An Algorithm for the Calculation of the Szeged Index of Benzenoid Hydrocarbons

Ivan Gutman*,[†,‡] and Sandi Klavžar[§]

Institute of Organic Chemistry, The Hebrew University of Jerusalem, Jerusalem 91904, Israel, and Department of Mathematics, PF, University of Maribor, Koroška cesta 160, 62000 Maribor, Slovenia

Received April 24, 1995[⊗]

An algorithm is designed for the calculation of the Szeged index of benzenoid hydrocarbons, based on the examination of their elementary cuts. The method requires the finding of all elementary cuts of a benzenoid system and counting the vertices lying on each side of these cuts—a task significantly simpler than the calculation of the Szeged index directly from its definition.

1. INTRODUCTION

The Szeged index (Sz) is a recently proposed[1,2] structural descriptor, based on the distances of the vertices of the molecular graph. In order to be able to study the properties of this novel topological index it would be advantageous to possess an easy method for its calculation. The calculation of Sz directly from its definition (see below) is quite cumbersome, especially in the case of large polycyclic molecules. In this paper we put forward a procedure for computing Sz that is significantly simpler. The algorithm, in the form elaborated in this paper, applies to benzenoid molecules, but its extension to other types of polycyclic systems would easily be possible.

THE SZEGED INDEX

Let G be a connected graph and x and y two of its vertices. Let $d(x,y|G)$ be the distance between x and y, i.e., the number of edges in a shortest path that connects x and y.

Let e be an edge of the graph G, connecting the vertices u and v. Define two sets $\mathcal{N}_1(e|G)$ and $\mathcal{N}_2(e|G)$ as

$$\mathcal{N}_1(e|G) = \{x|x \in \mathbf{V}(G),\quad d(x,u|G) < d(x,v|G)\}$$

$$\mathcal{N}_2(e|G) = \{x|x \in \mathbf{V}(G),\quad d(x,v|G) < d(x,u|G)\}$$

and denote by $n_1(e) = n_1(e|G)$ and $n_2(e) = n_2(e|G)$ their cardinalities (= number of elements). In other words, $n_1(e|G)$ is the number of vertices closer to u than to v, and $n_2(e|G)$ is the number of vertices closer to v than to u; vertices equidistant to u and v are not counted.

The Szeged index is defined as[1,2]

$$\mathrm{Sz} = \mathrm{Sz}(G) = \sum_{e} n_1(e|G)n_2(e|G) \tag{1}$$

where the summation goes over all edges of the graph G.

In the case of bipartite graphs there are no vertices equidistant to the both ends of an edge and therefore for all edges e of a bipartite graph G

$$n_2(e|G) + n_2(e|G) = n(G) = \text{number of vertices of } G \tag{2}$$

Notice that eq 2 applies, in particular, to benzenoid systems.

For further details concerning the definition of the Szeged index, for the reasons to introduce a new topological index via eq 1 as well as for the basic properties of Sz see the previous papers.[1,2] In the article[2] also the correct pronunciation of "Szeged" and "Sz" is explained.

BENZENOID SYSTEMS AND THEIR NORMAL CUTS

Throughout this paper the term *benzenoid system* (or *benzenoid graph*) is used for graphs constructed in the following manner.[3] Consider the hexagonal (graphite) lattice $\mathcal{H}$. Let Z be a circuit on this lattice. Then benzenoid systems are formed by the vertices and edges of $\mathcal{H}$, lying on some circuit Z or in its interior. The vertices and edges belonging to Z form the *perimeter* (sometimes called boundary) of the respective benzenoid system. The vertices (if any) not belonging to the perimeter are said to be the *internal vertices* of the respective benzenoid system.

The number of vertices, internal vertices, edges, and hexagons of a benzenoid system B will be denoted by[3] $n = n(B)$, $n_i = n_i(B)$, $m = m(B)$, and $h = h(B)$, respectively. The following relations between these parameters are well-known[3]

$$n = 4h + 2 - n_i \tag{3}$$

$$m = 5h + 1 - n_i \tag{4}$$

In Figure 1 we illustrate the above definition on the case of the benzenoid graph B_0 for which $h = 8$, $n_i = 6$, $n = 28$, $m = 35$.

