Two Variants of the Topological Index and the Relations between Them

Ivan GUTMAN,* Haruo HOSOYA,† Gordana URAKOVIĆ, and Lidija RISTIĆ
Faculty of Science, University of Kragujevac, P. O. Box 60, YU-34000 Kragujevac, Yugoslavia
† Department of Chemistry, Ochanomizu University, Bunkyo-ku, Tokyo 112
(Received August 28, 1991)

The two topological indices σ (recently proposed by Merrifield and Simmons) and Z (put forward by one of the present authors, *Bull. Chem. Soc. Jpn.*, **44**, 2332 (1971)) are shown to be both conceptually and numerically related. For alkanes, a linear correlation exists between ln σ and ln Z. For unbranched catacondensed benzenoid hydrocarbons σ and Z are linearly correlated. In any case σ and Z exhibit a reverse dependency on molecular structures: structural transformations having an increasing effect on Z decrease σ and vice versa.

The topological index Z was introduced in 1971 by one of the present authors.[1] Numerous subsequent researches have documented its usefulness in the characterization of molecular structure[2-4] and in the prediction of various physico-chemical properties of both saturated[4-9] and conjugated hydrocarbon systems.[10-12]

If G is a molecular graph then the respective topological index $Z=Z(G)$ is defined[1] as:

$$Z(G) = \sum_{k=0}^{m} p(G, k), \qquad (1)$$

where $p(G,k)$ stands for the number of ways in which k disjoint *edges* can be chosen in the graph G; by definition $p(G,0)=1$ and $p(G,1)=m$ (number of edges of G).

In a series of recently published works[13-16] and especially in the book[17] Merrifield and Simmons developed a novel mathematical approach to the molecular structure. They introduced a molecular topological space whose cardinality (denoted by σ) plays a significant role in their theory. It can be shown[15,17] that the quantity $\sigma=\sigma(G)$ is determined by the structure of the molecular graph G in the following way;

$$\sigma(G) = \sum_{k=0}^{n} q(G, k), \qquad (2)$$

where $q(G,k)$ stands for the number of ways in which k disjoint *vertices* can be chosen in the graph G; by definition $q(G,0)=1$ and $q(G,1)=n$ (number of vertices of G).

Evidently, there is a great deal of similarity in the definitions of the two topological indices, Z and σ. In the language of graph theory, $Z(G)$ and $\sigma(G)$ count the independent edge- and vertex-sets, respectively, of the graph G. In view of this, Z and σ can be considered as two variants of the same concept, reflecting the same structural features of the molecule considered. Having in mind the numerous chemical applications of the topological index Z, it would be of some interest to explore the possible usages of σ in the study of structure-property relations. Before this, however, it is necessary to test the (mathematical and/or statistical) independence of the two topological indices. If, namely, the two indices are independent (i.e., not related in some exact or approximate manner), then they could be used for modelling different chemical phenomena. Otherwise the domains of applicability of σ and Z would essentially coincide.

In this paper we communicate such results showing that a high degree of correlation exists between σ and Z, in particular in the case of alkanes and unbranched catacondensed benzenoids. These correlations are reverse. Namely, large σ values correspond to small Z values and vice versa.

I. Mathematical Properties of Z and σ

The calculation of the numerical values of the topological indices Z and σ of chemical graphs directly from Eqs. 1 and 2 is extremely tedious and impractical, except for very small m and n. Fortunately, however, several recursion formulas have been obtained which make these calculations relatively easy.[1]

Before presenting our theory let us define several subgraphs of graph G in which a pair of vertices u and v are connected to form edge uv. Either u or v is surrounded by the sets of the first neighbors, the second neighbors, etc. as depicted by the concentric semicircles in Fig. 1. Subgraph $G-uv$ is derived from G by deleting only edge uv leaving vertices u and v. $G-(uv)$ is derived from G-uv by deleting all the edges radiating from u and v to their first neighbor vertices. On the other hand,

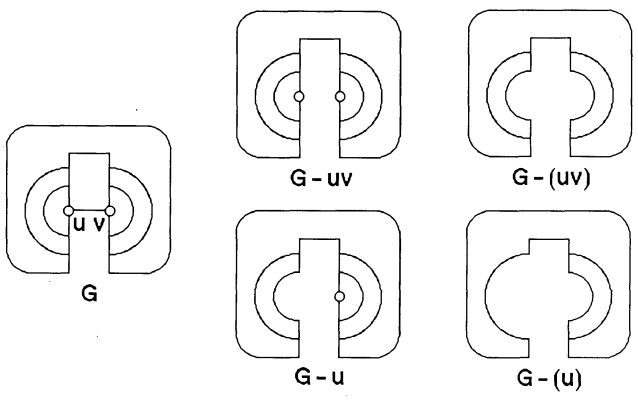


Fig. 1. Definitions of various subgraphs of graph G.

