty} \frac{F(x)}{(n-n_0)x^{-2} \ln x} = 1. \tag{9}$$

In the theory of total π-electron energy a method for constructing approximate formulas for E_π was elaborated,[25] which we now apply to Eqs. 4 and 6.

Suppose that we have a continuous function $F^*(x)$ that satisfies conditions analogous to Eqs. 7–9, i.e.,

Table 1. The Parameters A and B Obtained by Least-Squares Fitting, for the Correlation $ES \approx A \ln Z + B$ for Isomeric Alkanes with n Carbon Atoms, Whose Molecular Graphs Possess n_0 Zero Eigenvalues; N = Number of Data Points, R = Correlation Coefficient

n	n_0	N	A	B	R
6	0	2	—	—	—
6	2	3	0.839	1.066	0.99988
7	1	6	0.906	1.271	0.98675
7	3	3	0.789	1.279	0.99999
8	0	5	0.582	2.707	0.99999
8	2	11	0.920	1.286	0.99525
8	4	2	—	—	—
9	1	19	0.927	1.601	0.96875
9	3	15	0.870	1.525	0.99816
9	5	1	—	—	—
10	0	14	0.601	3.326	0.99970
10	2	47	0.936	1.625	0.98556
10	4	14	0.849	1.679	0.99914

Table 2. Same Data as in Table 1 for n-Vertex Trees

n	n_0	N	A	B	R
6	0	2	—	—	—
6	2	3	0.839	1.066	0.99988
6	4	1	—	—	—
7	1	6	0.906	1.271	0.98655
7	3	4	0.814	1.211	0.99988
7	5	1	—	—	—
8	0	5	0.582	2.707	0.99999
8	2	12	0.873	1.439	0.99448
8	4	5	0.807	1.319	0.99989
8	6	1	—	—	—
9	1	20	0.877	1.796	0.97673
9	3	20	0.855	1.580	0.99701
9	5	6	0.801	1.424	0.99993
9	7	1	—	—	—
10	0	15	0.600	3.333	0.99981
10	2	52	0.885	1.842	0.98559
10	4	31	0.841	1.710	0.99830
10	6	7	0.803	1.508	0.99995
10	8	1	—	—	—

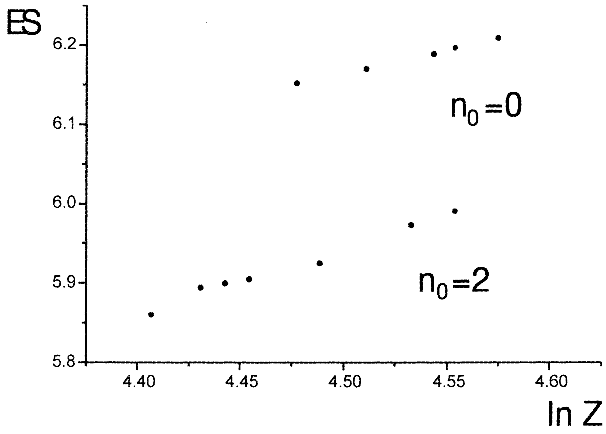

Fig. 3. The eigenvalue sums (ES) of the constitutional isomers of ethyldimethylcyclohexane versus the logarithm of the topological index Z; the respective molecular graphs have either no or two zero eigenvalues.

$$\lim_{x \to 0} F^*(x) = m \tag{10}$$

$$F^*(1) = \ln Z \tag{11}$$

$$\lim_{x \to +\infty} \frac{F^*(x)}{(n-n_0)x^{-2}\ln x} = 1. \tag{12}$$

Then the quantity ES^*, calculated by means of the integral

$$ES^*(G) = \frac{1}{\pi}\int_0^{+\infty} F^*(x)\,dx \tag{13}$$

is an approximation of the eigenvalue sum ES.

A class of functions with the required property is given as follows:

$$F^*(x) = \begin{cases} m-(m-\ln Z)x^5 & \text{for } 0 \leq x \leq 1 \\ [\ln Z + (n-n_0)x^t \ln x]x^{-(t+2)} & \text{for } x \geq 1 \end{cases} \tag{14}$$

Table 3. The Parameters s, t, and γ That Need To Be Used in Formula (Eq. 16) in the Case of Isomeric Alkanes, and Statistical Indicators of the Accuracy of the Approximation $ES \approx ES^*$; n = Number of Carbon Atoms, N = Number of Constitutional Isomers of C_nH_{2n+2}, i.e., Number of n-Vertex Chemical Trees, R = Correlation Coefficient for ES vs ES^*, ARE = Average Relative Error (in %)

n	N	s	t	γ	R	ARE
6	5	−0.23	−0.19	0.46	0.9996	0.028
7	9	−0.30	−0.30	0.47	0.9990	0.023
8	18	−0.30	−0.30	0.43	0.9986	0.018
9	35	−0.32	−0.33	0.43	0.9985	0.015
10	75	−0.33	−0.34	0.41	0.9979	0.003
11	159	−0.31	−0.34	0.42	0.9981	0.0001
12	355	−0.32	−0.35	0.40	0.9976	0.0001

