CROATICA CHEMICA ACTA CCACAA **70** (3) 863–871 (1997)

ISSN–0011-1643
CCA–2460

Original Scientific Paper

Szeged Index: Formulas For Fused Bicyclic Graphs

István Lukovits

Central Research Institute for Chemistry, Hungarian Academy of Sciences, P.O.Box 17, H-1525 Budapest, Hungary.

Received January 2, 1997; revised April 1, 1997; accepted April 2, 1997

The Szeged index (Sz) is a variant of the well known Wiener index W. It has been shown that $Sz(\mathrm{B}_N)$ is a third order polynomial in terms of the sizes of three chains making up a fused bicyclic graph B_N. Analytical formulas have been derived for $Sz(\mathrm{B}_N)$.

INTRODUCTION

The Szeged index (Sz) is a graph invariant which was introduced by Gutman.[1] Its definition is based on the original definition of the well known Wiener index.[2] According to this definition for an acyclic graph T (where T refers to a tree, in Wiener's paper an acyclic hydrocarbon), $W(\mathrm{T})$ is equal to the sum of »bond contributions«:

$$W(\mathrm{T}) = \sum_{(a,b)} n_a \times n_b \tag{1}$$

where a and b are adjacent vertices of T and (a,b) denotes the edge connecting vertices a and b, n_a and n_b denote the number of vertices in the respective disconnected graphs obtained by removing edge (a,b) from T. Only carbons are considered, hydrogens are neglected. The definition is valid for trees only.

The original definition given by Wiener could be extended to any graph G, including cycle-containing graphs, by using the fact (which was also found out by Wiener) that $W(\mathrm{G})$ is equal to the sum of distances between all pairs of vertices.

$$W(\mathrm{G}) = \sum_{(i<j)} D_{i,j}. \tag{2}$$

Eq. (2) is Hosoya's formula,[3] where $D_{i,j}$ denotes the length of the shortest path between vertices i and j. The entries $D_{i,i}$ are zero.

Since Eq. (2) may be used for cycle-containing structures, and because of its simplicity, Hosoya's equation, Eq. (2), is used when W has to be calculated. It has to be noted that, many years after Wiener's paper had appeared, the »bond-contribution« picture (Eq. (1)) could be extended for cycle-containing structures, too.[4–6] It was shown that W can be computed as a sum of bond contributions (*i.e.* $W = \Sigma\, W_e$, where W_e denotes the contribution of bond e and the summation has to be performed for all bonds), and each bond-contribution term is equal to

$$W_e = \sum_{i \leq j} K^e_{ij} / K_{ij} \tag{3}$$

where K^e_{ij} denotes the number of shortest paths between vertices i and j which contain e and K_{ij} is the total number of the shortest paths between i and j.

Gutman[1] proposed a different interpretation of the »bond-contribution« approach (Eq. (1)). In his approach Eq. (1) remains valid but the definitions of n_a and n_b are replaced by the following formulas:

$$n_a = |\{x | D_{x,a} < D_{x,b}\}| \tag{4}$$

and

$$n_b = |\{x | D_{x,b} < D_{x,a}\}|. \tag{5}$$

According to Eq. (4), n_a is the number of vertices (including vertex a itself) nearer to vertex a than to b, and a similar definition holds for n_b (Eq. (5)). Note that vertices x for which $D_{x,a} = D_{x,b}$ will not be considered.

The third variant of W, being equivalent to W in trees but different from it in cycle-containing graphs, is the detour index[7–9] w. w may be obtained by using an analogue of Eq. (2) in which the distances $D_{i,j}$ are replaced by entries $\Delta_{i,j}$, *i.e.* by the lengths of the *longest* paths between vertices i and j.

The fourth possibility of extending W to cyclic structures might be accomplished by observing the fact that in trees the sum of the positive eigenvalues of the Laplacian matrix $\boldsymbol{L}$ is equal[10] to W. Therefore the new invariant, being the sum of positive eigenvalues of $\boldsymbol{L}$(G), would be equivalent to W in trees, but would be different from it in cycle-containing graphs.

In 1994, Randić introduced[11] the »hyper-Wiener index« WW by modifying Wiener's definition. Originally, WW could only be obtained for acyclic graphs but it was later shown that WW may be obtained for cycle-containing graphs G, too, by using the following formula:[12]

$$WW(G) = 1/2 \sum_{i \leq j} (D_{i,j} + D_{i,j}^2). \tag{6}$$

Later on, the »hyper« variants of Sz and w were also introduced.[13,14]

A further variant of W was proposed by Rouvray,[15] but it was found later on that this index is equal to $2W$. An expanded version of W was introduced by Tratch *et al.*,[16] in which for each pair of vertices i and j, the distance $D_{i,j}$ is multiplied by the number of »superpaths« containing the corresponding shortest path between i and j, and the products are added. In cycle-containing structures each product has to be multiplied by the number of shortest i–j paths. A special version of W is based on the »resistance distance«[17] model and another one on asymmetric matrices that were obtained from the distance matrix.[18] The mathematical aspects and numerous applications of W have been reviewed several times.[19–22] A special review was devoted to the distance matrix.[23]

Because of the importance of W, other closely related graph invariants, like the Szeged (and the detour) index, are also interesting. In fact, many interesting mathematical properties of Sz have been discovered thus far.[24–29] In particular, it has been demonstrated that $Sz(G) = W(G)$ if and only if every block of G is a complete graph.[29] The aim of the present paper is to show that Sz is a third order polynomial in terms of the sizes of chains making up a fused bicyclic graph. Formulas for N vertex fused bicyclic graphs B_N (Figure 1) have been derived. Sz is also in this respect related to W, w, WW and the hyper detour index, for which analytical formulas were obtained earlier and these are also polynomials.[14,30–37]

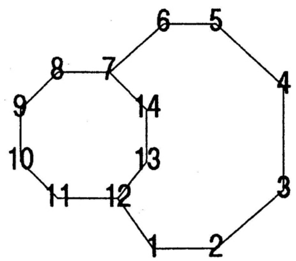

Figure 1. An example of a fused bicyclic graph with $k = 8$, $m = 6$ and $n = 4$.

DERIVATION OF THE POLYNOMIAL

Expressions like »graph« and »structural formula«, »valence« and »degree« are synonyms. Therefore, corresponding expressions used in chemistry and in graph theory, like »chemical bond« and »edge«, »atom« and »vertex«

will be used interchangeably hereafter. A »graph invariant« is the number which depends on graph G but does not depend on the numbering of the vertices in G. Hydrogen suppressed graphs will be considered only.

First, it has to be shown that $Sz(B_N)$ is a polynomial. Consider an example of a fused bicyclic graph B_N (Figure 1). There are three chains – subgraphs of B_N–k, m and n, which connect the branching vertices. k, m and n include the branching vertices. Examples: 1. The (hydrogen suppressed) graph of ethane is a single chain consisting of two vertices (endpoints), 2. The (hydrogen suppressed) graph of isobutane consists of three chains, each containing two vertices (the branching atom and the endpoint). Since branching vertices are considered several times in B_N, the total number of vertices N is therefore equal to $k+m+n-4$. Note that throughout this paper letters k, m and n will be used to denote the respective chains *and* the number of vertices (which will be referred to as the »size« of the chain) in k, m and n, respectively. Throughout this paper it will be assumed that $k \geq m \geq n$.

Figure 1 illustrates a special case but the procedure can easily be applied to any B_N. Bonds (12,1), (1,2), (2,3) and (3,4) will be considered only; a similar procedure could be used for the rest of the bonds of chain k and for m and n. Observe that $k+n-2$ denotes the size of a cycle. Two cases will be distinguished.

Case 1: $i+n-1 < (k+n-2)/2$, where the running index i refers to the first figures in the parentheses. Then for bonds (12,1) and (1,2): $n_a = (k+n-2)/2+m-2 = 9$ and $n_b = (k+n-2)/2 = 5$, where n_a refers to the first figure in the parentheses and n_b to the second.

Case 2: $i+n-1 \geq (k+n-2)/2$. For bonds (2,3) and (3,4), the following expressions are valid: $n_a = (2k+m+n-4)/2-i-1$, and $n_b = (2k+m+n-4)/2-(k-i-1)$.

In general, in the first case the first vertex is nearer to the both branching vertices than the second, and in the second case the second vertex is nearer to one of the branching vertices than the first one.

The product $n_a n_b$ is a second order polynomial in terms of k, m, n and i. The sum of these products is a third order polynomial in terms of numbers k, m and n and the running index i drops out. $Sz(B_N)$ is therefore a polynomial in terms of k, m and n. Although the derivation was given for a special case, where k, m and n are even, the procedure can be performed for other cases as well. The following theorem ensures that $Sz(B_N)$ is always a third order polynomial.

THEOREM. $$Sz(B_N) \leq N^2(N+1)/4 \tag{7}$$

Proof. The proof uses a result by Dobrynin[38] according to which for any graph G, and therefore also for B_N:

$$Sz(\mathrm{B}_N) = 0.25\,[\sum_{(a,b)} (N-n_{ab})^2 - \sum_{(a,b)} (n_a-n_b)^2] \tag{8}$$

where n_{ab} is the number of vertices being equally far from a and b. But:

$$Sz(\mathrm{B}_N) \leq 0.25 \sum_{(a,b)} (N-n_{ab})^2 \leq 0.25 \sum_{(a,b)} N^2 = 0.25N^2(N+1). \tag{9}$$

Q.E.D.

Because of this result, $Sz(\mathrm{B}_N)$ cannot be a fourth order polynomial in terms of *k, m* and *n*, unlike the Szeged index of a complete bipartite graph which is a fourth order polynomial[38] in terms of *N*.

Derivation of the formulas $Sz(\mathrm{B}_N)$ in terms of *k, m* and *n* was done by solving a system of 20 simultaneous linear equations with 20 unknowns (*A, B...T*).

$$Sz(\mathrm{B}_N) = Ak^3 + Bm^3 + Cn^3 + Dk^2m + Ek^2n + Fm^2n + \\ + Gm^2k + Hn^2k + In^2m + Jkmn + Kk^2 + Lm^2 + Mn^2 + \\ + Nkm + Okn + Pmn + Qk + Rm + Sn + T. \tag{10}$$

The equations were solved for known values of $Sz(\mathrm{B}_N)$ and by using the values of k^3, m^3, n^3, k^2m, *etc.* and of course the result cannot depend on the special choice of *k, m* and *n* values. Quite often, however, the matrix composed of powers of *k, m* and *n* is singular, therefore we have listed those combinations of the values of *k, m* and *n*, which did not yield a singular matrix (Table I). The values of *Sz* were obtained by using an algorithm similar to that proposed by Žerovnik.[39]

Case I. *k, m* and *n* are even, or *k, m* and *n* are odd.

$$Sz(\mathrm{B}_N) = (k^3 + m^3 + 2k^2m + 4k^2n + 4m^2n + 2m^2k + 3n^2k + \\ + 3n^2m + 6kmn - 11k^2 - 11m^2 - 8n^2 - 14km - 26kn - 26mn + \\ + 36k + 36m + 46n - 48)/4. \tag{11}$$

Case II. *k* and *m* are even and *n* is odd, or *k* and *m* are odd and *n* is even.

$$Sz(\mathrm{B}_N) = (k^3 + m^3 + 2k^2m + 4k^2n + 4m^2n + 2m^2k + 3n^2k + \\ + 3n^2m + 6kmn - 14k^2 - 14m^2 - 6n^2 - 22km - 32kn - 32mn + \\ + 65k + 65m + 60n - 102)/4. \tag{12}$$

Case III. *k* is even and *m* and *n* are odd or *k* is odd, and *m* and *n* are even.

TABLE I

List of paths and the respective values of Sz used to derive formulas 11–14.

Case 1				Case 2			
k	m	n	Szeged index	k	m	n	Szeged index
2	2	2	3	5	5	4	198
4	2	2	20	7	5	2	166
6	4	2	132	9	7	4	819
4	4	2	59	9	9	8	2352
6	2	2	63	7	7	6	875
8	4	2	251	9	5	4	562
8	2	2	144	7	3	2	89
10	4	4	723	11	9	4	1591
10	6	2	639	11	5	4	855
8	6	2	408	9	7	6	1252
6	4	4	267	5	5	2	81
6	6	4	446	7	7	4	546
8	6	4	695	9	9	4	1158
8	8	2	627	11	11	2	1401
10	8	6	2002	13	5	4	1238
8	4	4	458	13	7	4	1635
6	6	6	705	11	7	6	1731
8	8	4	1010	13	7	2	1054
8	6	6	1050	3	3	2	9
8	8	8	2016	9	9	6	1707

Case 3				Case 4			
k	m	n	Szeged index	k	m	n	Szeged index
8	5	3	359	9	6	5	897
10	7	5	1272	11	8	5	1749
8	7	7	1274	9	8	7	1769
6	5	5	361	11	10	5	2304
4	3	3	40	5	4	3	103
8	7	5	885	13	10	5	2997
8	3	3	216	13	6	5	1809
10	9	3	1192	11	8	3	1208
10	5	5	909	7	6	5	588
8	7	3	564	9	4	3	379
6	5	3	204	15	8	5	3069
10	9	5	1733	15	10	3	2845
12	7	3	1212	15	6	3	1749
12	3	3	624	13	10	3	2180
12	5	5	1318	17	6	5	3193
14	5	3	1268	17	8	3	2939
10	7	7	1779	11	8	7	2382
12	7	5	1761	13	12	3	2799
10	9	7	2366	11	10	7	3079
10	9	9	3091	11	10	9	3958

$$Sz(B_N) = (k^3 + m^3 + 2k^2m + 4k^2n + 4m^2n + 2m^2k + 3n^2k + 3n^2m + 6kmn - 13k^2 - 12m^2 - 11n^2 - 18km - 32kn - 34mn + 55k + 56m + 76n - 99)/4. \quad (13)$$

Case IV. k and n are even and m is odd or k and n are odd and m is even.

$$Sz(B_N) = (k^3 + m^3 + 2k^2m + 4k^2n + 4m^2n + 2m^2k + 3n^2k + 3n^2m + 6kmn - 12k^2 - 13m^2 - 11n^2 - 18km - 34kn - 32mn + 56k + 55m + 76n - 99)/4. \quad (14)$$

DISCUSSION

All equations were derived by using one set of conditions imposed on k, m and n (*e.g.* even values of k, m and n were used to derive Eq. (11)), but they remain valid for the second set of conditions as well (in case I, for k, m and n are odd). Similarly, Eqs. (12)–(14) are also valid for cases fulfilling the respective second sets of conditions. Example: calculate $Sz(k,m,n)$ for $k = 9$, $m = 7$ and $n = 5$. By using Eq. (11), we obtain

$$Sz(9,7,5) = (9^3 + 7^3 + 2*9^2*7 + 4*9^2*5 + 4*7^2*5 + 2*7^2*9 + 3*5^2*9 + 3*5^2*7 + 6*9*7*5 - 11*9^2 - 11*7^2 - 8*5^2 - 14*9*7 - 26*9*5 - 26*7*5 + 36*9 + 36*7 + 46*5 - 48)/4 = (729 + 343 + 1134 + 1620 + 980 + 882 + 675 + 525 + 1890 - 891 - 539 - 200 - 882 - 1170 - 910 + 324 + 252 + 230 - 48)/4 = 4944/4 = 1236.$$

Observe that restrictions imposed in Case I and Case II are equivalent with the statements that the number of vertices is even or odd in both cycles, respectively. In Case III the number of vertices ($k+n-2$) is odd in the larger cycle and even ($m+n-2$) in the smaller cycle, and in Case IV. $k+n-2$ is even and $m+n-2$ is odd.

Eq. (11) remains valid if $k = m = n = 1$, and $Sz(1,1,1) = 0$. This graph corresponds to a »hypothetical methane« with a single vertex and three edges starting and ending at this vertex. Because $Sz(1,1,1) = 0$, the sum of the coefficients of Eq. 11 must be zero. In fact 1+1+2+4+4+2+3+3+6–11–11–8–14–26–26+36+36+46–48 = 0.

Comparison of the coefficients of Eqs. (11)–(14) shows that the coefficients of the third order terms are equal. In contrast, in equations obtained for $W(B_N)$ and $w(B_N)$, the coefficients of the second order terms are also equal.[33,35] Note that n^3 does not appear in Eqs. (11)–(14).

A formula for $Sz(C_N)$, where C_N denotes an N vertex cycle, may be derived easily:

$$Sz(C_N) = N(N-1)^2/4 \qquad (N \text{ is odd}) \qquad (15)$$

and

$$Sz(C_N) = N^3/4 \qquad (N \text{ is even}). \qquad (16)$$

Comparison of these formulas with Eqs. (11) and (13) reveals that $Sz(C_N)$ is not equal to $Sz(B_N)$ with $k = N$, and $m = n = 2$, whereas $W(C_N) = W(N,2,2)$, and also $w(C_N) = w(N,2,2)$.[33,35]

Eqs. (11)–(14) can easily be programmed. The result is an example of algebraic expressions derived by using a numerical method. Formulas of graph invariants may be applied to solve – at least within a given set of structures – the inverse »QSPR problem«, namely to find structures corresponding to a numerical value of graph invariant.[40,41]