$G-u$ and $G-(u)$ are derived, respectively, by deleting all the edges radiating from u to the first and second neighbor vertices of u.

Then the fundamental recurrence relations for Z and σ read

$$Z(G) = Z(G-uv) + Z(G-(uv)) \quad (3)$$

and

$$\sigma(G) = \sigma(G-u) + \sigma(G-(u)).^{4,17} \quad (4)$$

In addition to this we have

$$Z(G_1 \cup G_2) = Z(G_1)Z(G_2) \quad (5)$$

$$\sigma(G_1 \cup G_2) = \sigma(G_1)\sigma(G_2), \quad (6)$$

where $G_1 \cup G_2$ is the graph composed of two disjoint components G_1 and G_2.

Suppose now that the vertices u and v of the graph G are not adjacent to each other as in Fig. 2. Denote by $G+uv$ the graph obtained from G by adding a new edge uv. Then a simple reformation of Eq. 3 yields

$$Z(G+uv) = Z(G) + Z(G-u-v). \quad (7)$$

In order to obtain the analogous expression for the σ-index, observe that $G+uv-u=G-u$ and $G+uv-(u)=G-(u)-v$ as illustrated in Fig. 2.

Then Eq. 4 implies

$$\sigma(G+uv) = \sigma(G-u) + \sigma(G-(u)-v). \quad (8)$$

Applying Eq. 4 to $G-(u)$ one gets

$$\sigma(G-(u)) = \sigma(G-(u)-v) + \sigma(G-(u)-(v)). \quad (9)$$

Combining Eqs. 4, 8, and 9 we arrive at

$$\sigma(G+uv) = \sigma(G) - \sigma(G-(u)-(v)). \quad (10)$$

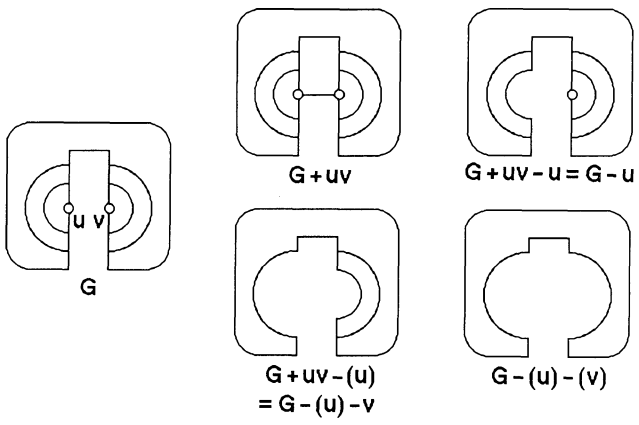


Fig. 2. Graphs G, $G+uv$, and the relevant subgraphs.

By comparing Eqs. 7 and 10 we reach the following elementary conclusion.

Proposition 1. The transformation $G \rightarrow G+uv$ has a positive effect on $Z(G)$ and a negative effect on $\sigma(G)$.

We anticipate that Proposition 1 is just a special case of a much more general property of Z and σ:

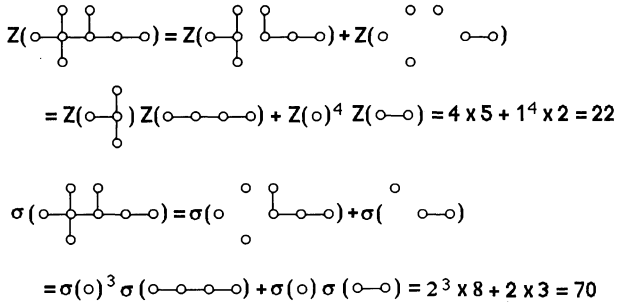
Proposition 2. A transformation of the graph G which does not change the number of vertices either increases $Z(G)$ and decreases $\sigma(G)$, or vice versa.

Proposition 2 means that the Z- and σ-indices of graphs having equal n and m (e.g., of molecular graphs of isomeric species) should be somehow related. Furthermore, large (or small) Z-values should correspond to small (or large) σ-values, and vice versa.

Although we are not able to provide a general proof of Proposition 2, we managed to demonstrate its validity in the case of two important classes of organic molecules, i.e., saturated and benzenoid hydrocarbons. The pertinent details are outlined in the two subsequent sections.

II. Relations between the Z- and σ-Indices of Alkanes

In the case of the molecular graphs of alkanes the recurrence relations 3—6 enable a very easy pencil-and-paper calculation of both Z and σ. We illustrate this on the example of 2,2,3-trimethylpenatne:



In such calculations extensive use can be made of the facts[3,17] that $Z(P_n)=F_n$ and $\sigma(P_n)=F_{n+1}$, where P_n is the path-graph with n vertices and F_n is the n-th Fibonacci number. (Recall that $F_1=1$, $F_2=2$, $F_3=3$, $F_4=5$, $F_5=8$, $F_n=F_{n-1}+F_{n-2}$.)

In Fig. 3 and Table 1 are presented the results of our studies of the correlation between Z and σ of the alkane isomers.

As seen from Fig. 3, a good linear correlation exists between the logarithms of Z and σ, and apparently, the lines obtained for different groups of isomers are parallel. A more careful analysis of the data collected in Table 1 reveals that the regression lines are not only parallel (i.e., the parameter a is practically independent of the number n of carbon atoms), but that these lines are equidistant (i.e., the parameter b is a linear function of the number n of carbon atoms).

Table 1. The Coefficients a and b in the Formula $\ln \sigma = a \ln Z + b$ Obtained by Least Squares Fitting for the Alkanes C_nH_{2n+2}

n	Number of isomers	a	b	Correlation coefficient
6	5	−0.562	4.484	−0.996
7	9	−0.554	5.216	−0.997
8	18	−0.553	5.958	−0.997
9	35	−0.565	6.753	−0.997
10	75	−0.562	7.490	−0.997
11	159	−0.559	8.228	−0.996
12	355	−0.561	8.992	−0.996

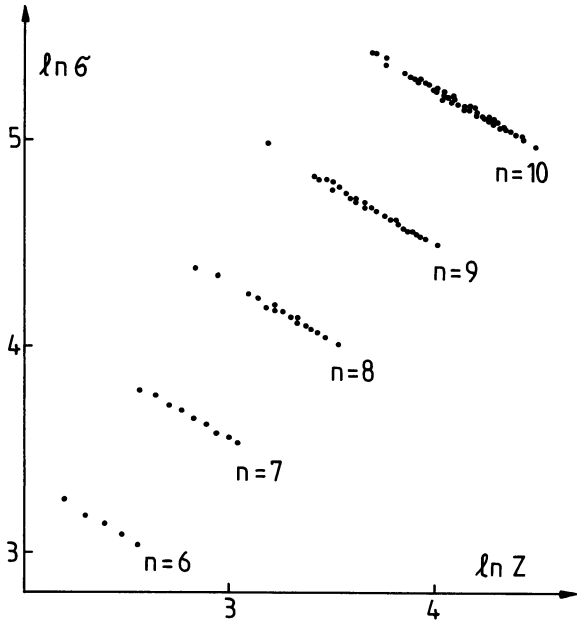
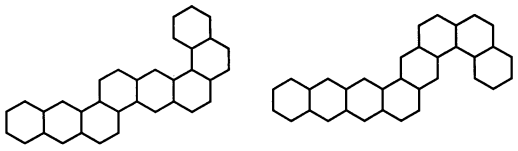


Fig. 3. Correlation between the topological indices Z and σ of isomeric alkanes C_nH_{2n+2} for n=6—10.

III. Relations between the Z- and σ-Indices of Unbranched Catacondensed Benzenoid Hydrocarbons

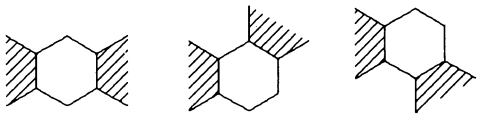
Unbranched catacondensed benzenoid hydrocarbons form an important class of benzenoid compounds. Numerous representatives of this class are chemically well-characterized substances.[18,19] All catacondensed systems with h hexagons have the formula $C_{4h+2}H_{2h+4}$ and are thus isomers.

The molecular graphs of unbranched catacondensed benzenoid hydrocarbons (sometimes referred to as hexagonal chains or animals) are composed of hexagons which either share a common edge or are disjoint.[20] Furthermore, no three hexagons can share a common vertex and each hexagon has not more than two adjacent hexagons.[20] Two example of such molecular graphs are U_1 and U_2.



U_1 U_2

The number of hexagons in a hexagonal chain is denoted by h. Two of these hexagons are terminals, i.e., with just one neighbor. The remaining $h-2$ non-teminal hexagons have two neighbors and can be of the following three types:



0 1 2

A hexagon of type 0 is said to be linearly annelated whereas hexagons of types 1 and 2 are annelated in two different angular modes.[20] By indicating the annelation types of the non-terminal hexagons we obtain the so-called 3D (=three digit) code, introduced by Balaban.[21,22] It is easy to see that the 3D-code fully characterizes the respective unbranched catacondensed benzenoid system. For instance, the 3D-codes of U_1 and U_2, counting from the second leftmost hexagons, are 012011 and 001022, respectively.

In the general case the 3D-code of an h-cyclic unbranched catacondensed benzenoid system U can be written as $3D(U)=[S_2,S_3,\cdots,S_{h-1}]$, where S_i (i=2,3,$\cdots$,$h-1$) symbolizes the annelation mode of the i-th hexagon of U, i.e, S_i=0, 1, or 2.

The topological indices Z and σ of unbranched benzenoids can be obtained by an appropriately modified operator technique, originally proposed by Hosoya and Ohkami[23] and elaborated in a few subsequent papers.[24-27] When applied to the Z- and σ-indices this method yields the following results.

Let U be an unbranched benzenoid system with h hexagons, $h \geqslant 3$, whose 3D-code is $[S_2,S_3,\cdots,S_{h-1}]$. Let further $\boldsymbol{L}_Z, \boldsymbol{M}_Z, \boldsymbol{N}_Z, \boldsymbol{L}_\sigma, \boldsymbol{M}_\sigma$, and $\boldsymbol{N}_\sigma$ be square matrices of order 4, defined as

$$\boldsymbol{L}_Z=\begin{bmatrix}5&3&3&2\\2&2&1&1\\2&1&2&1\\1&1&1&1\end{bmatrix},\quad \boldsymbol{M}_Z=\begin{bmatrix}5&3&3&2\\3&0&2&0\\2&2&1&1\\2&0&1&0\end{bmatrix},\quad \boldsymbol{N}_Z=\begin{bmatrix}5&3&3&2\\2&1&2&1\\3&2&0&0\\2&1&0&0\end{bmatrix},$$

$$\boldsymbol{L}_\sigma=\begin{bmatrix}3&2&2&1\\2&2&1&1\\2&1&2&1\\1&1&1&1\end{bmatrix},\quad \boldsymbol{M}_\sigma=\begin{bmatrix}3&2&2&1\\3&0&2&0\\2&2&1&1\\2&0&1&0\end{bmatrix},\quad \boldsymbol{N}_\sigma=\begin{bmatrix}3&2&2&1\\2&1&2&1\\3&2&0&0\\2&1&0&0\end{bmatrix}.$$

Proposition 3. Let $\boldsymbol{T}_Z$ be a square matrix of order 4, defined as

$$T_Z = X_1 X_2 \cdots X_h,$$

where $X_1 = X_h = L_Z$ and

$$X_i = \begin{cases} L_Z & \text{if } S_i = 0 \\ M_Z & \text{if } S_i = 1 \\ N_Z & \text{if } S_i = 2 \end{cases} \quad \text{for } i = 2, \cdots, h-1.$$

Then

$$Z(\mathrm{U}) = (T_Z Q_Z)_{11}, \tag{11}$$

where $Q_Z = (2\ 1\ 1\ 1)^t$ (t:transposed matrix)

Proposition 4. Let $T_\sigma = Y_1 Y_2 \cdots Y_h$, where $Y_1 = Y_h = L_\sigma$ and

$$Y_i = \begin{cases} L_\sigma & \text{if } S_i = 0 \\ M_\sigma & \text{if } S_i = 1 \\ N_\sigma & \text{if } S_i = 2 \end{cases} \quad \text{for } i = 2, \cdots, h-1.$$

Then

$$\sigma(\mathrm{U}) = (T_\sigma Q_\sigma)_{11}, \tag{12}$$

where $Q_\sigma = (3\ 2\ 2\ 1)^t$.

Propositions 3 and 4 can be considered as special cases of results communicated in Refs. 26 and 27, respectively. Therefore we state them without (otherwise lengthy) proofs. As an illustration of these propositions we apply them to the systems U_1 and U_2. Then

$$T_Z(\mathrm{U}_1) = L_Z L_Z M_Z N_Z L_Z M_Z M_Z L_Z$$

$$T_Z(\mathrm{U}_2) = L_Z L_Z L_Z M_Z L_Z N_Z N_Z L_Z$$

$$T_\sigma(\mathrm{U}_1) = L_\sigma L_\sigma M_\sigma N_\sigma L_\sigma M_\sigma M_\sigma L_\sigma$$

$$T_\sigma(\mathrm{U}_2) = L_\sigma L_\sigma L_\sigma M_\sigma L_\sigma N_\sigma N_\sigma L_\sigma.$$

After performing the necessary matrix multiplications (which, of course, have to be done on a computer) we obtain $Z(\mathrm{U}_1)=47483194$, $Z(\mathrm{U}_2)=46516602$, $\sigma(\mathrm{U}_1)=7233243$, and $\sigma(\mathrm{U}_2)=7328121$.

Based on Eqs. 11 and 12 we have computed the topological indices Z and σ for all unbranched catacondensed benzenoids with 9 and less hexagons. A very good linear correlation between them has been found, as seen from Fig. 4 and the data collected in Table 2.

The regression lines for different values of h are not parallel. Their slopes (i.e., the parameter a in Table 2) seem to have no simple dependence on h.

The most striking feature of the Z/σ correlation is the grouping of the points according to the number of linearly annelated hexagons. This is clearly seen from Fig. 4.

If l denotes the number of linearly annelated hexagons (i.e., the number of digits 0 in the 3D-code) then not a single violation of the inequalities

$$Z(l+1) > Z(l)$$

$$\sigma(l+1) < \sigma(l)$$

has been observed, provided the number of hexagons is fixed. In particular, the linear polyacenes (for which $l = l_{\max} = h-2$) have maximum Z- and minimum σ-values among all unbranched catacondensed benzenoid hydrocarbons.

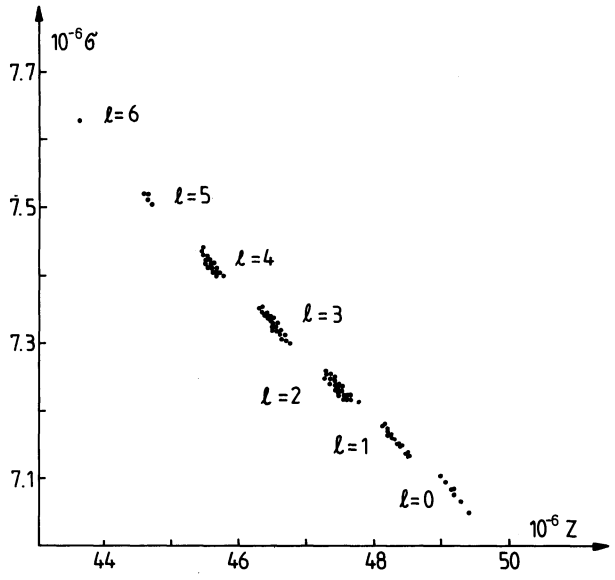


Fig. 4. Correlation between the topological indices Z and σ of isomeric unbranched catacondensed benzenoid hydrocarbons with 8 hexagons; l=number of linearly annelated hexagons.

Table 2. The Coefficients a and b in the Formula $\sigma = aZ + b$ Obtained by Least Squares Fitting for the Catafusenes $C_{4h+2}H_{2h+4}$

h	Number of isomers	a	b	Correlation coefficient
4	4	−3.580	2.641×10^4	−0.9998
5	10	−4.658	2.176×10^5	−0.9995
6	24	−6.091	1.799×10^6	−0.9994
7	67	−8.222	1.516×10^7	−0.9989
8	182	−10.557	1.239×10^8	−0.9996
9	520	−13.793	1.023×10^9	−0.9992

IV. Conclusion

We have analyzed the behavior of the topological indices Z and σ in two classes of organic molecules: (acyclic) alkanes and (polycyclic) benzenoids. In both cases Z and σ are strongly correlated, although the form of the correlation is completely different. These studies were possible because there exist special, convenient

mathematical techniques for the calculation of Z and σ for acyclic graphs and hexagonal chains. For other types of organic molecules or molecular graphs the comparative examination of Z and σ has not been undertaken, mainly because of the lack of an appropriate computational method. Nevertheless, we deem that already the present two examples provide a strong indication in favor of the conclusion that Z and σ are not independent topological indices. In other words, Z and σ are influenced by the same structural factors of the respective molecule and, consequently, they will be successful in the modelling of the same physico-chemical properties of organic compounds.

References

1) H. Hosoya, *Bull. Chem. Soc. Jpn.*, **44**, 2332 (1971).
2) H. Hosoya, *J. Chem. Doc.*, **12**, 181 (1972).
3) H. Hosoya, *Fibonacci Quart.*, **11**, 255 (1973).
4) I. Gutman and O. E. Polansky, "Mathematical Concepts in Organic Chemistry," Springer-Verlag, Berlin (1986), pp. 127—134.
5) H. Hosoya, K. Kawasaki, and K. Mizutani, *Bull. Chem. Soc. Jpn.*, **45**, 3415 (1972).
6) H. Narumi and H. Hosoya, *Bull. Chem. Soc. Jpn.*, **53**, 1228 (1980).
7) H. Narumi and H. Hosoya, *Bull. Chem. Soc. Jpn.*, **58**, 1778 (1985).
8) I. Gutman, *Z. Phys. Chem. (Leipzig)*, **267**, 1152 (1986).
9) Y-D. Gao and H. Hosoya, *Bull. Chem. Soc. Jpn.*, **61**, 3093 (1988).
10) H. Hosoya, K. Hosoi, and I. Gutman, *Theor. Chim. Acta*, **38**, 37 (1975).
11) H. Hosoya and K. Hosoi, *J. Chem. Phys.*, **64**, 1065 (1976).
12) I. Gutman, *Theor. Chim. Acta*, **45**, 79 (1977).
13) H. E. Simmons and R. E. Merrifield, *Proc. Natl. Acad. Sci.*, **74**, 2616 (1977).
14) R. E. Merrifield and H. E. Simmons, *Theor. Chim. Acta*, **55**, 55 (1980).
15) R. E. Merrifield and H. E. Simmons, *Proc. Natl. Acad. Sci.*, **78**, 692 (1981).
16) R. E. Merrifield and H. E. Simmons, *Proc. Natl. Acad. Sci.*, **82**, 1 (1985).
17) R. E. Merrifield and H. E. Simmons, "Topological Methods in Chemistry," Wiley, New York (1989).
18) E. Clar, "Polycyclic Hydrocarbons," Academic Press, London (1964).
19) J. R. Dias, "Handbook of Polycyclic Hydrocarbons. Part A. Benzenoid Hydrocarbons," Elsevier, Amsterdam (1987).
20) I. Gutman and S. J. Cyvin, "Introduction to the Theory of Benzenoid Hydrocarbons," Springer-Verlag, Berlin (1989).
21) A. T. Balaban, *Rev. Roum. Chim.*, **15**, 1243 (1970).
22) A. T. Balaban, *Pure Appl. Chem.*, **52**, 1409 (1980); **54**, 1075 (1982).
23) H. Hosoya and N. Ohkami, *J. Comput. Chem.*, **4**, 585 (1983).
24) H. Hosoya and A. Motoyama, *J. Math. Phys.*, **26**, 157 (1985).
25) O. E. Polansky, M. Randić, and H. Hosoya, *MATCH*, **24**, 3 (1989).
26) M. Randić, H. Hosoya, and O. E. Polansky, *J. Comput. Chem.*, **10**, 683 (1989).
27) I. Gutman, *Rev. Roum. Chim.*, in press.