Table 4. Same Data as in Table 3 for Trees with a Given Number n of Vertices; N = Number of n-Vertex Trees

n	N	s	t	γ	R	ARE
6	6	−0.33	−0.30	0.43	0.9996	0.064
7	11	−0.27	−0.27	0.48	0.9997	0.006
8	23	−0.29	−0.29	0.44	0.9994	0.019
9	47	−0.27	−0.29	0.46	0.9995	0.002
10	106	−0.28	−0.30	0.44	0.9993	0.004
11	235	−0.28	−0.31	0.44	0.9993	0.0001
12	551	−0.28	−0.31	0.43	0.9991	0.0002

Relations (10)-(12) are obeyed for any choice of the parameters s and t provided $s > 0$ and $t > -2$.

In Fig. 4 $F^*(x)$ curves with two different sets of s and t parameters for the 2,2,3,3-tetramethybutane graph are compared with the $F(x)$ curves. It is generally observed that, for other alkane graphs with different n and n_0 values, $F^*(x)$ is a good approximation to $F(x)$ for various combinations of s and t parameters.

By substituting Eq. 14 back into Eq. 13 and by performing elementary integration one obtains

$$ES^* = \frac{1}{\pi}\left[m+n+n_0-\frac{m-\ln Z}{s+1}+\frac{\ln Z}{t+1}\right]$$

i.e.,

$$ES^* = \frac{1}{\pi}\left[(1-\frac{1}{s+1})m+n-n_0+(\frac{1}{s+1}+\frac{1}{t+1})\ln Z\right].(15)$$

Formula (Eq. 15) is in good qualitative agreement with the empirical findings specified in the preceding section. Indeed, its right-hand side is a linear function of the logarithm of the Z index, and various n_0-values result in different (yet mutually parallel) lines.

Numerical testing revealed that in order to achieve *quantitative* agreement with the observed ES-ln Z relations (cf. Figs. 1 and 2), the model (Eq. 15) had to be somewhat modified:

$$ES^* = \frac{1}{\pi}\left[(1-\frac{1}{s+1})m+n-\gamma n_0+(\frac{1}{s+1}+\frac{1}{t+1})\ln Z\right] \quad (16)$$

where γ is a parameter measuring the magnitude of the effect of zero eigenvalues on the ES–ln Z relations. The optimal value of γ is around 0.45 (see Tables 3 and 4), implying that formula (Eq. 15) overestimates the influence of zero eigenvalues by more than 100%. In addition to this, in Eq. 16 we permit negative values of s provided $s + 1 > 0$.

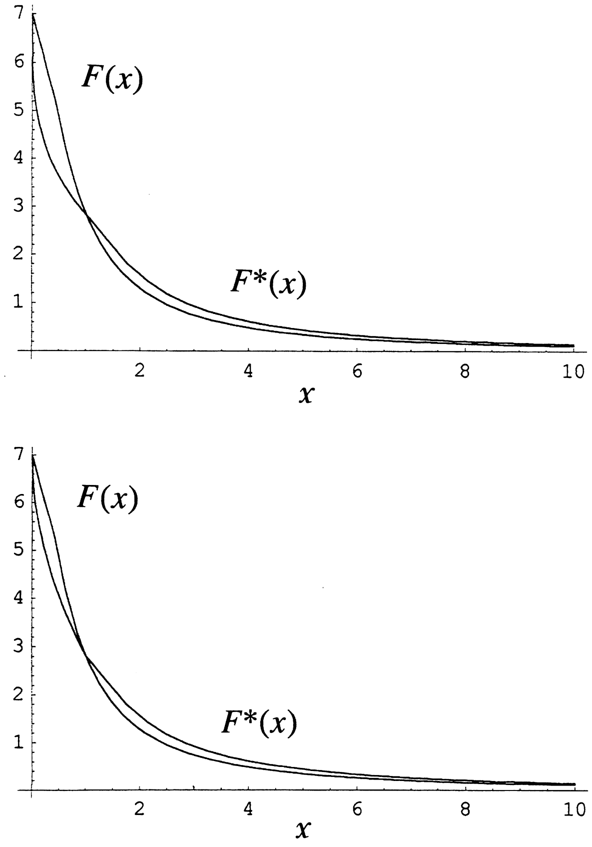

Fig. 4. $F(x)$ and $F^*(x)$ of 2,2,3,3-tetramethylbutane graph (n = 8 and n_0 = 4). (a) s =0.3 and t = −0.3, (b) s =0.5 and t = 0. $F^*(x)$ curves generally drop down more sharply than $F(x)$ curves, while for larger x just the opposite inclination can be observed.