If a benzenoid system is considered as a geometric figure in the plane then an *elementary cut* is defined as follows. Choose an edge e of the benzenoid system and draw a straight line through the center of e, orthogonal on e. This line will intersect the perimeter in two[4] points P_1 and P_2. The straight line segment C whose end points are P_1 and P_2 is the elementary cut, intersecting the edge e. Clearly, C intersects not only the edge e but also all edges that lie between P_1 and P_2 (inclusive the two edges on the perimeter to which the points P_1 and P_2 belong).

In Figure 2 the elementary-cut-concept is illustrated on the example of the benzenoid system B_0.

* Address correspondence to the author at Institute of Physical Chemistry, Attila Jozsef University, P.O. Box 105, H-6701, Szeged, Hungary.
† On leave from the Faculty of Science, University of Kragujevac, Yugoslavia.
‡ The Hebrew University of Jerusalem.
§ University of Maribor.
⊗ Abstract published in *Advance ACS Abstracts,* September 1, 1995.

0095-2338/95/1635-1011$09.00/0

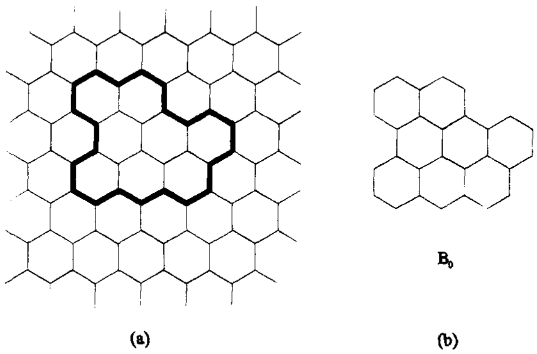

Figure 1. (a) The hexagonal lattice and a circuit Z on it. (B) The benzenoid system B_0 determined by the circuit Z; B_0 has six internal vertices and its perimeter is of size 22.

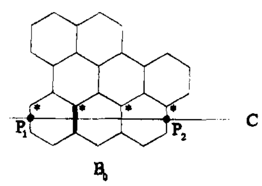

Figure 2. The elementary cut C corresponding to an edge of the benzenoid system B_0, indicated by a heavy line; observe that C is an elementary cut intersecting the edges of B_0 marked by asterisks.

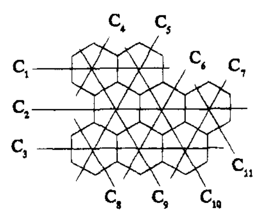

Figure 3. All elementary cuts of the benzenoid system B_0; $\mathbf{C}(B_0) = \{C_1, C_2, ..., C_{11}\}$.

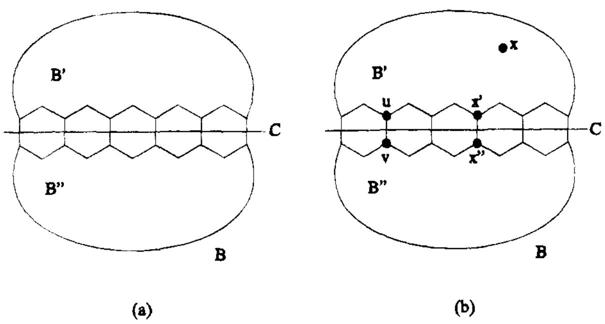

Figure 4. Structural details needed for the formulation and proof of Lemma 1.

Elementary cuts play an important role in the theory of benzenoid systems.[3,5,6]

The set of elementary cuts of a benzenoid system B, that involves all the edges of B, will be called a *complete set of elementary cuts* (CSEC) and will be denoted by $\mathbf{C} = \mathbf{C}(B)$. It should be noticed that $\mathbf{C}(B)$ possesses much fewer elements that there are edges in B. For instance, B_0 has 35 edges, but its CSEC possesses only 11 elements (see Figure 3).

ELEMENTARY CUTS AND THE NUMBERS $n_1(e|B)$ and $n_2(e|B)$

If C is an elementary cut of the benzenoid system B, then by deleting from B the edges intersected by C, we obtain a subgraph of B consisting of two disconnected parts, B' and B'' (see Figure 4a). These subgraphs may, but need not, be benzenoid systems themselves.

Let $n(B')$ and $n(B'')$ be the number of vertices of the fragments B' and B'', respectively. Of course,

$$n(B') + n(B'') = n(B) \tag{5}$$

Let e be an edge intersected by C, that connects the vertices u and v (see Figure 4b). Then we have the following simple, but important auxiliary result:[6]

Lemma 1. $n_1(e|B) = n(B')$ and $n_2(e|B) = n(B'')$.

Proof. Suppose there is a vertex x in B' which lies closer to v than to u (see Figure 4b). Then a shortest path between x and v must intersect C at an edge different from $e = (u,v)$ and, clearly, there can be only one such edge. Let this be the edge $e' = (x',x'')$, see Figure 4b. Then

$$\mathrm{d}(x,v|B) = \mathrm{d}(x,x'|B') + \mathrm{d}(x'',v|B'') + 1$$

On the other hand, $\mathrm{d}(x'',v|B'') = \mathrm{d}(x',u|B')$ because the shortest paths between x'' and v as well as between x' and u evidently go along the elementary cut C (see Figure 4b). From Figure 4b is also seen that $\mathrm{d}(x,x'|B') + \mathrm{d}(x',u|B') \geq \mathrm{d}(x,u|B')$. Therefore

$$\mathrm{d}(x,v|B) = \mathrm{d}(x,x'|B') + \mathrm{d}(x',u|B') + 1 > \mathrm{d}(x,u|B)$$

which is in contradiction with the supposition that x lies closer to v than to u. Hence any vertex in B' lies closer to u than to v and, consequently

$$n_1(e|B) \geq n(B') \tag{6}$$

In an analogous manner we conclude that

$$n_2(e|B) \geq n(B'') \tag{7}$$

Bearing in mind eqs 2 and 5 we immediately see that in (6) and (7) only the equality sign can occur. By this we arrive at Lemma 1.

Lemma 1 applies to all edges that are intersected by the elementary cut C. This observation immediately yields the following:

Lemma 2. Let $e_1, e_2, ..., e_r$, $r \geq 2$, be the edges of a benzenoid system B, intersected by the elementary cut C. Then for all $i = 1,2,...,r$

$$n_1(e_i|B) = n(B') \quad \text{and} \quad n_2(e_i|B) = n(B'')$$

where B' and B'' are the fragments of B, depicted in Figure 4a.

THE ALGORITHM

In view of Lemma 2, an algorithm for the calculation of the Szeged index, eq 1, of benzenoid systems is readily conceived.

Let C be an elementary cut that divides the benzenoid system B into components $B'(C)$ and $B''(C)$. Let C intersects $r(C)$ distinct edges of B. Then

$$\mathrm{Sz}(B) = \sum_C r(C)n(B'(C))n(B''(C)) \tag{8}$$

with the summation going over the CSEC of B. When applying formula (8) we need to count only the vertices of the fragment B', because the number of vertices of B'' is determined via eq 5.

As an illustration of our algorithm we calculate $Sz(B_0)$ using the CSEC from Figure 3. Recall that $n(B''_0) = n(B_0) - n(B'_0) = 28 - n(B'_0)$.

elementary cut C	r	$n(B'_0)$	$n(B''_0)$	$r\,n(B')n(B'')$
C_1	3	5	23	$345 = 3 \times 5 \times 23$
C_2	4	13	15	780
C_3	4	7	21	588
C_4	2	3	25	150
C_5	4	10	18	720
C_6	3	11	17	561
C_7	3	5	23	345
C_8	2	3	25	150
C_9	4	10	18	720
C_{10}	4	10	18	720
C_{11}	2	3	25	150
				total: 5229

Hence, $Sz(B) = 5229$.

Needless to say that we may always choose for B' the smaller of the two fragments obtained by dissecting B along the elementary cut C.

In large benzenoid system the counting of the vertices of the fragments B' and B'' may become somewhat tedious and error prone. Formula (8) can be further simplified in the following manner: instead of counting the vertices of B' and B'' we may count the hexagons and the internal vertices of B, lying in B' and B''.

Let B' and B'' possess $h(B')$ and $h(B'')$ hexagons, respectively. Let $n_i(B')$ and $n_i(B'')$ be the number of internal vertices of the benzenoid system B that belong to B' and B'', respectively. Then

$$h(B') + h(B'') + r - 1 = h \tag{9}$$

$$n_i(B') + n_i(B'') = n_i \tag{10}$$

where h and n_i refer to the benzenoid system B and where $r = r(C)$ is the number of edges connecting B' and B'', intersected by the elementary cut C. Bearing in mind eq 3, we arrive at

$$n(B') = 4h(B') + 2r - 1 - n_i(B') \tag{11}$$

$$n(B'') = 4h(B'') + 2r - 1 - n_i(B'') \tag{12}$$

In order to obtain eq 11 consider the (true) benzenoid system which embraces the hexagons of B' and the $r - 1$ hexagons intersected by C. This benzenoid system has $h(B') + r - 1$ hexagons, $n_i(B')$ internal vertices, and $2r - 1$ vertices more than B'. Equation 3 has to be applied to it, and then the number of vertices has to be diminished by $2r - 1$ in order to obtain $n(B')$. Equation 12 is deduced in a fully analogous manner.

When eqs 11 and 12 are substituted back into (8), we obtain after a straightforward, but lengthy calculation

$$Sz(B) = \sum r[4h(B') - n_i(B')][4h(B'') - n_i(B'')] - (4h - n_i + 3)\sum r + 2(4h + 4 - n_i)\sum r^2 - 4\sum r^3 \tag{13}$$

All the four summations on the right-hand side of (13) go over the CSEC of the benzenoid system B. Observe that $\sum r = m$, and therefore using eq 4

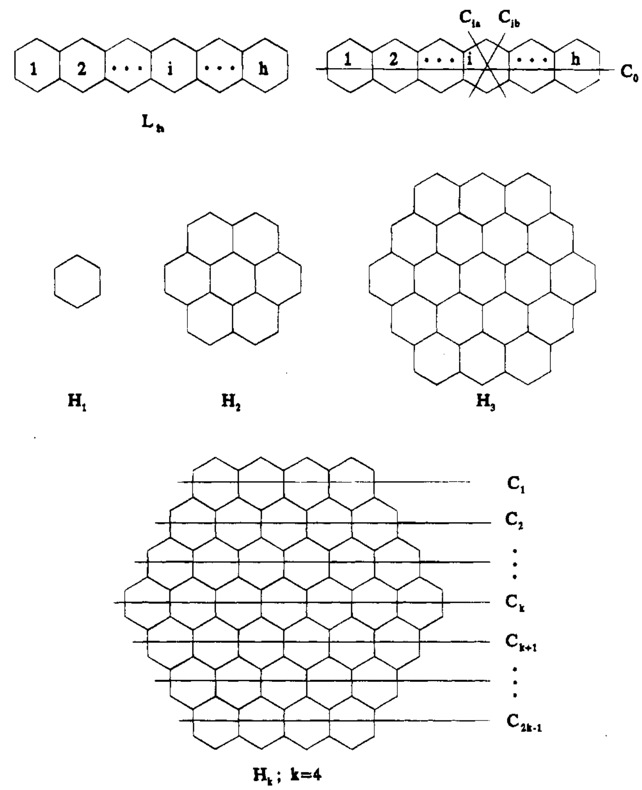

Figure 5. The polyacene (L_h) and the coronene/circumcoronene series (H_h) and their elementary cuts.

$$Sz(B) = \sum r[4h(B') - n_i(B')][4h(B'') - n_i(B'')] - (4h - n_i + 3)\cdot(5h + 1 - n_i) + 2(4h + 4 - n_i)\sum r^2 - 4\sum r^3 \tag{14}$$

Now, although (13) and (14) have a seemingly more complicated form than eq 8, the finding of $h(B')$ and $n_i(B')$ is usually an easier task than the counting of the vertices of B'. As before, it is not necessary to independently search for $h(B'')$ and $n_i(B'')$, because these quantities are directly obtainable from eqs 9 and 10.

In the case of catacondensed benzenoid systems, which are characterized by the condition[3] $n_i = 0$, formulas 13 and 14 are significantly simplified

$$Sz(B) = 16\sum rh(B')h(B'') - (4h + 3)\sum r + 8(h + 1)\sum r^2 - 4\sum r^3 = 16\sum rh(B')h(B'') - (4h + 3)(5h + 1) + 8(h + 1)\sum r^2 - 4\sum r^3$$

TWO APPLICATIONS

To further illustrate the efficiency of our algorithm, we find general expressions for the Szeged index of the polyacenes (L_h) and of the members of the coronene/circumcoronene series (H_k). The respective structures and elementary cuts are depicted in Figure 5.

Polyacenes. Using the notation from Figure 5 we immediately see that the CSEC of the polyacenes is given by $\mathbf{C}(L_h) = \{C_0, C_{ia}, C_{ib} | i = 1,2,...,h\}$. For C_0: $r = h + 1$, $n(B') = n(B'') = 2h + 1$. For both C_{ia} and C_{ib}: $r = 2$, $n(B') = 4(i - 1) + 3 = 4i - 1$, $n(B'') = 4h + 2 - (4i - 1) = 4h - 4i + 3$. Applying eq 8 we get

$$\mathrm{Sz}(L_h) = (h+1)(2h+1)^2 + 2\sum_{i=1}^{h} 2(4i-1)(4h-4i+3)$$

which by direct calculation yields

$$\mathrm{Sz}(L_h) = \frac{1}{3}(44h^3 + 72h^2 + 43h + 3)$$

The Coronene/Circumcoronene Series. In Figure 5 are indicated only the $2k + 1$ horizontal elementary cuts of H_k. There exist two additional groups of $2k + 1$ symmetry-equivalent elementary cuts, obtained by rotating the former group by +60° and by −60°. Therefore, if one applies eq 8 to only the horizontal elementary cuts, the result will be equal to $^1/_3\ \mathrm{Sz}(H_k)$.

It should also be observed that because of symmetry, the contribution to the right-hand side of (8) of the elementary cut C_i is equal to the contribution of C_{2k-i}, $i = 1,2,...,k-1$.

It can be shown that $n(H_k) = 6k^2$ and that for the elementary cut C_i: $r = k + i$, $n(B') = i(2k + i)$, $i = 1,2,...,k$.

Taking into account all the above properties of the elementary cuts of H_k and using eq 8 we obtain

$$\frac{1}{3}\mathrm{Sz}(H_k) = 18k^5 + 2\sum_{i=1}^{k-1}(k+i)i(2k+i)[6k^2 - i(2k+i)]$$

After a lengthy calculation this results in the formula

$$\mathrm{Sz}(H_k) = \frac{3}{2}k^2(36k^4 - k^2 + 1)$$

ACKNOWLEDGMENT

This research was supported in part by the Mathematical Institute in Belgarde (I.G.) and by the Ministry of Science and Technology of Slovenia (S.K.). In Spring 1995 I.G. was a Lady Davis Fellow and he gratefully acknowledges the support by the Lady Davis Fellowship Trust which enabled him to stay at the Hebrew University in Jerusalem, Israel.

REFERENCES AND NOTES

(1) Gutman, I. A Formula for the Wiener Number of Trees and Its Extension to Graphs Containing Cycles. *Graph Theory Notes New York* **1994**, *27*, 9−15.
(2) Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Dobrynin, A.; Gutman, I.; Dömötör, G. The Szeged index and an analogy with the Wiener index. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 547.
(3) Gutman, I.; Cyvin, S. J. *Introduction to the Theory of Benzenoid Hydrocarbons*; Springer-Verlag: Berlin, 1989.
(4) It may happen that the straight line intersects the perimeter in more than two points. Then we choose as P_1 and P_2 the intersection points lying on opposite sides of, and being nearest to, the edge *e*.
(5) Sachs, H. Perfect matchings in hexagonal systems. *Combinatorica* **1984**, *4*, 89−99.
(6) Klavžar, S.; Gutman, I.; Mohar, B. Labeling of benzenoid systems which reflects the vertex-distance relations. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 590.

CI9500391