REFERENCES

1. I. Gutman, *Graph Theory Notes (New York)* **27** (1994) 9.
2. H. Wiener, *J. Am. Chem. Soc.* **69** (1947) 17.
3. H. Hosoya, *Bull. Chem. Soc. Japan* **44** (1971) 2332.
4. I. Lukovits, *Quant. Struct.-Act. Relat.* **9** (1990) 227.
5. I. Lukovits and I. Gutman, *Commun. Math. Chem.* **31** (1994) 133.
6. T. Pisanski and J. Žerovnik, *J. Chem. Inf. Comput. Sci.* **34** (1994) 395.
7. D. Amić and N. Trinajstić, *Croat. Chem. Acta* **68** (1995) 53.
8. P. John, *Commun. Math. Chem.* **32**, (1995) 207.
9. O. Ivanciuc and A. Balaban, *Commun. Math. Chem.* **30** (1994) 141.
10. N. Trinajstić, D. Babić, S. Nikolić, D. Plavić, D. Amić, and Z. Mihalić, *J. Chem. Inf. Comput. Sci.* **34** (1994) 368.
11. M. Randić, *Chem. Phys. Lett.* **211** (1993) 478.
12. D. J. Klein, I. Lukovits, and I. Gutman, *J. Chem. Inf. Comput. Sci.* **35** (1995) 50.
13. M. V. Diudea, O. M. Minailiuc, G. Katona, and I. Gutman, *Commun. Math. Chem.* **35** (1997) 000.
14. I. Lukovits, *Croat. Chem. Acta* **69** (1996) 873.
15. D. Rouvray, *Commun. Math. Chem.* **1** (1975) 125.
16. S. S. Tratch, M. I. Stankevich, and N. S. Zefirov, *J. Comput. Chem.* **11** (1990) 899.
17. D. J. Klein and M. Randić, *J. Math. Chem.* **12** (1993) 81.
18. M. V. Diudea, *J. Chem. Inf. Comput. Sci.* **36** (1996) 535.
19. I. Gutman, Y. N. Yeh, S. L. Lee, and Y. L. Luo, *Indian J. Chem.* **32A** (1993) 651.
20. S. Nikolić, N. Trinajstić, and Z. Mihalić, *Croat. Chem. Acta* **68** (1995) 105.
21. H. Y. Zhu, D. J. Klein, and I. Lukovits, *J. Chem. Inf. Comput. Sci.* **36** (1996) 420.
22. A. Graovac and T. Pisanski, *J. Math. Chem.* **8** (1991) 53.
23. Z. Mihalić, S. Nikolić, and N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **32** (1992) 28.

24. P. V. Khadikar, N. V. Deshpande, P. P. Kale, A. A. Dobrynin, I. Gutman, and Gy. Dömötör, *J. Chem. Inf. Comput. Sci.* **35** (1995) 547.
25. I. Gutman, *Bull. Acad. Serbe Sci. Arts (Ser. Math. Nature)* **111** (1996) 19.
26. A. A. Dobrynin and I. Gutman, *Croat. Chem. Acta.* **69** (1996) 845.
27. A. A. Dobrynin and I. Gutman, and G. Dömötör, *Appl. Math. Lett.* **8** (1995) 57.
28. S. Klavžar, A. Rajapakse, and I. Gutman, *Appl. Math. Lett.* **9** (1996) 45.
29. A. A. Dobrynin, I. Gutman, *Graph Theory New York* **28** (1995) 21.
30. R. C. Entringer, D. E. Jackson, and D. A. Snyder, *Czech. Math. J.* **26** (1976) 283.
31. D. Bonchev, O. Mekenyan, and N. Trinajstić, *Int. J. Quantum Chem.* **17** (1980) 845.
32. I. Gutman, *J. Mol. Struct. (Theochem)* **285** (1993) 137.
33. I. Lukovits, *J. Chem. Inf. Comput. Sci.* **31** (1991) 503.
34. I. Lukovits, *J. Chem. Inf. Comput. Sci.* **34** (1994) 1079.
35. I. Lukovits and M. Razinger, *J. Chem. Inf. Comput. Sci.* **37** (1997) 283.
36. W. Linert and I. Lukovits, *Commun. Math. Chem.* **35** (1997) 65.
37. I. Lukovits, *Comput. & Chem.* **19** (1995) 27.
38. A. A. Dobrynin, *Croat. Chem. Acta* (to appear)
39. J. Žerovnik, *Croat. Chem. Acta* **69** (1996) 837.
40. L. B. Kier, L. H. Hall, and J. W. Frazer, *J. Chem. Inf. Comput. Sci.* **33** (1993) 143.
41. M. I. Skvortsova, I. I. Baskin, O. L. Slovokhotova, V. A. Palyulin, and N. S. Zefirov, *J. Chem. Inf. Comput. Sci.* **33** (1993) 630.

SAŽETAK

Szegedov indeks: formule za kondenzirane bicikličke grafove

István Lukovits

Szegedov indeks (Sz) varijanta je poznatog Wienerova indeksa *W*. Pokazano je da je $Sz(B_N)$ polinom trećeg reda u veličinama triju lanaca koji grade kondenzirani biciklički graf B_N. Izvedene su analitičke formule za $Sz(B_N)$.