Molecular Topology. 17. Layer Matrices of Walk Degrees

Mircea V. Diudea*
Department of Chemistry, "Babes-Bolyai" University, 3400 Cluj, Romania

Mihai Topan
Department of Computers, Technical University, Cluj, Romania

Ante Graovac
Ruder Boskovic' Institute, 41001 Zagreb, Croatia

Received November 16, 1993®

Layer matrices of walk degrees, $\mathbf{LW}^{(e)}$ of various elongation (e), proposed by Diudea,[1] are further exploited. Walk degrees (e.g., atomic walk counts), $w_i^{(e)}$, are calculated by a simple summation procedure (analogous with the Morgan's algorithm[2]) iterated on the adjacency/connectivity matrix and next weighted by an electronegativity factor. Formulas for counting walk degrees in some particular graphs are given. The matrices $\mathbf{LW}^{(e)}$ are used as a ground for building a centric topological index, $C(\mathbf{LW}^{(e)})$ and for investigating the topological symmetry in some sets of molecular graphs. A computer program TOP-W for calculating walk degrees and layer matrices $\mathbf{LW}^{(e)}$ is also presented.

INTRODUCTION

The occurrence of nonsymmetrical matrices[1,3-7] in molecular topology in the last few years was associated with real number descriptors in 2D and 3D spaces and was aimed to improve representation of molecular graphs. These arrays, called layer matrices[1,3,5,6] are constructed on the basis of partitions of a graph, G, with respect to its vertices. The i-relative partition of graph, $G(i)$, can be written as[1]

$$G(i) = \{G(u)_j; j = 0, 1, ..., \mathrm{ecc}_i\} \quad (1)$$

where ecc_i (eccentricity of vertex i) is the maximal distance from i in graph G. The jth layer of vertex i, $G(u)_j$, is the set of all u vertices in G being at the distance j from i, i.e.

$$G(u)_j = \{u:\ d_{iu} = j\} \quad (2)$$

Pictorial representation of partitions is shown in Figure 1 for the graph G_1.

In the present work we advocate a further use of random walks in the characterization of molecular graphs by means of layer matrices $\mathbf{LW}^{(e)}$.[1] A simple summation algorithm for calculating walk degrees $w_i^{(e)}$ and formulas for particular graphs are given. Topological symmetry, in a set of graphs, is investigated by means of the layer matrices $\mathbf{LW}^{(e)}$. A computer program for counting walk degrees and generating layer matrices $\mathbf{LW}^{(e)}$ is also presented.

ITERATIVE ALGORITHM FOR COMPUTING WEIGHTED WALKS

Random walk, $w^{(e)}$, is defined[8-10] as a continuous sequence of edges, allowing edges and vertices to be revisited. The terminal vertex of a random walk is in general different from the starting vertex, but they could coincide as well. The number of edges traversed is called the elongation of walk, e.

The number of walks of elongation e starting at the vertex i is called the walk degree $w_i^{(e)}$ (e.g., atomic walk count, see refs 11 and 12). Its evaluation is a simple task[13-15] as it is easily obtained from the eth power $\mathbf{A}^e$ of the adjacency matrix $\mathbf{A}$ as

$$w_i^{(e)} = \sum_{u \in V(G)} (a_{i,u})^e \quad (3)$$

where $V(G)$ is the vertex set of G, and $(a_{i,u})^e$ are the entries of $\mathbf{A}^e$.

The calculus of $w_i^{(e)}$ can also be accomplished from vertex degrees k (which are identical with the walk degrees of elongation 1, $w_i^{(1)}$) by iterative summation over all neighbors, as Morgan[2] proposed by its extended connectivities, ECs. Razinger[15] first, and very recently Rücker and Rücker[11] and Figueras,[13] demonstrated the identity between EC_i and $w_i^{(e)}$.

We propose here an algorithm which, basically, follows Morgan's algorithm[2] (see also ref 13).

The algorithm is implemented on the connectivity matrix, $\mathbf{C}$ (in particular, on the adjacency matrix, $\mathbf{A}$), whose entries are $c_{ii} = 0$ and c_{ij} = connectivity (bond orders).

One defines the matrix $\mathbf{CW}^{(e)}$ as the sum between the $\mathbf{C}$ matrix and a diagonal matrix, $\mathbf{W}^{(e)}$ (of walk degrees)

$$\mathbf{C} + \mathbf{W}^{(e)} = \mathbf{CW}^{(e)} \quad (4)$$

where the matrix elements of $\mathbf{CW}^{(e)}$ are given by

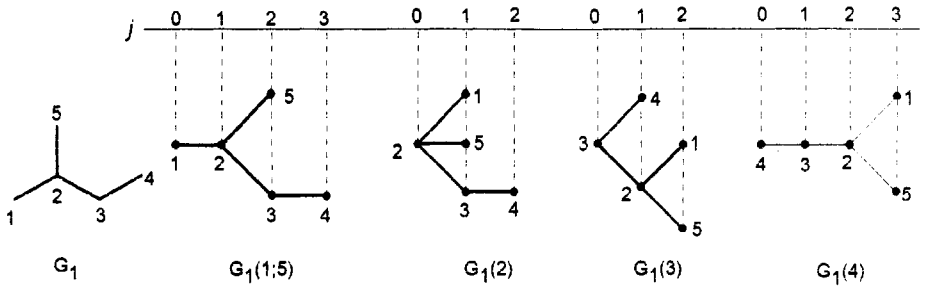
$$cw_{ii}^{(e+1)} = \sum_j (c_{ij} cw_{jj}^{(e)}) \quad \text{with } cw_{jj}^{(0)} = 1 \quad (5)$$

$$cw_{ij}^{(e+1)} = cw_{ij}^{(e)} = c_{ij} \quad (6)$$

One observes that the diagonal entries $cw_{ii}^{(e)}$ are just the walk degrees $w_i^{(e)}$. The algorithm accounts for multiple bonds

* Abstract published in *Advance ACS Abstracts*, September 1, 1994.

0095-2338/94/1634-1072$04.50/0



$G_1(1) = \{\{1\},\{2\},\{3,5\},\{4\}\}$
$G_1(2) = \{\{2\},\{1,3,5\},\{4\}\}$
$G_1(3) = \{\{3\},\{2,4\},\{1,5\}\}$
$G_1(4) = \{\{4\},\{3\},\{2\},\{1,5\}\}$
$G_1(5) = \{\{5\},\{2\},\{1,3\},\{4\}\}$

Figure 1. Relative partitions of G_1.

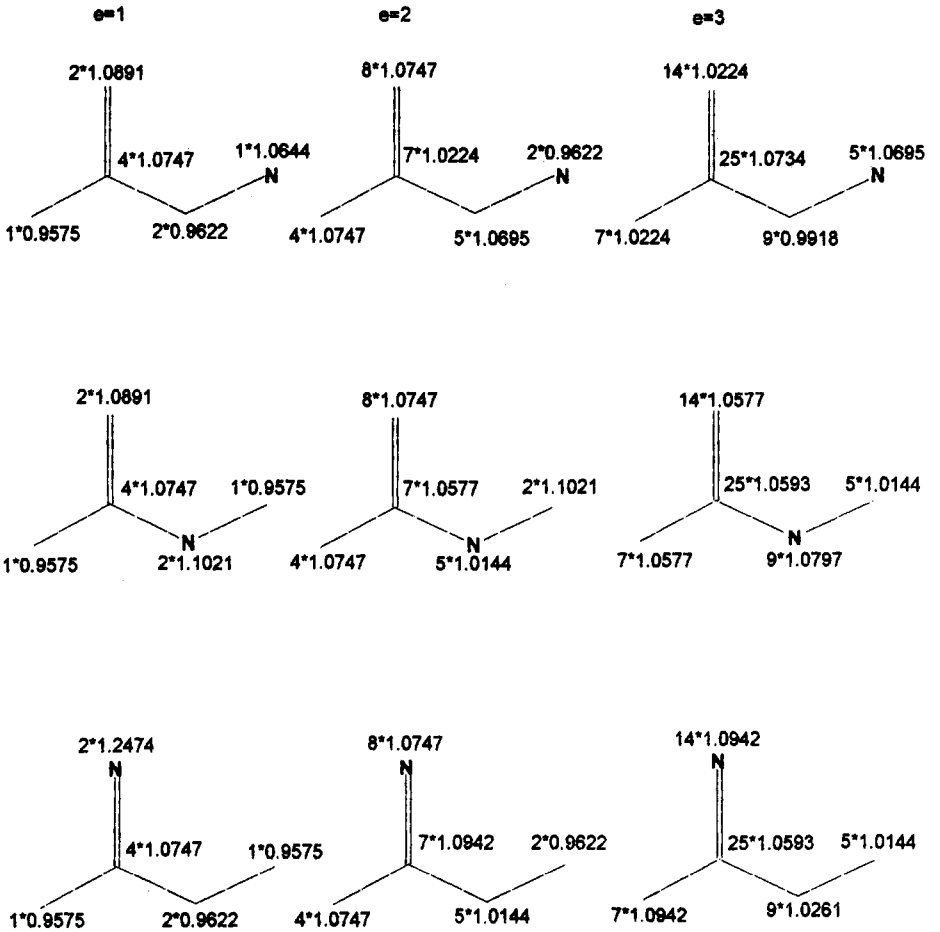


Figure 2. The count of weighted walk degrees $tw_i^{(e)}$ (as $cw_{ii}^{(e)}t_i^{(e)}$) for e = 1–3; the values $t_i^{(1)}$ are taken from ref 16.

by means of the c_{ij} entries. For specifying the chemical nature of vertices we introduced a weighting factor t_i. Thus, the weighted walk degree tw_i is given by

$$tw_i^{(e)} = cw_{ii}^{(e)}t_i^{(e)} \tag{7}$$

where the weighting factor consists of a carbon related Sanderson type electronegativity.[16] It is taken as a geometric mean between the values corresponding to the first neighbors of atom i

$$t_i^{(e+1)} = (\prod_j (t_j^{(e)})^{c_{ij}})^{1/\sum_j c_{ij}} \tag{8}$$

with $t_j^{(1)}$ as in ref 16 and in the Appendix. For nonweighted walks $t_i = 1$ and $tw_i^{(e)} = w_i^{(e)}$. Figure 2 shows in a "pictorial" manner (by means of some vertex weighted graphs) the iterative summation procedure for counting weighted walk degrees.

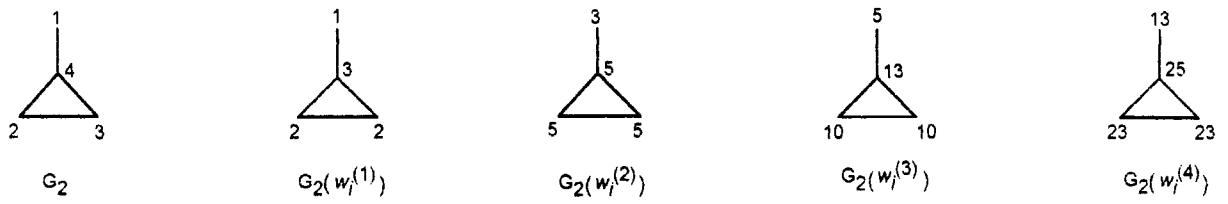
LAYER MATRICES OF WALK DEGREES, $\mathbf{LW}^{(e)}$

Let us now consider a given vertex i and its jth layer cf. eq 2. The sum of walk degrees $w_u^{(e)}$ for all vertices u in the jth layer of i, $G(u)_j$, defines the matrix element in the ith row and jth column of the layer matrix $\mathbf{LW}^{(e)}$

$$lw_{i,j}^{(e)} = \sum_{u \in G(u)_j} w_u^{(e)} \tag{9}$$

and subsequently, the layer matrix $\mathbf{LW}^{(e)}$

$$\mathbf{LW}^{(e)} = \{lw_{i,j}^{(e)};\ i \in [1,n];\ j \in [0,d]\} \tag{10}$$



G_2 $G_2(w_i^{(1)})$ $G_2(w_i^{(2)})$ $G_2(w_i^{(3)})$ $G_2(w_i^{(4)})$

partitions	$LW^{(1)}$			$LW^{(2)}$			$LW^{(3)}$			$LW^{(4)}$		
$G_2(1)$: {{1},{4},{2,3}}	1	3	4	3	5	10	5	13	20	13	25	46
$G_2(2)$: {{2},{3,4},{1}}	2	5	1	5	10	3	10	23	5	23	48	13
$G_2(3)$: {{3},{2,4},{1}}	2	5	1	5	10	3	10	23	5	23	48	13
$G_2(4)$: {{4},{1,2,3},{0}}	3	5	0	5	13	0	13	25	0	25	59	0
$GW^{(e)}$:	4			9			19			42		

Figure 3. Weighted graphs, $G_2\{w_i^{(e)}\}$; partitons, $G_2(i)$; matrices $\mathbf{LW}^{(e)}$; and $GW^{(e)}$ values for the graph G_2.

with $n = |V(G)|$ and d being the diameter of G, i.e., the largest distance in G; (e) stands for a selected elongation.

Thus, the matrix $\mathbf{LW}^{(e)}$ is a rectangular matrix (and accordingly nonsymmetric) with n rows and $(d + 1)$ columns. The first column refers to vertices at distance zero in each relative partition of the graph (0th layer), i.e., the vertices themselves. Its elements represent just the walk degrees (of elongation e), $w_i^{(e)}$. When weighted walks $tw_i^{(e)}$ are used, the entries in $\mathbf{LW}^{(e)}$ will be real figures (see below).

Further, as far as rows are concerned, one easily sees that the first $(\mathrm{ecc}_i + 1)$ elements in some ith row are nonzero; the remaining $(d - \mathrm{ecc}_i)$ elements in that row have to be zero.

Figure 3 shows the partitions for the graph G_2 along with the corresponding first four (e = 1–4) matrices $\mathbf{LW}^{(e)}$.

By adding all vertex degrees $w_i^{(e)}$ of graph G, each walk of elongation e is counted twice. Therefore, the following quantity

$$GW^{(e)} = {}^1/_2\sum_{i=1}^{n} w_i^{(e)} = {}^1/_2\sum_{i=1}^{n} lw_{i0}^{(e)} = {}^1/_2\sum_{j=0}^{d} lw_{ij}^{(e)} \quad (11)$$

represents the total number of walks of elongation e in G (e.g., the graph walk count or molecular walk count). The values $GW^{(e)}$, e = 1–4, for the graph G_2 are also given in Figure 3. One can see that they equal to the half sum of the entries in the first column in $\mathbf{LW}^{(e)}(G_2)$ as well as in each row of the corresponding matrices. Some other relationships are given

$$\sum_i lw_{i1}^{(e)} = \sum_i lw_{i0}^{(e+1)} \quad (12)$$

$$\sum_i (lw_{i1}^{(e)} + lw_{i2}^{(e)}) = \sum_i lw_{i1}^{(e+1)} \quad (13)$$

$$\sum_i (lw_{i(jm-2)}^{(e)}) = \sum_i (lw_{i(jm-1)}^{(e+1)}) \quad (14)$$

$$\sum_i (lw_{i(jm-1)}^{(e)} - lw_{i(jm)}^{(e)}) = \sum_i (lw_{i(jm)}^{(e+1)}) \quad (15)$$

where jm equals the diameter of graph.

Relation 12 holds true in graphs without multiple bonds; relation 13 holds true for trees, and relations 14 and 15 work just for paths (n-alkanes).

COUNTING OF WALKS IN SPECIAL GRAPHS

Special graphs show regularities in the counting of their walks. We present here some relationships along with the partitions, as **LC** matrices[1] (which entries $lc_{i,j} = \sum_{u \in G(u)_j} 1_u$ means the cardinality of layers $G(u)_j$ at distance j from the vertex i) and the corresponding $\mathbf{LW}^{(e)}$ matrices, in line form[1] (i.e., the rows are written in decreasing order of their length and in lexicographic order in case of equal length).

Stars, $K_{1,k}$; $1 + k = n$

$$w_{\mathrm{center}}^{(e)} = k^{p+z}; \quad w_{\mathrm{radius}}^{(e)} = k^p; \quad e = 2p + z \quad (16)$$

$z = 0$ for e = even; $z = 1$ for e = odd

$$GW^{(e)} = {}^1/_2k^p(k^z + k) \quad (17)$$

$$\mathbf{LC} = \{(1;k);\ k[1;1;(k-1)]\} \quad (18)$$

$$\mathbf{LW}^{(e)} = \{(k^{p+z};k^{p+1});\ k[k^p;k^{p+z};(k-1)k^p]\} \quad (19)$$

Cycles, Cy_n; $n = 2p + z$

$$w_i^{(e)} = 2^e \quad (20)$$

$$GW^{(e)} = n2^{e-1} \quad (21)$$

$$\mathbf{LC} = \{(2p + z)(1;2 \times 1^j;1 - z)\} \quad (22)$$

$$\mathbf{LW}^{(e)} = \{(2p + z)[2^e;2^{e+1} \times 1^j;2^e(1 - z)]\}; \quad j \in [1, p + z - 1] \quad (23)$$

Complete graphs, K_n

$$w_i^{(e)} = k^e = (n-1)^e \quad (24)$$

$$GW^{(e)} = {}^1/_2nk^e = {}^1/_2n(n-1)^e \quad (25)$$

$$\mathbf{LC} = \{n[1;(n-1)]\} \quad (26)$$

$$\mathbf{LW}^{(e)} = \{n[k^e;(n-1)k^e]\} = \{n[(n-1)^e;(n-1)^{e+1}]\} \quad (27)$$

Paths (n-alkanes); layer matrices as sums on columns

$$\mathbf{LC} = \{n;2(n-j)\};\ j \in [1,n-1] \quad (28)$$

$$\mathbf{LW}^{(1)} = \{2q;4(q-j) + 2\};\ j \in [1,q] \quad (29)$$

$$GW^{(1)} = q \quad (30)$$

with q being the number of edges in graph.

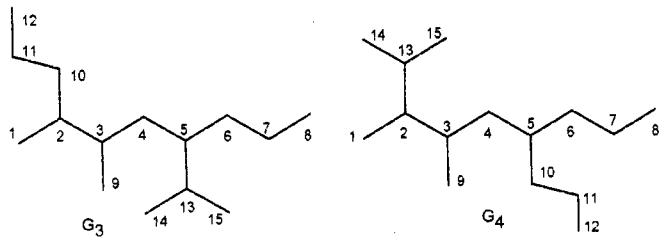
Regular graphs

$$w_i^{(e)} = k^e \quad (31)$$

$$GW^{(e)} = {}^1/_2nk^e \quad (32)$$

TOPOLOGICAL SYMMETRY

Many procedures for finding the automorphism group and orbits in graphs have been introduced.[17-22] Some of them,



G_3; $c(\mathbf{LW}^{(1)})_i$	[dsp = 20]	G_4; $c(\mathbf{LW}^{(1)})_i$
1 0.1131899		1 0.1114086
2 0.1371097		2 0.1347560
3 0.1664271		3 0.1632493
4 0.1688010		4 0.1995842
5 0.1423490		5 0.1632493
6 0.1187676		6 0.1323354
7 0.1009639		7 0.1102646
8 0.0872190		8 0.0937584
9 0.1331460		9 0.1307811
10 0.1150171		10 0.1323354
11 0.0979365		11 0.1102646
12 0.0846760		12 0.0937584
13 0.1185889		13 0.1128431
14 0.1002372		14 0.0956220
15 0.1002372		15 0.0956220
$C(\mathbf{LW}^{(1)})$: 1.784665888		$C(\mathbf{LW}^{(1)})$: 1.879832594

Figure 4. Discrimination of vertices in the graphs G_3 and G_4 by the centric index, $c(\mathbf{LW}^{(1)})_i$ (vertices 10–15 being pairwise isocodal—see ref 12).

such as the well-known HOC algorithm,[17,22] are directly related to the walk counts in molecular graphs.

In a recent paper, Rücker and Rücker,[11] tested the discriminating ability of atomic walk counts (actually walk degrees $w_i^{(e)}$) and related parameters (or sequences of such as invariants) on various sets of trees. They recognized the limits of $w_i^{(e)}$ (and of atomic walk count sequences) in discriminating the endospectral or isospectral atoms. More successful were their ordered atomic walk count sums sequences, OAWCS, proposed by analogy with the OSC (ordered structural code) of Barisz and Trinajstic.[23] Despite the OAWCSs differentiation of all the tested trees (including the Ivanciuc/Balaban[12] trees having pairwise identical OSCs), they require the counting of walks until the elongation $n - 1$ (in agreement with the Cayley–Hamilton theorem).

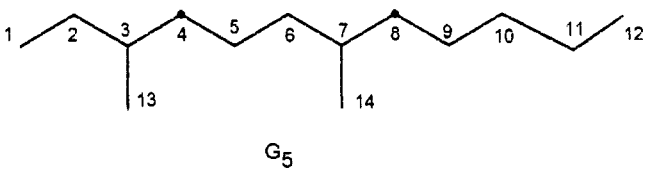
To elude this drawback we used the layer matrices $\mathbf{LW}^{(e)}$ and a centric index, $c(\mathbf{LW}^{(e)})_i$, constructed on

$$c(\mathbf{LW}^{(e)})_i = [\sum_{j=1}^{ecc_i} (lw_{ij})^{j/dsp}]^{-1} \qquad (33)$$

where dsp is a specified topological distance, usually larger than the diameter of the graph[1,5,6,20] (here dsp = 10, unless otherwise specified).

Summation over all vertices in G will provide the global topological index, $C(\mathbf{LW}^{(e)})$ (denoted by capitals).

We tested the ability of $\mathbf{LW}^{(e)}$, by means of $c(\mathbf{LM}^{(e)})_i$, to discriminate the pairs of isocodal vertices (labeled from 10 to 15) in the graphs G_3 and G_4 of Ivanciuc/Balaban:[12] at elongation $e = 1$, they are already separated (see Figure 4). Endospectral vertices (like no. 4 and no. 8 in G_5, which show the identical walk degree sequence 2, 5, 9, 22, 39, 95, 168, ...) are also easily discriminated (Figure 5). