communications in mathematical
match
and in computer chemistry

no. 35, March 1997

ISSN 0340-6253

MATCDY (35) 129-143 (1997)

Szeged Matrices and Related Numbers

Mircea V. Diudea, Ovidiu M. Minailiuc, Gabriel Katona
Department of Chemistry, "Babes-Bolyai" University, 3400 Cluj, Romania
Ivan Gutman
Institute of Physical Chemistry, Attila Jozsef University, P.O. Box 105, H-6701 Szeged, Hungary; permanent address: Faculty of Science, University of Kragujevac, YU-3400 Kragujevac, Yugoslavia

(Received April , 1996)

Abstract: A new unsymmetric matrix, $\mathbb{SZ}_u$ (unsymmetric Szeged matrix) is proposed by analogy to the matrix $\mathbb{CJ}_u$ (unsymmetric Cluj matrix) [1,2]. It is defined and exemplified both for acyclic and cycle-containing structures. Its relation with the $\mathbb{CJ}_u$ matrix is discussed. The derived Szeged numbers are compared to the Wiener matrix- and Cluj matrix-derived numbers and tested for separating and correlating ability on selected sets of graphs.

Introduction

The Wiener number, W, [3] is one of the most studied topological indices, both from theoretical point of view and applications (the reader can consult two recent reviews [4,5]). It is connected to the problem of distances in graph.

In acyclic structures, the Wiener number and its extension, hyper-Wiener number [6], can be defined as edge/path contributions, $I_{e/p}$, to a global number, I:

$$I = I(G) = \sum_{e/p} I_{e/p} = \sum_{e/p} N_{i,\,e/p} N_{j,\,e/p} \qquad (1)$$

where N_i and N_j denote the number of vertices lying on the two sides of the edge/path, **e/p** (having the endpoints **i** and **j**). Edge/path contributions, $I_{e/p}$ are just the entries in the Wiener matrices, $\mathbb{W}_e$ and $\mathbb{W}_p$ [7,8], from which I can be calculated as

$$I = (1/2) \sum_i \sum_j [\mathbb{W}_{e/p}]_{ij} \quad (2)$$

Within this paper, numbers (*i.e.* topological indices) are denoted by the corresponding boldface italic symbols, whereas matrices (and their entries) by special capital letters. By such a notation we try to suggest the different graph-theoretical properties that different matrices (providing identical Wiener-type numbers) collect (see below).

When is defined on edge (*i.e.* **(i,j)**∈**E(G)**, where **E(G)** is the set of edges in graph), the index I is a Wiener, W_e , number. When is defined on path, (*i.e.* **(i,j)**∈**P(G)**, with **P(G)** being the set of paths in graph) it is a hyper-Wiener, W_p , number (also denoted ***WW*** and ***R*** - see [7,8]). The numbers W_e and W_p count all "external" paths passing through the two endpoints of all edges/paths, **(i,j)** in graph.

Attempts have been made to extend the "edge contribution" definition (see eq **1**) to cycle-containing structures, [9-11] such as

$$I_e = (1/2) \sum_i \sum_j C^e_{ij} / C_{ij} \quad (3)$$

where C_{ij} is the number of the shortest paths between vertices **i** and **j**, and C^e_{ij} denotes the number of those shortest paths between **i** and **j** which contain the edge **e**. For the I_p contributions, Lukovits and Linert [9] have proposed a definition which resulted in a variant of hyper-Wiener number.

An alternative definition [12,13] of Wiener-type numbers is based on the distance matrix, $\mathbb{D}$

$$I = (1/2) \sum_i \sum_j [\mathbb{D}_{e/p}]_{ij} \quad (4)$$

where $\mathbb{D}_e$ is just the classical $\mathbb{D}$ matrix, whereas $\mathbb{D}_p$ is the "distance path" matrix [13]. The Wiener number, D_e , and the hyper-Wiener number, D_p , count all "internal" paths of length

|**e/p**| contained in all shortest paths in graph [1,13]. Note that, in acyclic structures $\boldsymbol{D_e}$=$\boldsymbol{W_e}$ and $\boldsymbol{D_p}$=$\boldsymbol{W_p}$.

$\mathbb{D}_p$ matrix is defined [13] as

$$[\mathbb{D}_P]_{ij} = \binom{[\mathbb{D}_e]_{ij} + 1}{2} \qquad \textbf{(5)}$$

and since it is derived from the $\mathbb{D}_e$ matrix, the definition (**4**) is valid both for acyclic and cyclic structures. The $\mathbb{D}_p$ matrix is illustrated in Figure **1**, for the graph **1** (1-Methyl,2-ethylcyclopropane).

Szeged Matrices

One of us has proposed a Wiener analogue, referred to as the Szeged number, SZ , [14-19]. It is defined according to eq **1** but the sets N_i and N_j are defined so that eq **1** remains valid both for acyclic and cycle-containing graphs

$$N_{i,e/p} = |\{\mathbf{u} \mid \mathbf{u} \in \mathbf{V(G)};\ [\mathbb{D}_e]_{iu} < [\mathbb{D}_e]_{ju}\}| \qquad \textbf{(6)}$$

$$N_{j,e/p} = |\{\mathbf{u} \mid \mathbf{u} \in \mathbf{V(G)};\ [\mathbb{D}_e]_{ju} < [\mathbb{D}_e]_{iu}\}| \qquad \textbf{(7)}$$

Thus, N_i and N_j represent the cardinality of the sets of vertices closer to i and to j, respectively; vertices equidistant to **i** and **j** are not counted.

By analogy to the Wiener matrices, the entries in the Szeged matrices can be defined as edge/path contributions,

$$[\mathbb{SZ}_{e/p}]_{ij} = N_{i,e/p} N_{j,e/p} \qquad \textbf{(8)}$$

and consequently the global index $\boldsymbol{I}$ can be written as

$$I = (1/2) \sum_i \sum_j [\mathbb{SZ}_{e/p}]_{ij} \tag{9}$$

When is defined on edge, the index I is a Szeged, SZ_e , number; when is defined on path, it is a hyper-Szeged, SZ_p , number and, finally, when path is larger than unity (*i.e.* $|\mathbf{p}|>1$) it is a Δ_{SZ} number (see Figure **1**).

By analogy to the Cluj matrix, $\mathbb{CJ}_u$, [1,2] a Szeged analogue matrix, $\mathbb{SZ}_u$, can be defined

$$[\mathbb{SZ}_u]_{ij} = \left|\left\{\mathbf{u} \;\middle|\; \mathbf{u} \in \mathbf{V(G)};\; [\mathbf{D}_e]_{iu} < [\mathbf{D}_e]_{ju}\right\}\right| \tag{10}$$

$\mathbb{SZ}_u$ is a square (in general) unsymmetric matrix, of dimensions **N**x**N**. It is constructed by the principle of unsymmetric (*i.e.* single point) characterization of a path [13]. The Cluj matrix, $\mathbb{CJ}_u$, differs from $\mathbb{SZ}_u$, by the supplementary condition: $\mathbf{(i,u)\cap(i,j)=max(\{i\})}$.

$\mathbb{SZ}_u$ matrix allows the construction of the symmetric matrices $\mathbb{SZ}_{e/p}$, (see Figure **1**) by relation

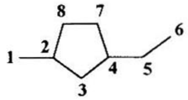
$$[\mathbb{SZ}_{e/p}]_{ij} = [\mathbb{SZ}_u]_{ij}\,[\mathbb{SZ}_u]_{ji} \tag{11}$$

and the derivation of two Wiener-type indices, as

$$I = (1/2)\sum_i \sum_j [\mathbb{SZ}_{e/p}]_{ij} = \sum_{(i,j)} [\mathbb{SZ}_u]_{ij}[\mathbb{SZ}_u]_{ji} \tag{12}$$

The index defined on edge is identical both in Szeged and Cluj matrices while the index defined on path is different. In other words, the above mentioned supplementary condition acts only when I is defined on path. Thus, $SZ_e = CJ_e$ in any graph whereas $SZ_p \neq CJ_p$ in most graphs (except, *e.g.* three- to five-membered simple cycles). In tree graphs, the index defined on edge is identical in Wiener, Szeged and Cluj matrices. Values of Wiener and Szeged indices in the set of octane isomers are listed in Table **1**.

Figure 1. Matrices $\mathbb{D}_e$, $\mathbb{D}_p$, SZ_e, SZ_p and SZ_u and derived indices in the graph 1.