In Table 3 are given the values of the parameters s, t and γ in Eq. 16 that optimize the approximation $ES \approx ES^*$ in the case of isomeric alkanes. In Table 4 are found the analogous data for trees. As seen, the accuracy of formula (Eq. 16) is remarkably good.

Discussion

By means of the theory outlined in the preceding sections, the main features of the relation between the Z index and the eigenvalue sum (i.e., between Z and total π-electron Energy) have been established and explained. By this, a research project of the present authors, started long time ago,[1,9,24] has now been completed.

Yet, a few details remain unresolved.

First of all, our considerations cannot explain why the ($n_0 = 0$)-line is not parallel with the other lines, cf. Figs. 1 and 2. This indicates that the model put forward in this work may and should be further upgraded.

The optimized values of the parameters s and t are remarkably close to each other and are negative-valued; in addition, s, t, γ vary only slightly with n (cf. Tables 3 and 4). It would be good to know why this is so.

We trust that answers to these questions will be obtained in the future. We hope that it will not be necessary to wait an additional 30 years to arrive at these answers.

References

1 H. Hosoya, *Bull. Chem. Soc. Jpn.*, **44**, 2332 (1971).

2 H. Hosoya, K. Kawasaki, and K. Mizutani, *Bull. Chem. Soc. Jpn.*, **45**, 3415 (1972).

3 H. Narumi and H. Hosoya, *Bull. Chem. Soc. Jpn.*, **53**, 1228 (1980).

4 H. Narumi and H. Hosoya, *Bull. Chem. Soc. Jpn.*, **58**, 1778 (1985).

5 I. Gutman, *Z. Phys. Chem. (Leipzig)*, **267**, 1152 (1986).

6 Y. D. Gao and H. Hosoya, *Bull. Chem. Soc. Jpn.*, **61**, 3093 (1988).

7 H. Hosoya, M. Gotoh, M. Murakami, and S. Ikeda, *J. Chem. Inf. Comput. Sci.*, **39**, 192 (1999).

8 H. Hosoya, *Theor. Chim. Acta*, **25**, 215 (1972).

9 H. Hosoya, K. Hosoi, and I. Gutman, *Theor. Chim. Acta*, **38**, 37 (1975).

10 H. Hosoya and M. Murakami, *Bull. Chem. Soc. Jpn.*, **48**, 3512 (1975).

11 H. Hosoya and K. Hosoi, *J. Chem. Phys.*, **64**, 1065 (1976).

12 I. Gutman, T. Yamaguchi, and H. Hosoya, *Bull. Chem. Soc. Jpn.*, **49**, 1811 (1976).

13 H. Hosoya, in "Mathematics and Computational Concepts in Chemistry," ed by N. Trinajstić, Horwood, Chichester (1986), pp. 110–123.

14 I. Gutman and O. E. Polansky, "Mathematical Concepts in Organic Chemistry," Springer-Verlag, Berlin (1986).

15 R. Todeschini and V. Consonni, "Handbook of Molecular Descriptors," Wiley-VCH, Weinheim (2000), pp. 215–219.

16 J. Aihara, *J. Org. Chem.*, **41**, 2488 (1976).

17 I. Gutman and N. Trinajstić, *Topics Curr. Chem.*, **42**, 49 (1973).

18 D. Cvetković, I. Gutman, and N. Trinajstić, *J. Mol. Struct.*, **28**, 289 (1975).

19 J. R. Dias, "Molecular Orbital Calculations Using Chemical Graph Theory," Springer-Verlag, Berlin (1993).

20 M. Randić and M. Vračko, *J. Chem. Inf. Comput. Sci.*, **40**, 599 (2000).

21 M. Randić, M. Vračko, and M. Nović, in "QSPR/QSAR Studies by Molecular Descriptors," ed by M. V. Diudea, Nova, Huntington, N.Y. (2001), pp. 147–211. They report that good correlation is observed for *ES* of octane isomers with their boiling point, entropy, heat of formation, surface tension, etc.

22 N. Trinajstić, "Chemical Graph Theory," CRC Press, Boca Raton, Fl. (1983).

23 A. Graovac, I. Gutman, and N. Trinajstić, "Topological Approach to the Chemistry of Conjugated Molecules," Springer-Verlag, Berlin (1977).

24 I. Gutman, *Theor. Chim. Acta*, **45**, 79 (1977).

25 I. Gutman, *J. Chem. Phys.*, **66**, 1652 (1977).