Real Number Vertex Invariants: Regressive Distance Sums and Related Topological Indices

Alexandru T. Balaban*

Department of Organic Chemistry, Polytechnic University, Splaiul Independentei 313, 77206 Bucharest, Romania

Mircea V. Diudea

Department of Organic Chemistry, "Babes-Bolyai" University, Strada Arany Janos 11, 3400 Cluj, Romania

Received September 18, 1992

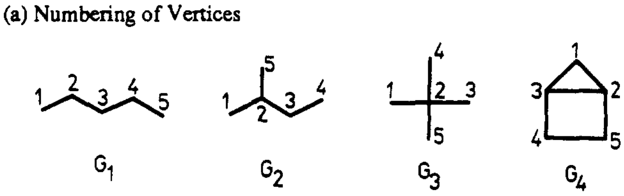
A new type of layer matrix, called **R** matrix, was constructed on the basis of distance sums of vertices. This matrix was operated with two classes of operators: one of "centricity" ("c") type and the other of "centrocomplexity" ("x") type, the last one taking into account the "more important" vertices in molecular graphs. The matrix invariants are computed with a TURBO PASCAL, TOPIND 10 program for various examples, and finally an intercorrelating matrix is given for the proposed topological indices in the set of heptane isomers.

INTRODUCTION

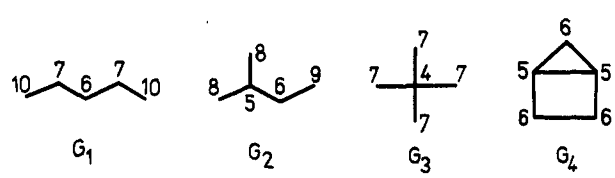
Various papers[1–14] advocated the use of real number vertex invariants for obtaining less degenerate topological indices (TIs). First-generation TIs had been integer numbers obtained on the basis of integer-number local vertex invariants (LOVIs) (e.g. the Wiener index[15]). Second-generation TIs were real numbers obtained from integer LOVIs using more sophisticated operations; such TIs are Randić's molecular connectivity,[16] Kier and Hall's extended molecular connectivity,[17] all information-theoretic indices,[18–20] and the average distance sum connectivity, J.[21]

The newest (third-generation) TIs are real numbers based on real-number LOVIs. Among these ones, in a recent paper,[22] we introduced new LOVIs called regressive vertex degrees. We started from the idea that TIs based on the distance matrix of graphs reflect the more distant relationships between graph

(a) Numbering of Vertices

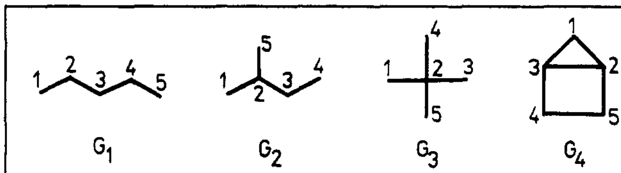


(b) Distance Sums



	G1 for k = 0–4					G2 for k = 0–3				G3 for k = 0–2			G4		
i	0	1	2	3	4	0	1	2	3	0	1	2	0	1	2
1	10	7	6	7	10	8	5	14	9	7	4	21	6	10	12
2	7	16	7	10	0	5	22	9	0	4	28	0	5	17	6
3	6	14	20	0	0	6	14	16	0	7	4	21	5	17	6
4	7	16	7	10	0	9	6	5	16	7	4	21	6	11	11
5	10	7	6	7	10	8	5	14	9	7	4	21	6	11	11

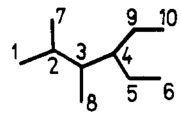
Figure 1. **R** matrix in graphs G$_1$–G$_4$.



		G1	G2	G3	G4
		J = 2.1906	J = 2.5395	J = 3.0237	J = 2.1939
r_i	1	10.070 6071	8.051 409	7.0421	6.1012
	2	7.160 7100	5.220 900	4.2800	5.1706
	3	6.142 0000	6.141 600	7.0421	5.1706
	4	7.160 7200	9.060 516	7.0421	6.1111
	5	10.070 6071	8.051 409	7.0421	6.1111
r_i*	1	0.099 30	0.124 20	0.142 00	0.163 90
	2	0.139 65	0.191 54	0.233 64	0.193 40
	3	0.162 81	0.162 82	0.142 00	0.193 40
	4	0.139 65	0.110 37	0.142 00	0.163 64
	5	0.099 30	0.124 20	0.142 00	0.163 64
	R*	0.640 71	0.713 13	0.801 66	0.877 98
rc_i	1	0.143 88	0.208 20	0.334 77	0.344 51
	2	0.208 74	0.343 17	0.716 61	0.362 52
	3	0.320 25	0.328 61	0.334 77	0.362 52
	4	0.208 74	0.205 20	0.334 77	0.346 46
	5	0.143 88	0.208 20	0.334 77	0.346 46
	RC	1.025 49	1.293 38	2.055 70	1.762 45
rx_i	1	0.099 30	0.124 20	0.142 00	0.327 80
	2	0.279 30	0.574 61	0.934 58	0.580 20
	3	0.325 63	0.325 65	0.142 00	0.580 20
	4	0.279 30	0.110 37	0.142 00	0.327 27
	5	0.099 30	0.124 20	0.142 00	0.327 27
	RX	1.082 83	1.259 03	1.502 59	2.142 76
rj_i	1	0.117 76	0.154 24	0.182 15	0.356 08
	2	0.268 55	0.485 07	0.728 60	0.549 34
	3	0.301 58	0.310 65	0.182 15	0.549 34
	4	0.268 55	0.134 05	0.182 15	0.341 53
	5	0.117 76	0.154 24	0.182 15	0.341 53
	RJ	1.074 19	1.238 26	1.457 19	2.137 83
dj_i	1	0.119 52	0.158 11	0.188 98	0.365 15
	2	0.273 83	0.498 80	0.755 93	0.565 15
	3	0.308 61	0.318 66	0.188 98	0.565 15
	4	0.273 83	0.136 08	0.188 98	0.349 24
	5	0.119 52	0.158 11	0.188 98	0.349 24
	DJ	1.095 30	1.269 77	1.511 86	2.193 93

Figure 2. Local and global invariants based on **R** matrix in G$_1$–G$_4$.