1

$\mathbb{D}_e$

	1	2	3	4	5	6	7	8
1	0	1	2	3	4	5	3	2
2	1	0	1	2	3	4	2	1
3	2	1	0	1	2	3	2	2
4	3	2	1	0	1	2	1	2
5	4	3	2	1	0	1	2	3
6	2	4	3	2	1	0	3	4
7	3	2	2	1	2	3	0	1
8	2	1	2	2	3	4	1	0

$W = D_e = (1/2)\sum_i \sum_j [\mathbb{D}_e]_{ij} = 63$

SZ_e

	1	2	3	4	5	6	7	8
1	0	7	0	0	0	0	0	0
2	7	0	12	0	0	0	0	6
3	0	12	0	12	0	0	0	0
4	0	0	12	0	12	0	8	0
5	0	0	0	12	0	7	0	0
6	0	0	0	0	7	0	0	0
7	0	0	0	8	0	0	0	12
8	0	6	0	0	0	0	12	0

$SZ_e = (1/2)\sum_i \sum_j [SZ_e]_{ij} = 76$

$\mathbb{D}_p$

	1	2	3	4	5	6	7	8
1	0	1	3	6	10	15	6	3
2	1	0	1	3	6	10	3	1
3	3	1	0	1	3	6	3	3
4	6	3	1	0	1	3	1	3
5	10	6	3	1	0	1	3	6
6	15	10	6	3	1	0	6	10
7	6	3	3	1	3	6	0	1
8	3	1	3	3	6	10	1	0

$WW = D_p = (1/2)\sum_i \sum_j [\mathbb{D}_p]_{ij} = 119$

SZ_p

	1	2	3	4	5	6	7	8
1	0	7	5	10	12	12	10	5
2	7	0	12	8	12	10	12	6
3	5	12	0	12	8	12	6	8
4	10	8	12	0	12	6	8	12
5	12	12	8	12	0	7	8	12
6	12	10	12	6	7	0	12	10
7	10	12	6	8	8	12	0	12
8	5	6	8	12	12	10	12	0

$SZ_p = (1/2)\sum_i \sum_j [SZ_p]_{ij} = 266$

SZ_u

	1	2	3	4	5	6	7	8
1	0	1	1	2	3	4	2	1
2	7	0	3	2	4	5	3	3
3	5	4	0	3	4	6	3	4
4	5	4	4	0	6	6	4	4
5	4	3	2	2	0	7	2	3
6	3	2	2	1	1	0	2	2
7	5	4	2	2	4	6	0	4
8	5	2	2	3	4	5	3	0

$SZ_e = \sum_{i<j} [SZ_u]_{ij}[SZ_u]_{ji} = 76; \ (i,j) \in E(G)$

$SZ_\Delta = \sum_{i<j} [SZ_u]_{ij}[SZ_u]_{ji} = 190; \ (i,j) \notin E(G)$

$SZ_p = \sum_{i<j} [SZ_u]_{ij}[SZ_u]_{ji} = 266; \ (i,j) \in P(G)$

Table 1. Topological Data in Octanes.

Graph	W_e	W_p	SZ_e	SZ_Δ	SZ_p
C8	84	210	84	256	340
2MC7	79	185	79	241	320
3MC7	76	170	76	231	307
4MC7	75	165	75	219	294
3EC6	72	150	72	200	272
25M2C6	74	161	74	234	308
24M2C6	71	147	71	217	288
23M2C6	70	143	70	212	282
34M2C6	68	134	68	200	268
2E2MC5	67	129	67	175	242
22M2C6	71	149	71	209	280
33M2C6	67	131	67	183	250
234M3C5	65	122	65	193	258
3E3MC5	64	118	64	156	220
224M3C5	66	127	66	188	254
223M3C5	63	115	63	179	242
233M3C5	62	111	62	172	234
2233M4C4	58	97	58	174	232

Szeged numbers in line graphs, L_N and simple cycles, C_N:

The examination of Szeged indices in line graphs, L_N, and simple cycles, C_N, leads to the following relations

$$SZ_e(L_N) = N(N^2 -1) / 6 \quad (13)$$

$$SZ_p(L_N) = (N^3 + 3N^2 + 2N) / 6 \quad (14)$$

$$SZ_e(C_N) = [(1-z)N^3 + zN(N-1)^2] / 4 \quad (15)$$

$$SZ_p(C_N) = \binom{N}{2}[(N-1)^2 + 3(1-z)] / 32 \tag{16}$$

where $z = N \bmod 2$.

For comparison, the corresponding relations for Wiener, D_e, and hyper-Wiener, D_p, indices are given

$$D_e(L_N) = N(N^2 - 1) / 6 \tag{17}$$

$$D_p(L_N) = N(N - 1)(N + 1)(N + 2) / 24 \tag{18}$$

$$D_e(C_N) = [N^3 + z(z-2)N] / 8 \tag{19}$$

$$D_p(C_N) = \{N^4 + 3N^3 + (2-3z)N^2 + z[(2z-3)(3N + 1) + 1]\} / 48 \tag{20}$$

It is easily seen that only in trees and only when defined on edge, the two indices: Seged, SZ_e and Wiener, D_e, coincide. Otherwise, they are different quantities and coincide only accidentally or in special cases (*e.g.* in complete graphs, $SZ_e(K_N) = D_e(K_N) = N(N-1) / 2$; $N > 2$ [14]).

Separating and correlating tests

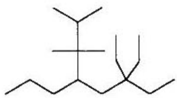
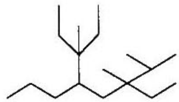
Three pairs of 4-trees (graphs **2-7**) constructed by Ivanciuc [20] so that they show pairwise identical distance degree sequence, **DDS**, were investigated. Results are presented in Figure **2**.

One can see that only the first pair shows distinct values for SZ_p whereas the Wiener indices and SZ_e are degenerate. The other pairs show degeneracies for all these indices.

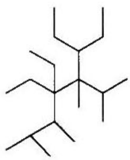
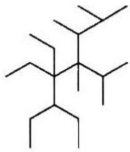
The spiro-graphs **8** and **11** in Figure **3** show identical values for D_e and SZ_e but different values for the corresponding hyper-indices.

Figure 2. (a) Pairs of graphs with pairwise identical distance degree sequence, DDS; (b) Topological indices: W_e, W_p, SZ_e, SZ_p.

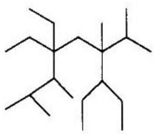
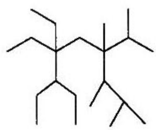
(a)

2 **3**

4 **5**

6 **7**

DDS:

(2;3): 34, 48, 58, 50, 52, 46, 18

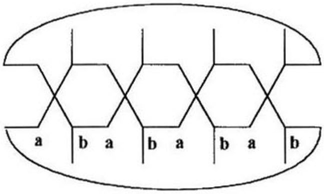
(4;5): 38, 56, 82, 92, 72, 32, 8

(6;7): 40, 58, 76, 78, 72, 56, 32, 8

(b) Topological indices:

graph:	2	3	4	5	6	7
W_e :	583	583	686	686	840	840
W_p :	1638	1638	1797	1797	2445	2445
SZ_e :	583	583	686	686	840	840
SZ_p :	7286	7264	9610	9610	13848	13848

Figure 3. Spiro-belts; DDS and Topological Indices.



8

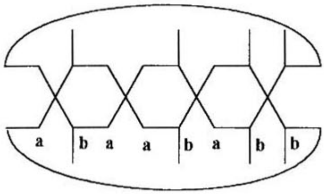
DDS: 64, 112, 152, 152, 144, 84, 32, 16

D_e = 1464

D_p = 4114

Sz_e = 4776

Sz_p = 52878



9

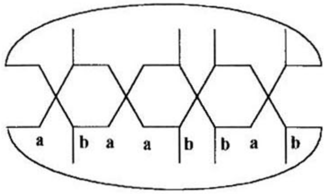
DDS: 64, 112, 152, 160, 128, 92, 48

D_e = 1456

D_p = 4054

Sz_e = 4744

Sz_p = 52682



10

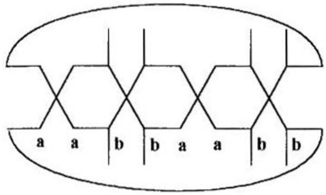
DDS: 64, 112, 152, 160, 128, 92, 40, 8

D_e = 1460

D_p = 4086

Sz_e = 4760

Sz_p = 52538



11

DDS: 64, 112, 152, 168, 112, 100, 32, 16

D_e = 1464

D_p = 4122

Sz_e = 4776

Sz_p = 51878

Table 2. All possible catafusenes with five 6-membered rings [21]; Szeged and Wiener indices.

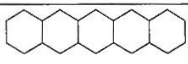
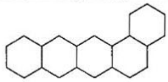
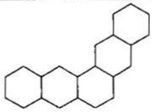
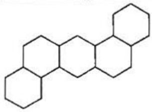
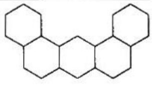
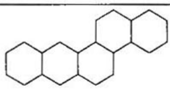
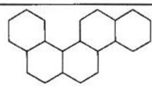
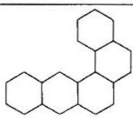
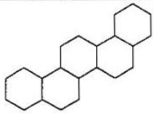
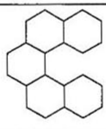
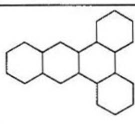
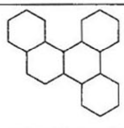
No.	Graph	SZ_e	SZ_Δ	SZ_p	D_e	D_Δ	D_p
1		2506	18119	20625	1011	2435	3396
2		2410	17830	20240	987	2279	3266
3		2378	17601	19979	979	2227	3206
4		2354	17639	19993	971	2191	3162
5		2290	17437	19727	955	2063	3018
6		2354	17554	19908	971	2191	3162
7		2234	17044	19278	931	1931	2862
8		2258	17055	19313	939	1967	2906

Table 2. (continued)

9		2330	17459	19789	963	2155	3118
10		2138	16481	18619	899	1723	2622
11		2290	16247	18537	907	1775	2682
12		2125	16358	18483	883	1643	2526

It is obvious that, despite the limits above illustrated, the hyper-Szeged number, SZ_p , is more discriminating than the SZ_e number. This is also the conclusion of a test performed on the set of all possible catafusenes with five 6-membered rings [21], listed in Table **2**. No degeneracy has occurred for SZ_p , whereas for SZ_e two pairs of degenerate values were encountered (entries **4; 6** and **5, 11**). The pair **4; 6** shows also degeneracy for W_e , W_p and the **Harary** number (see [21]); it is a consequence of the identity of **DDS** **(52,76,82,68, 56,44,34,24,16,8,2)**.

The correlating ability of Szeged indices was tested on a dozen of simple cycles (C_6 - C_{17} - see Table **3**) *vs.* the enthalpy of formation [22]. Statistics of single- and two variable regressions are given in Table **4**.

Table 3. Enthalpy of Formation and Topological Descriptors in Simple Cycles.

No.	Graph	ΔHF* (KJ/mol)	SZ_e	SZ_Δ	SZ_p	D_e	D_p
1	C_6	-123.40	54	51	105	27	42
2	C_7	-118.03	63	126	189	42	70
3	C_8	-124.40	128	236	364	64	120
4	C_9	-132.80	144	432	576	90	180
5	C_{10}	-154.30	250	695	945	125	275
6	C_{11}	-179.40	275	1100	1375	165	385
7	C_{12}	-230.20	432	1614	2046	216	546
8	C_{13}	-246.40	468	2340	2808	273	728
9	C_{14}	-289.00	686	3227	3913	343	980
10	C_{15}	-301.40	735	4410	5145	420	1260
11	C_{16}	-321.70	1024	5816	6840	512	1632
12	C_{17}	-364.30	1088	7616	8704	612	2040

* from ref. [22]

Table 4. Statistics of Single and Two Variable Regression of Parameters from Table 3.

$$\mathbf{y = a + \textstyle\sum_i b_i\, x_i}$$

Y	X	R	S	F	a	b
ΔHF	SZ_e	0.9813	17.8143	259.5516	-109.1849	-0.2345
	SZ_Δ	0.9573	26.7403	109.6318	-137.1480	-0.0340
	SZ_p	0.9623	25.1641	125.0873	-133.3277	-0.0299
	D_e	0.9834	16.8032	292.9670	-107.8221	-0.4470
	D_p	0.9694	22.7084	155.8838		
	NC	0.9764	19.9675	204.5503	59.1899	-23.8812

Table 4 (continued).

ΔHF	SZ_e NC	0.9875	15.3364	171.3451	-33.6631	-0.1382 -10.4530
	SZ_Δ NC	0.9859	16.3116	156.2516	-6.7371	-0.0130 -15.5345
	SZ_p NC	0.9862	16.1169	160.1597	-10.4474	-0.01216 -14.9176
	SZ_e D_e	0.9912	12.8899	252.9226	-83.4038	0.0325 -0.8599
	SZ_Δ D_e	0.9836	17.4057	36.6774	-107.9855	-0.0761 -0.3056
	SZ_p D_e	0.9910	13.0534	246.5145	-83.2722	0.0326 -0.9215

The correlations involving the Szeged indices are at least as good as those shown by the Wiener indices. Other correlations of the Szeged indices with physico-chemical properties of alkanes and cycloalkanes are given in ref. [2].

Conclusions

A new unsymmetric matrix, $\mathbb{SZ}_u$ (unsymmetric Szeged matrix) was proposed and exemplified both for acyclic and cycle-containing structures. Its relation with the $\mathbb{CJ}_u$ matrix was discussed. The derived Szeged numbers were compared to the Wiener matrix- and Cluj matrix-derived numbers and tested for separating and correlating ability on selected sets of graphs. The Szeged numbers represent an alternative to the Wiener numbers and provide new tools in chemical graph theory.

References

[1] M.V. Diudea, Wiener and Hyper-Wiener Numbers in a Single Matrix, *J. Chem. Inf. Comput. Sci* (in press).

[2] M.V. Diudea, Cluj Matrix Invariants, *J. Chem. Inf. Comput. Sci* (submitted).

[3] H. Wiener, Structural Determination of Paraffin Boiling points. *J.Am.Chem.Soc.* **1947,** *69*, 17-20.

[4] I. Gutman; Y.N. Yeh; S.L. Lee; Y.L. Luo, Some Recent Results in the Theory of the Wiener Number, *Indian J. Chem.* **1993**, *32A*, 651-661.

[5] S. Nikolić; N. Trinajstić; Z. Mihalić, The Wiener Index: Development and Applications. *Croat. Chem. Acta* **1995,** *68,* 105-129.

[6] M. Randić, Novel Molecular Descriptor for Structure-Property Studies. *Chem. Phys. Lett.,* **1993**, *211*, 478-483.

[7] M. Randić; X. Guo; T. Oxley; H. Krishnapriyan, Wiener Matrix: Source of Novel Graph Invariants. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 700-716.

[8] M. Randić; X. Guo; T. Oxley; H. Krishnapriyan; L. Naylor, Wiener Matrix Invariants. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 361-367.

[9] I. Lukovits; W. Linert, A Novel Definition of the Hyper-Wiener Index for Cycles. *J. Chem. Inf. Comput. Sci.* **1994**, *34,* 899-902.

[10] I. Lukovits; I. Gutman, Edge-Decomposition of the Wiener Number, *MATCH* **31** (1994) 133-144.

[11] I. Lukovits, An Algorithm for Computation of Bond Contributions of the Wiener Index. *Croat. Chem. Acta* **1995**, *68*, 99-103.

[12] H. Hosoya, Topological Index. A Newly Proposed Quantity Characterizing the Topological Nature of Structural Isomers of Saturated Hydrocarbons. *Bull. Chem. Soc. Jpn.* **1971,** *44,* 2332-2339.

[13] M.V. Diudea, Walk Numbers ${}^{e}W_{M}$: Wiener-Type Numbers of Higher Rank, *J. Chem. Inf. Comput. Sci.* **1996** (in press).

[14] I. Gutman, A Formula for the Wiener Number of Trees and Its Extension to Graphs Containing Cycles, *Graph Theory Notes New York* **1994**, *27*, 9-15.

[15] A.A. Dobrynin; I. Gutman, On a Graph Invariant Related to the Sum of all Distances in a Graph, *Publ. Inst. Math. (Beograd)* **1994**, *56*, 18-22.

[16] A.A. Dobrynin; I. Gutman; G. Dömötör, A Wiener-Type Graph Invariant for Some Bipartite Graphs, *Appl. Math. Lett.* **1995**, *8*, 57-62.

[17] A.A. Dobrynin; I. Gutman, Solving a Problem Connected with Distances in Graphs, *Graph Theory Notes New York* **1995**, *28*, 21-23.

[18] P.V. Khadikar; N.V. Deshpande; P.P. Kale; A.A. Dobrynin; I. Gutman; G. Dömötör, The Szeged Index and an Analogy with the Wiener Index, *J.Chem.Inf. Comput. Sci.* **1995**, *35*, 547-550.

[19] I. Gutman; S. Klavžar, An Algorithm for the Calculation of the Szeged Index of Benzenoid Hydrocarbons, *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 1011-1014.

[20] O. Ivanciuc; T.S. Balaban; A.T. Balaban, Chemical Graphs with Degenerate Topological Indices Based on Information on Distances, *J. Math. Chem.* **1993**, *14*, 21-33.

[21] D. Bonchev; A.T. Balaban; X. Liu; D.J. Klein, Molecular Cyclicity and Centricity of Polycyclic Graphs.I. Cyclicity Based on Resistance Distances or Reciprocal Distances, *Int. J. Quantum Chem.* **1994**, *50*, 1-20.

[22] C. Csunderlik; L. Cotarca; H.H. Glatt, Structure and Properties of Organic Compounds, Ed. Tehnica, Bucharest, 1987, vol. 2, p. 62 (in Romanian).