0095-2338/93/1633-0421$04.00/0 © 1993 American Chemical Society



G_5

(a) Actual Parameters

	r_i	r_i^*	rc_i	rx_i	rj_i	dj_i
no. 3	17.063 104 062	0.058 605 984	0.133 44	0.175 817 95	0.159 74	0.160 26
no. 4	17.063 108 058	0.058 605 970	0.134 32	0.175 817 91	0.159 46	0.159 97

(b) Other Parameters, Including the Eccentricity, ecc_i

	ecc_i	f_{ik}[22]	b_{ik}[22]	self returning walks,[23]	random walks[23]
no. 3	3	3, 4, 2	3, 7, 6, 2	0, 3, 0, 13, 0, 59, 0	3, 7, 15, 33, 73, 159
no. 4	3	3, 4, 2	3, 7, 6, 2	0, 3, 0, 13, 0, 61, 0	3, 7, 15, 35, 73, 169

Figure 3. Vertex discrimination in 234MMEC$_6$, G$_5$.

vertices. In this respect, the distance sums for each vertex are dominated by the more remote vertices; on the other hand, most TIs based on the adjacency matrix emphasize only the immediate vicinity relationships. By means of the regressive vertex degrees, we extended the neighborhood relationship to also include more distant vertices but in attenuated form, their contribution decreasing with increasing distance.[22]

In the present paper we perform a similar operation for distance sums, in order to convert them into real numbers. The new LOVIs herein proposed may include information about multiple bonding and about heteroatoms, as will be shown below.

REGRESSIVE DISTANCE SUMS

One starts by calculating the distance sums D_i for all vertices in graph G, i.e. by summing entries over rows or columns in the distance matrix ($D_i = \sum_j d_{ij}$). Next, one writes a new matrix which will be called the **R** matrix (for regressive distance sums) according to the various shells around each vertex i: the entry in column $k = 0$ is just the distance sum, D_i. The next columns will sum all distance sums, D_j, of vertices j, belonging to a shell at distance $d_{ij} = k$, around the vertex i. Thus, the entries in **R** matrix will be

$$r_{ik} = \sum_{j,d_{ij}=k} D_j \tag{1}$$

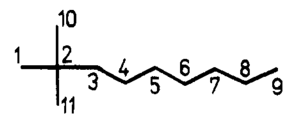
The number of columns in **R** is equal to the largest distance in G, (i.e. the graph diameter). It is obvious that the sums over each row in **R** are all equal to twice the Wiener index (the sum of all distances in G). We exemplify on four graphs with five vertices, as in Figure 1.

By analogy with the regressive degrees[22] (which count the decreasing contributions of more remote vertices to the classical degree of a vertex, as their distance to that vertex increases), we propose new real-number LOVIs, *regressive distance sums*, defined as

$$r_i = \sum_{k=0}^{\text{diam}} 10^{-nk} r_{ik} \tag{2}$$

where diam is the diameter and n denotes the number of digits for the maximal r_{ik} value in G.

This vertex invariant is directly related to the second criterion (D_i = min), established by Bonchev et al.[23] for the center



G_6

rc_i, bc_i	5, 4, 6, 3, 7, 2, 8, 1 = 10 = 11, 9
r_i^*	4, 5, 3, 6, 2, 7, 1 = 10 = 11, 8, 9
rx_i, rj_i	2, 4, 5, 3, 6, 7, 8, 1 = 10 = 11, 9
dj_i	2, 4, 5 = 3, 6, 7, 8, 1 = 10 = 11, 9
bx_i	2, 3, 4, 5, 6, 7, 8, 1 = 10 = 11, 9

Figure 4. Vertex ordering in 22MMC$_9$, G$_6$.

identification in G, so that it can be used for centric ordering of vertices (see refs 24 and 25). Since the r_i parameters become cumbersome in large graphs, for an easier handling of the **R** matrix we propose four operators, defining four other LOVIs, whose first letter is r, as follows:

$$r_i^* = [\sum_{k=0}^{\text{diam}} 10^{-nk} r_{ik}]^{-1} = (r_i)^{-1} \tag{3}$$

$$rc_i = [\sum_{k=1}^{\text{diam}} (r_{ik})^{k/d_{\text{spec}}}]^{-1} \tag{4}$$

$$rx_i = [r_i/dg_i - m_i]^{-1} w_i \tag{5}$$

$$m_i = f_i[r_{i0}/10 + r_{i1}/100] \tag{6}$$

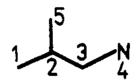
$$f_i = \sum_j (c_{ij} - 1) \tag{7}$$

$$rj_i = \sum_{(i,j)} [r_i/(w_i c_i) r_j/(w_j c_j)]^{-1/2} \tag{8}$$

$$c_i = 1 + f_i \tag{9}$$

where d_{spec} is a specified distance value, usually larger than the largest path in G[24,25] (in the following, $d_{\text{spec}} = 10$ unless otherwise specified); m_i is the local parameter for multiple bonds; c_i, f_i refer to the connectivity around the vertex i; c_{ij} is the conventional bond order, 1, 2, 3, and 1.5 for single, double, triple, and aromatic bonds, respectively; and w_i is a weighting factor, accounting for heteroatoms, as defined in refs 22 and 26.

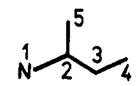
Summation of the new LOVIs over all i vertices in G provides the corresponding global indices (TIs), denoted by capital



G_7

$J_{het} = 2.5030$

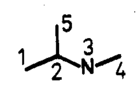
	rx_i		rj_i		dj_i		bx_i
2	0.734 45	2	0.678 94	2	0.698 05	2	4.358 51
3	0.450 28	3	0.438 32	3	0.449 35	3	3.346 21
1	0.201 47	1	0.222 09	1	0.227 67	1	2.159 01
5	0.201 47	5	0.222 09	5	0.227 67	5	2.159 01
4	0.184 03	4	0.203 55	4	0.206 63	4	2.054 27
RX	1.771 70	*RJ*	1.764 99	*DJ*	1.809 36	*BX*	14.077 00



G_8

$J_{het} = 2.4971$

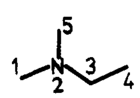
	rx_i		rj_i		dj_i		bx_i
2	0.734 45	2	0.682 03	2	0.701 21	2	4.358 51
3	0.450 28	3	0.435 54	3	0.446 52	3	3.346 21
1	0.207 10	1	0.225 17	1	0.230 83	1	2.219 34
5	0.201 47	5	0.222 09	5	0.227 67	5	2.159 01
4	0.179 03	4	0.200 77	4	0.203 80	4	1.998 42
RX	1.772 32	*RJ*	1.765 58	*DJ*	1.810 03	*BX*	14.081 49



G_9

$J_{het} = 2.4540$

	rx_i		rj_i		dj_i		bx_i
2	0.734 45	2	0.682 30	2	0.701 52	2	4.358 51
3	0.463 23	3	0.441 76	3	0.452 89	3	3.442 44
1	0.201 47	1	0.222 09	1	0.227 67	1	2.159 01
5	0.201 47	5	0.222 09	5	0.227 67	5	2.159 01
4	0.179 03	4	0.203 63	4	0.206 71	4	1.998 42
RX	1.779 64	*RJ*	1.771 86	*DJ*	1.816 46	*BX*	14.117 39



G_{10}

$J_{het} = 2.4057$

	rx_i		rj_i		dj_i		bx_i
2	0.768 17	2	0.694 36	2	0.713 90	2	4.558 67
3	0.450 28	3	0.440 87	3	0.452 03	3	3.346 21
1	0.201 47	1	0.227 13	1	0.232 84	1	2.159 01
5	0.201 47	5	0.227 13	5	0.232 84	5	2.159 01
4	0.179 03	4	0.200 77	4	0.203 80	4	1.998 42
RX	1.800 42	*RJ*	1.790 25	*DJ*	1.835 40	*BX*	14.221 32

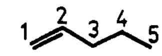
Figure 5. Heteroatom perception.

letters, corresponding to the respective LOVIs.

If r_i is replaced by D_i in eq 8, a new vertex invariant can be designed:

$$dj_i = \sum_{(i,j)} [D_i/(w_i c_i) D_j/(w_j c_j)]^{-1/2} \qquad (10)$$

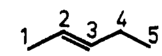
It is easily seen that, when $w_i = 1$ and $c_i = 1$, the corresponding global index, DJ, is related to the J index[21]



G_{11}

$J = 2.4017$

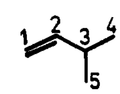
	rx_i		rj_i		dj_i		bx_i
2	0.367 60	2	0.448 76	2	0.457 26	2	2.551 00
3	0.325 63	3	0.364 03	3	0.372 52	3	2.420 00
4	0.279 30	4	0.268 55	4	0.273 83	4	2.321 00
1	0.111 10	1	0.235 52	1	0.239 05	1	1.342 10
5	0.099 30	5	0.117 76	5	0.119 52	5	1.222 10
RX	1.182 93	*RJ*	1.434 62	*DJ*	1.462 18	*BX*	9.856 20



G_{12}

$J = 2.6224$

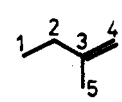
	rx_i		rj_i		dj_i		bx_i
3	0.429 00	3	0.514 82	3	0.526 82	3	2.660 00
2	0.367 60	2	0.468 11	2	0.477 64	2	2.551 00
4	0.279 30	4	0.331 01	4	0.337 74	4	2.321 00
1	0.099 30	1	0.166 54	1	0.169 03	1	1.222 10
5	0.099 30	5	0.117 76	5	0.119 52	5	1.222 10
RX	1.274 50	*RJ*	1.598 24	*DJ*	1.630 76	*BX*	9.976 20



G_{13}

$J = 2.8257$

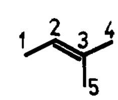
	rx_i		rj_i		dj_i		bx_i
3	0.574 61	3	0.558 22	3	0.574 43	3	3.410 00
2	0.429 04	2	0.517 86	2	0.530 36	2	2.660 00
4	0.124 20	1	0.268 11	1	0.272 17	1	1.352 00
5	0.124 20	4	0.154 24	5	0.158 11	4	1.331 00
1	0.123 45	5	0.154 24	4	0.158 11	5	1.331 00
RX	1.375 50	*RJ*	1.652 67	*DJ*	1.693 18	*BX*	10.084 00



G_{14}

$J = 2.8474$

	rx_i		rj_i		dj_i		bx_i
3	0.980 10	3	0.776 35	3	0.798 03	3	3.750 00
2	0.325 65	2	0.383 80	2	0.394 28	2	2.420 00
4	0.138 86	4	0.308 48	4	0.316 23	4	1.461 00
5	0.124 20	5	0.218 13	5	0.223 61	5	1.331 00
1	0.110 37	1	0.134 05	1	0.136 08	1	1.232 00
RX	1.679 18	*RJ*	1.820 81	*DJ*	1.868 23	*BX*	10.194 00



G_{15}

$J = 3.1943$

	rx_i		rj_i		dj_i		bx_i
3	0.980 10	3	0.789 45	3	0.812 36	3	3.750 00
2	0.429 04	2	0.542 78	2	0.557 60	2	2.660 00
4	0.124 20	4	0.218 13	4	0.223 61	4	1.331 00
5	0.124 20	5	0.218 13	5	0.223 61	5	1.331 00
1	0.110 37	1	0.189 58	1	0.192 45	1	1.232 00
RX	1.767 91	*RJ*	1.958 06	*DJ*	2.009 62	*BX*	10.304 00

Figure 6. Double bonding perception.

Table I. **R** Matrices and *RC* Index, for G_{16} and G_{17}, According to Their Centric Numbering (d_{spec} = 20)[a]

G_{16}: 42 Vertices and 51 Edges

1	161	507	1110	1640	1840	2112	1766	618						
2	161	499	1102	1640	1880	2112	1734	626						
3	169	539	945	1431	2081	2040	1369	867	313					
4	169	539	945	1431	2081	2040	1369	867	313					
5	173	547	957	1419	1812	1731	1626	1180	309					
6	173	547	957	1419	1812	1731	1626	1180	309					
7	181	559	977	1519	1912	1430	940	1056	867	313				
8	181	559	977	1519	1912	1430	940	1056	867	313				
9	181	563	1230	1559	1366	1390	1217	1056	883	309				
10	181	563	1230	1559	1366	1390	1217	1056	883	309				
11	197	394	840	1290	1242	1511	1731	1369	867	313				
12	197	394	840	1290	1242	1511	1731	1369	867	313				
13	205	406	583	1001	1491	1535	1418	1626	1180	309				
14	205	406	583	1001	1491	1535	1418	1626	1180	309				
15	209	418	848	1310	1318	1342	1133	940	1056	867	313			
16	209	418	848	1310	1318	1342	1133	940	1056	867	313			
17	209	667	1145	1057	768	1053	1390	1217	1056	883	309			
18	209	667	1145	1057	768	1053	1390	1217	1056	883	309			
19	225	695	948	651	511	933	1511	1731	1369	867	313			
20	225	695	948	651	511	933	1511	1731	1369	867	313			
21	233	446	659	900	796	933	1254	1418	1626	1180	309			
22	233	446	659	900	796	933	1254	1418	1626	1180	309			
23	237	707	948	647	543	1021	1342	1133	940	1056	867	313		
24	237	707	948	647	543	1021	1342	1133	940	1056	867	313		
25	241	719	944	635	539	768	1053	1390	1217	1056	883	309		
26	241	719	944	635	539	768	1053	1390	1217	1056	883	309		
27	245	490	735	864	744	768	1053	1390	1217	1056				
28	245	490	735	864	744	768	1053	1390	1217	1056	883	309		
29	261	522	743	651	342	511	933	1511	1731	1369	867	313		
30	261	522	743	651	342	511	933	1511	1731	1369	867	313		
31	273	546	731	639	350	543	1021	1342	1133	940	1056	867	313	
32	273	546	731	639	350	543	1021	1342	1133	940	1056	867	313	
33	277	554	723	631	354	539	768	1053	1390	1217	1056	883	309	
34	277	554	723	631	354	539	768	1053	1390	1217	1056	883	309	
35	281	558	486	422	587	744	768	1053	1390	1217	1056	883	309	
36	281	558	486	422	587	744	768	1053	1390	1217	1056	883	309	
37	297	570	498	434	378	342	511	933	1511	1731	1369	867	313	
38	297	570	498	434	378	342	511	933	1511	1731	1369	867	313	
39	309	570	498	434	378	350	543	1021	1342	1133	940	1056	867	313
40	309	570	498	434	378	350	543	1021	1342	1133	940	1056	867	313
41	313	558	486	442	386	354	539	768	1053	1390	1217	1056	883	309
42	313	558	486	442	386	354	539	768	1053	1390	1217	1056	883	309

	RC		*RC*		*RC*		*RC*		*RC*		*RC*		*RC*
1	0.027 3258	7	0.015 4987	13	0.014 5172	19	0.010 6084	25	0.007 7722	31	0.005 6155	37	0.005 3242
2	0.027 3198	8	0.015 4987	14	0.014 5172	20	0.010 6084	26	0.007 7722	32	0.005 6155	38	0.005 3242
3	0.021 2683	9	0.015 4075	15	0.011 0351	21	0.010 2561	27	0.007 7552	33	0.005 5346	39	0.004 0122
4	0.021 2683	10	0.015 4075	16	0.011 0351	22	0.010 2561	28	0.007 7552	34	0.005 5346	40	0.004 0122
5	0.020 7258	11	0.014 9788	17	0.010 9346	23	0.007 8730	29	0.007 5071	35	0.005 5187	41	0.003 9474
6	0.020 7258	12	0.014 9788	18	0.010 9346	24	0.007 8730	30	0.007 5071	36	0.005 5187	42	0.003 9474

global index, *RC*: 0.466 826 603

G_{17}: 42 Vertices and 51 Edges

1	161	503	1106	1640	1860	2112	1750	622				
2	161	503	1106	1640	1860	2112	1750	622				
3	169	539	949	1435	2081	2020	1369	883	309			
4	169	539	949	1435	2081	2020	1369	883	309			
5	173	547	953	1415	1812	1751	1626	1164	313			
6	173	547	953	1415	1812	1751	1626	1164	313			
7	181	559	977	1523	1916	1430	920	1056	883	309		
8	181	559	977	1523	1916	1430	920	1056	883	309		
9	181	563	1230	1555	1362	1390	1237	1056	867	313		
10	181	563	1230	1555	1362	1390	1237	1056	867	313		
11	197	394	840	1294	1246	1511	1711	1369	883	309		
12	197	394	840	1249	4	1246	1511	1711	1369	883	309	
13	205	406	583	997	1487	1535	1438	1626	1164	313		
14	205	406	583	997	1487	1535	1438	1626	1164	313		
15	209	418	848	1310	1322	1346	1133	920	1056	883	309	
16	209	418	848	1310	1322	1346	1133	920	1056	883	309	
17	209	667	1145	1057	764	1049	1390	1237	1056	867	313	
18	209	667	1145	1057	764	1049	1390	1237	1056	867	313	
19	225	695	948	651	515	937	1511	1711	1369	883		309
20	225	695	948	651	515	937	1511	1711	1369	883	309	
21	233	446	659	900	792	929	1254	1438	1626	1164	313	
22	233	446	659	900	792	929	1254	1438	1626	1164	313	
23	237	707	948	647	543	1025	1346	1133	920	1056	883	309

Table I (Continued)

G_{17}: 42 Vertices and 51 Edges

24	237	707	948	647	543	1025	1346	1133	920	1056	883	309		
25	241	719	944	635	539	764	1049	1390	1237	1056	867	313		
26	241	719	944	635	539	764	1049	1390	1237	1056	867	313		
27	245	490	73q5	864	744	764	1049	1390	1237	1056	867	313		
28	245	490	735	864	744	764	1049	1390	1237	1056	867	313		
29	261	522	743	651	342	515	937	1511	1711	1369	883	309		
30	261	522	743	651	342	515	937	1511	1711	1369	883	309		
31	273	546	731	639	350	543	1025	1346	1133	920	1056	883	309	
32	273	546	731	639	350	543	1025	1346	1133	920	1056	883	309	
33	277	554	723	631	354	539	764	1049	1390	1237	1056	867	313	
34	277	554	723	631	354	539	764	1049	1390	1237	1056	867	313	
35	281	558	486	422	587	744	764	1049	1390	1237	1056	867	313	
36	281	558	486	422	587	744	764	1049	1390	1237	1056	867	313	
37	297	570	498	434	378	342	515	937	1511	1711	1369	883	309	
38	297	570	498	434	378	342	515	937	1511	1711	1369	883	309	
39	309	570	498	434	378	350	543	1025	1346	1133	920	1056	883	309
40	309	570	498	434	378	350	543	1025	1346	1133	920	1056	883	309
41	313	558	486	442	386	354	539	764	1049	1390	1237	1056	867	313
42	313	558	486	442	386	354	539	764	1049	1390	1237	1056	867	313

	RC		*RC*		*RC*		*RC*		*RC*		*RC*		*RC*
1	0.027 3227	7	0.015 5020	13	0.014 5135	19	0.010 6066	25	0.007 7727	31	0.005 6159	37	0.005 3226
2	0.027 3227	8	0.015 5020	14	0.014 5135	20	0.010 6066	26	0.007 7727	32	0.005 6159	38	0.005 3226
3	0.021 2669	9	0.015 4065	15	0.011 0369	21	0.010 2542	27	0.007 7556	33	0.005 5353	39	0.004 0122
4	0.021 2669	10	0.015 4065	16	0.011 0369	22	0.010 2542	28	0.007 7556	34	0.005 5353	40	0.004 0122
5	0.020 7212	11	0.014 9777	17	0.010 9347	23	0.007 8739	29	0.007 5053	35	0.005 5194	41	0.003 9481
6	0.020 7212	12	0.014 9777	18	0.010 9347	24	0.007 8739	30	0.007 5053	36	0.005 5194	42	0.003 9481

global index, *RC*: 0.466 807 649

[a] Missing entries in shorter columns are zeros.

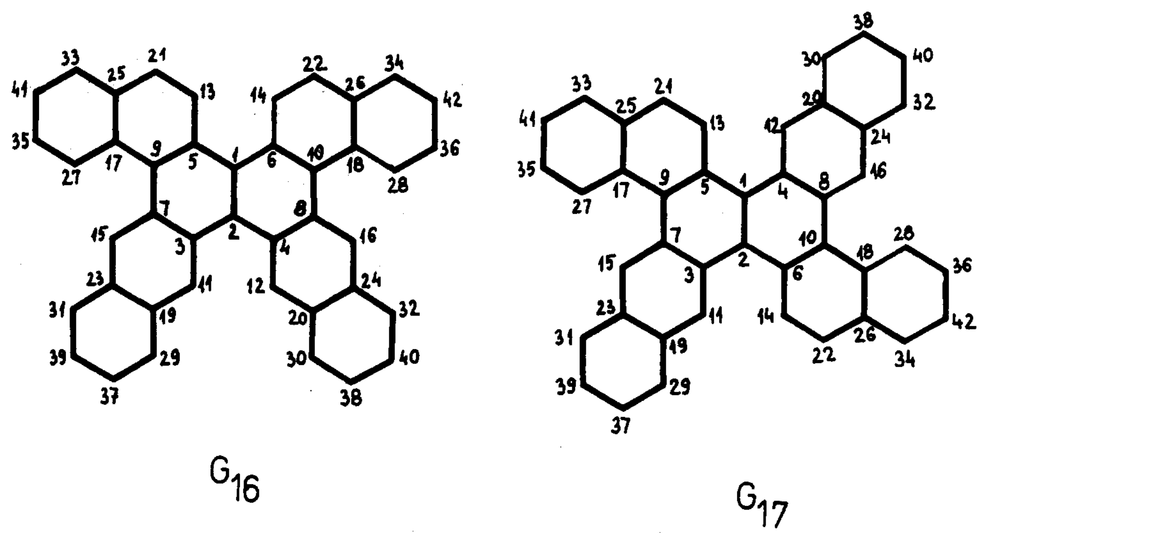


Figure 7. Centric numbering of graphs G_{16} and G_{17}.[27]

as follows:

$$DJ = 2J(\mu + 1)/q \qquad (11)$$

where μ is the cyclomatic number and q is the number of edges in G.

The above LOVIs and TIs are examplified for G_1–G_4 in Figure 2.

INTRAMOLECULAR ORDERING

That the **R** matrix is more powerful in discriminating nonequivalent vertices than $\mathbf{F}^{24}$ and $\mathbf{B}^{22}$ matrices and their derived invariants or other topological descriptors can be seen in Figure 3 for G_5.[23]

(i) "c"-Type versus "x"-Type Ordering. Our invariants are capable of ordering the vertices in molecular graphs either in terms of centricity ("*c*") or centrocomplexity ("*x*").[25] Figure 4 shows the ordering given by operators rc_i, r_i^*, rx_i, rj_i, and dj_i and the older operators bc_i and bx_i (the last one denoted as $BY^{(2)}$ in ref 22 and as BCX in ref 25).

(ii) Heteroatoms and Multiple Bonding. The operators rx_i, rj_i, dj_i, and bx_i are sensitive to the presence of heteroatoms and multiple bondings by means of the w_i and m_i (or c_i) factors. We exemplify this with a set of amines and alkenes, the LOVIs values being given in Figures 5 and 6, in decreasing order.

One can see that our invariants (the earlier *BX* included) emphasize the centricity of heteroatoms and multiple bonding, their values paralleling the centricity of "important" vertices in graphs.

INTERMOLECULAR ORDERING OF ISOMERIC GRAPHS

The **R** matrix surpasses the ability of the known layer matrices **F** and **B** to discriminate between isomeric graphs.

(a) New TIs

G	isomer	R^*	RC	RX	RJ	DJ
18	C_7	0.4490	0.6412	0.8035	0.8000	0.8158
19	$2MC_6$	0.4825	0.7802	0.8789	0.8730	0.8928
20	$3MC_6$	0.5040	0.8303	0.9306	0.9223	0.9439
21	$24MMC_5$	0.5212	1.0452	0.9692	0.9599	0.9844
22	$3EC_5$	0.5267	1.0782	0.9849	0.9740	0.9974
23	$23MMC_5$	0.5464	1.1103	1.0339	1.0208	1.0481
24	$22MMC_5$	0.5461	1.0770	1.0343	1.0211	1.0515
25	$33MMC_5$	0.5728	1.1931	1.1043	1.0865	1.1201
26	$223MMC_4$	0.5959	1.4062	1.1631	1.1422	1.1804

(b) Previous TIs and van der Waals Areas, A (Å^2)

G	isomer	BC[25]	BX[25]	τ[28]	$B(V)$[29]	$DM(1)$[30]	A[2]
18	C_7	1.324 95	14.395 06	23.6523	1	13.4246	334.36
19	$2MC_6$	1.472 44	14.615 04	24.5823	112	14.7656	322.91
20	$3MC_6$	1.540 56	14.636 82	25.0396	297	15.0821	316.67
21	$24MMC_5$	1.737 82	14.836 80	25.5709	364	16.3631	309.97
22	$3EC_5$	1.802 68	14.658 60	25.4757	561	15.3665	303.15
23	$23MMC_5$	1.838 00	14.876 40	26.2208	1402	16.9492	303.11
24	$22MMC_5$	1.787 26	15.054 60	26.2438	3546	17.9498	306.53
25	$33MMC_5$	1.938 40	15.094 20	26.8844	5472	18.4853	297.20
26	$223MMC_4$	2.148 56	15.312 00	27.6039	8508	20.5470	292.89

(c) Intermolecular Ordering of Heptane Isomers G_{18}–G_{26}

RX, RJ, DJ, $B(V)$: 18, 19, 20, 21, 22, 23, 24, 25, 26
R^*: 18, 19, 20, 21, 22, 24, 23, 25, 26
RC, BC: 18, 19, 20, 21, 24, 22, 23, 25, 26
BX, τ, $DM(1)$: 18, 19, 20, 22, 21, 23, 24, 25, 26

Figure 8. Global indices in heptane isomers (M = methyl; E = ethyl).

Dobrynin found a very interesting pair of catacondensed benzenoid graphs,[27] whose centric numbering is given in Figure 7.

These graphs show identical **F** and **B**, but not **R**, matrices. All TIs based on these matrices are identical, except the *RC* index. It is obvious that the degeneracy of matrix invariants induce the degeneracy of all derived TIs.

The two graphs show indeed different **R** matrices, but their sums on columns are identical, so that the degeneracy of TIs appears at the operational stage (simple summation over all vertices in graph). Only a more sophisticated function, that is the centric *RC* index, may discriminate between the two graphs. **R** matrices and the *RC* index for these graphs are presented in Table I.

Global ordering in agreement with the c and x concepts, in the set of heptane isomers, is shown in Figure 8.

On considering Figure 8, it may be seen that in the set of isomeric heptanes the new TIs lead to several distinct orderings. It is interesting to observe that indices *RX*, *RJ*, and *DJ* give the same ordering as the "ideal" one advocated by Bertz,[29] which is identical with that induced by the J index.[21]

DISCUSSION

The idea of "seeing" the total graph environment of each vertex/atom, developed by us in connection with layer matrices **B** and **R** (see refs 22, 25, and 31 and this work) was also considered by Hall and Kier[28] in constructing the "*topological state*" matrix and related τ-indices. They have defined the overall structural relationships of a vertex i, making use of all paths joining that vertex with each of the other vertices in G, and the geometric mean, GM_{ij}, of δ_j-values for the chemical nature of each vertex j belonging to a given path of n_{ij} vertices. Their algorithm provides a set of τ-indices with high discriminating power which are useful in topological equivalence perception, and also in QSAR.

On comparing the intramolecular vertex ordering in $22MMC_9$, G_6, one can see that c operators, rc_i and bc_i, order vertices alternatively vs central vertex no. 5. Notice that such operators find the center of the graph according to the first criterion of Bonchev et al.[23] (minimal eccentricity).

Conversely, bx_i (an x operator), "sees" vertex no. 2 (with degree 4) as the most important vertex in that graph, the remaining vertices being ordered according to their increasing distance from vertex no. 2. A quite similar ordering is given by operators rx_i, rj_i, and dj_i (which, explicitly or implicitly, all take into account the vertex degree).

The r_i^* index behaves differently, alternating the vertices relative to vertex no. 4 (considered as center, in agreement with the second criterion of Bonchev et al.:[23] minimal distance sums).

Supplementary examples are given in Figure 9 for illustrating the equivalent vertex discriminating power of our indices. Four cubic graphs are taken from ref 28 along with the corresponding S_i^v values ($S_i^v = \sum_j GM_{ij}/n_{ij}$, as LOVIs within τ-index[28]).

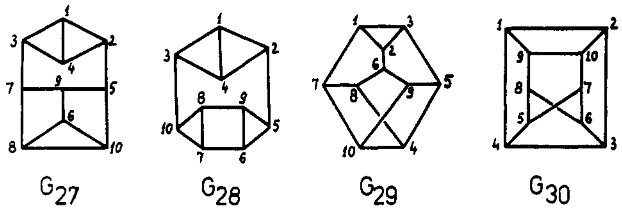
Despite different ranking of LOVIs, with one exception (G_{29}; operators R, RC, and RX), the equivalence classes were correctly found.

From the above examples, some remarks emerge:

(i) c operators enhance the contribution of more remote vertices, whereas x operators emphasize that of the nearer neighbors.

(ii) rc_i is the best c-type operator and bx_i is the best x-type one. Indices rx_i, rj_i, and dj_i are not pure x operators since they are based on D_i, which is a c parameter.

(iii) The r_i^* operator is a crude one and represents a compromise between c- and x-types.



G_{27}

	r_i^*		rc_i		rx_i		rj_i		dj_i		$S_i^{v\ 28}$
7	0.060 46	7	0.157 38	7	0.181 38	7	0.172 86	7	0.178 48	1	105.2143
5	0.060 46	5	0.157 38	5	0.181 38	5	0.172 86	5	0.178 48	4	105.2143
9	0.057 02	9	0.145 69	9	0.171 06	9	0.168 89	9	0.174 19	6	102.9429
10	0.053 89	3	0.142 13	10	0.161 68	10	0.161 00	10	0.165 91	8	99.9119
8	0.053 89	2	0.142 13	8	0.161 68	8	0.161 00	8	0.165 91	10	99.9119
3	0.053 86	10	0.141 45	3	0.161 59	3	0.159 42	3	0.164 33	9	96.4762
2	0.053 86	8	0.141 45	2	0.161 59	2	0.159 42	2	0.164 33	2	91.9786
4	0.048 63	4	0.098 36	4	0.145 89	6	0.151 51	6	0.155 79	3	91.9786
1	0.048 63	1	0.098 36	1	0.145 89	4	0.150 99	4	0.155 40	5	85.1429
6	0.046 43	6	0.092 66	6	0.139 31	1	0.150 99	1	0.155 40	7	85.1429

G_{28}

	r_i^*		rc_i		rx_i		rj_i		dj_i		$S_i^{v\ 28}$
10	0.056 98	10	0.147 91	10	0.170 96	10	0.166 31	10	0.171 49	1	104.5905
5	0.056 98	5	0.147 91	5	0.170 96	5	0.166 31	5	0.171 49	4	104.5905
9	0.053 95	3	0.142 77	9	0.161 85	9	0.163 35	9	0.168 27	6	98.4381
8	0.053 95	2	0.142 77	8	0.161 85	8	0.163 35	8	0.168 27	7	98.4381
7	0.053 95	9	0.141 94	7	0.161 85	7	0.163 35	7	0.168 27	8	98.4381
6	0.053 95	8	0.141 94	6	0.161 85	6	0.163 35	6	0.168 27	9	98.4381
3	0.053 89	7	0.141 94	3	0.161 67	3	0.160 41	3	0.165 31	2	91.4881
2	0.053 89	6	0.141 94	2	0.161 67	2	0.160 41	2	0.165 31	3	91.4881
4	0.051 14	4	0.140 35	4	0.153 42	4	0.156 14	4	0.160 77	5	89.2905
1	0.051 44	1	0.140 35	1	0.153 42	1	0.156 14	1	0.160 77	10	89.2905

G_{29}

	r_i^*		rc_i		rx_i		rj_i		dj_i		$S_i^{v\ 28}$
6	0.066 44	6	0.250 77	6	0.199 34	6	0.191 12	6	0.191 72	2	119.8452
9	0.062 31	7	0.161 54	9	0.186 93	8	0.187 11	9	0.187 68	1	118.5619
8	0.062 31	5	0.161 54	8	0.186 93	9	0.187 11	8	0.187 68	3	118.5619
7	0.062 30	9	0.160 45	7	0.186 91	7	0.183 21	7	0.183 76	4	118.0810
5	0.062 30	8	0.160 45	5	0.186 91	5	0.183 21	5	0.183 76	10	118.0810
10	0.058 65	3	0.150 81	10	0.175 96	2	0.179 73	2	0.180 26	8	117.1905
4	0.058 65	1	0.150 81	4	0.175 96	10	0.179 56	10	0.180 09	9	117.1905
2	0.058 65	10	0.150 29	2	0.175 96	4	0.179 56	4	0.180 09	5	113.8310
3	0.058 65	4	0.150 29	3	0.175 95	3	0.177 75	3	0.178 28	7	113.8310
1	0.058 65	2	0.150 29	1	0.175 95	1	0.177 75	1	0.178 28	6	113.1571

G_{30}

	r_i^*		rc_i		rx_i		rj_i		dj_i		$S_i^{v\ 28}$
8	0.060 64	10	0.161 77	8	0.181 94	8	0.181 93	8	0.187 50	1	122.9524
7	0.060 64	9	0.161 77	7	0.181 94	7	0.181 93	7	0.187 50	2	122.9524
6	0.060 64	4	0.161 77	6	0.181 94	6	0.181 93	6	0.187 50	5	121.6952
5	0.060 64	3	0.161 77	5	0.181 94	5	0.181 93	5	0.187 50	6	121.6952
10	0.060 61	8	0.160 75	10	0.181 83	10	0.180 10	10	0.185 63	7	121.6952
9	0.060 61	7	0.160 75	9	0.181 83	9	0.180 10	9	0.185 63	8	121.6952
4	0.060 61	6	0.160 75	4	0.181 83	4	0.180 10	4	0.185 63	3	120.9452
3	0.060 61	5	0.160 75	3	0.181 83	3	0.180 10	3	0.185 63	4	120.9452
2	0.057 15	2	0.151 47	2	0.171 46	2	0.174 87	2	0.180 09	9	120.9452
1	0.057 15	1	0.151 47	1	0.171 46	1	0.174 87	1	0.180 09	10	120.9452

Figure 9. Vertex equivalence perception in G_{27}–G_{30},[28] (LOVIs in decreasing order).

(iv) Heteroatoms and multiple bondings are located mainly as "*complexity*" subgraphs.

(v) An intercorreleting matrix (Figure 10) shows the *RC* index being a part of the x-type indices (τ-index[28] included). The proposed indices correlate well with van der Waals areas[2] in heptane isomers.

(vi) Our indices are good tools in vertex equivalence perception; when the connectivity is taken into account (*RJ* and *DJ*), their capability in discriminating the equivalence classes increases.

	RC	*R**	*RX*	*RJ*	*DJ*	τ^{28}	A^2
RC	1.0000	0.9748	0.9732	0.9736	0.9726	0.9690	0.9696
R		1.0000	0.9993	0.9995	0.9992	0.9969	0.9717
RX			1.0000	0.9999	0.9999	0.9972	0.9652
RJ				1.0000	0.9999	0.9974	0.9664
DJ					1.0000	0.9978	0.9632
τ^{28}						1.0000	0.9527
A^2							1.0000

Figure 10. Intercorrelating matrix between TIs and van der Waals areas, A (Å^2), in heptane isomers.

CONCLUSIONS

The **R** matrix (layer of distance sums) represents an extension of **F** (layer of neighbors/distance frequencies) and **B** (layer of degrees) matrices and is suitable for topological index design. Its discriminating power surpasses that of the previous layer matrices.

The intramolecular ordering of vertices, in c or x terms leads to real-number LOVIs, which can be used in QSAR/QSPR studies (see ref 32).

ACKNOWLEDGMENT

Thanks are addressed to Dr. A. Dobrynin, Institute of Mathematics, Russian Academy of Sciences, Siberian Branch, Novosibirsk, for a preprint. We also thank Dr. D. Horvath, Department of Chemistry, "Babes-Bolyai" University, Cluj, Romania, for computer assistance.

REFERENCES AND NOTES

(1) Balaban, A. T. Using real numbers as vertex invariants for third generation topological indexes. *J. Chem. Inf. Comput. Sci.* **1991**, *32*, 23–28. Balaban, A. T.; Balaban, T. S. Correlation using topological indexes based on real graph invariants. *J. Chim. Phys.* **1992**, *89*, 1735–1745.

(2) Labanowski, J.; Motoc, I.; Dammkoehler, R. A. The physical meaning of topological indices. *Comput. Chem.* **1991**, *15*, 47–53.

(3) Rouvray, D. H. The limits of applicability of topological indices. *J. Mol. Struct. (THEOCHEM)* **1989**, *54*, 187–201.

(4) Ivanciuc, O.; Balaban, T. S.; Balaban, A. T. Reciprocal distance matrix, related local vertex invariants and topological indices. *J. Math. Chem.*, in press.

(5) Filip, P.; Balaban, T. S.; Balaban, A. T. A new approach for devising local graph invariants: derived topological indices with low degeneracy and good correlation ability. *J. Math. Chem.* **1987**, *1*, 61–83.

(6) Sarkar, R.; Roy, A. B.; Sarkar, P. K. Topological information content of genetic molecules-1. *Math. Bio. Sci.* **1978**, *39*, 299–312.

(7) (a) Roy, S. K.; Basak, S. C.; Raychaudhury, C.; Roy, A. B.; Ghosh, J. J. A quantitative structure-activity relationship study of N-alkylnorketobemidones and triazinones using structural information content. *Arzneimt.-Forsch.* **1982**, *32*, 322–325. (b) Roy, S. K.; Basak, S. C.; Raychaudhury, C.; Roy, A. B.; Ghosh, J. J. The utility of information content, structural information content, hydrophobicity and van der Waals volume in the design of barbiturates and tumor inhibitory triazines. *Arzneimt.-Forsch.* **1983**, *33*, 352–356. (c) Roy, S. K.; Basak, S. C.; Raychaudhury, C.; Roy, A. B.; Ghosh, J. J. A quantitative structure-activity relationship (QSAR) analysis of carbamoyl piperidines, barbiturates and alkanes using information-theoretic topological indices. *Indian J. Pharmacol.* **1982**, *13*, 301–312.

(8) (a) Randić, M. On Computation of Optimal Parameters for Multivariate Analysis of Structure-Property Relationship. *J. Comput. Chem.* **1991**, *12*, 970–980. (b) Randić, M. Generalized Molecular Descriptors *J. Math. Chem.* **1991**, *7*, 155–168.

(9) Raychaudhury, C.; Roy, S. K.; Ghosh, J. J.; Roy, A. B.; Basak, S. C. Discrimination of isomeric structures using information theoretic topological indices. *J. Comput. Chem.* **1989**, *5*, 581–588.

(10) Klopman, G.; Raychaudbury, C. Vertex indices of molecular graphs in structure-activity relationships: a study of the convulsant-anticonvulsant activity of barbiturates and the carcinogenicity of unsubstituted polycyclic aromatic hydrocarbons. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 12–19.

(11) Klopman, G.; Raychaudbury, C.; Henderson, R. V. A new approach to structure-activity relationships using distance information content of graph vertices: a study with phenylalkylamines. *Math. Comput. Modelling* **1988**, *11*, 635–640.

(12) Balaban, A. T.; Balaban, T. S. New vertex invariants and topological indices of chemical graphs based on information distances. *J. Math. Chem.* **1991**, *8*, 383–397.

(13) Balaban, A. T.; Ciubotariu, D.; Medeleanu, M. Topological indices and real vertex invariants based on graph eigenvalues or eigenvectors. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 517–523.

(14) Balaban, A. T.; Catana, C. Search for non-degenerate real vertex invariants and derived topological indices. *J. Comput. Chem.* **1992**, *14*, 155–160.

(15) (a) Wiener, H. Structural determination of paraffin boiling point. *J. Am. Chem. Soc.* **1947**, *69*, 17. (b) Wiener, H. Correlation of heats of isomerization, and differences in heats of vaporization of isomers, among the paraffin hydrocarbons. *J. Am. Chem. Soc.* **1947**, *69*, 2636–2638.

(16) (a) Randić, M. On characterization of molecular branching *J. Am. Chem. Soc.* **1975**, *97*, 6609–6615. (b) Randić, M. Search for optimal molecular descriptors *Croat. Chem. Acta* **1991**, *64*, 43–54. (c) Randić, M. Correlation of enthalpy of octanes with Orthogonal Connectivity Indices *J. Mol. Struct. (THEOCHEM)* **1991**, *233*, 45–59. (d) Randić, M. Orthogonal molecular descriptors *New J. Chem.* **1991**, *15*, 517–525. (e) Randić, M. Resolution of ambiguities in structure-property studies by use of orthogonal descriptors. *J. Chem. Inf. Comput. Sci.* **1991**, *31*, 311–320.

(17) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Design*; Academic Press: New York, 1976. Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure-Activity Analysis*; RSP and Wiley: New York, 1986.

(18) Bonchev, D.; Trinajstic', N. Information theory, distance matrix, and molecular branching. *J. Chem. Phys.* **1977**, *67*, 4517–4533.

(19) Ivanciuc, O. Definition of a vertex topological index in the case of 4-trees. *Rev. Roum. Chim.* **1989**, *34*, 1361–1368.

(20) Basak, S. C.; Magnuson, V. R. Determining structural similarity of chemicals using graph-theoretic indices. *Discrete Appl. Math.* **1988**, *19*, 17–44.

(21) (a) Balaban, A. T. Highly discriminating distance-based topological index. *Chem. Phys. Lett.* **1982**, *89*, 399–404. (b) Balaban, A. T. Challenging problems involving benzenoid polycyclics and related systems. *Pure Appl. Chem.* **1982**, *54*, 1075–1096. (c) Balaban, A. T. Topological indices based on topological distances in molecular graphs. *Pure Appl. Chem.* **1983**, *55*, 199–206. (d) Balaban, A. T. Topological index J for heteroatom-containing molecules taking into account periodicities of element properties. *MATCH* **1986**, *21*, 115–122.

(22) Diudea, M. V.; Minailiuc, O.; Balaban, A. T. Regressive vertex degree (new graph invariants) and derived topological indices. *J. Comput. Chem.* **1991**, *12*, 527–535.

(23) Bonchev, D.; Balaban, A. T.; Randić, M. The graph center concept for polycyclic graphs. *Int. J. Quantum Chem.* **1981**, *19*, 61–82.

(24) Diudea, M. V.; Parv, B. A new centric connectivity index (CCI). *MATCH* **1988**, *23*, 65–87.

(25) Diudea, M. V.; Horvath, D.; Kacso, I. E.; Minailiuc, O. M.; Parv, B. Centricities in molecular graphs. The MOLCEN algorithm. *J. Math. Chem.* **1992**, *11*, 259–270.

(26) Diudea, M. V.; Silaghi-Dumitrescu, I. Valence group electronegativity as a vertex discriminator *Rev. Roum. Chim.* **1989**, *34*, 1175–1182.

(27) Dobrynin, A. Degeneracy of some matrix invariants and derived topological indices. *J. Math. Chem.*, in press.

(28) Hall, L. H.; Kier, L. B. Determination of Topological Equivalence in Molecular Graphs from the Topological State. *Quantum Struct.-Act. Relat.* **1990**, *9*, 115–131.

(29) Bertz, S. H. Branching in Graphs and Molecules. *Discrete Appl. Math.* **1988**, *19*, 65–83.

(30) Balaban, A. T.; Ciubotariu, D.; Ivanciuc, O. Design of topological indices. Part 2. Distance measure connectivity indices. *MATCH* **1990**, *25*, 41–70.

(31) Diudea, M. V.; Kacso, I. E.; Minailiuc, O. M. Y-indices in homogeneous dendrimers. *MATCH*, in press.

(32) Horvath, D.; Diudea, M. V.; Graovac, A. 3D-Distance matrices and related topological indices. *Comput. Chem.*, submitted for publication.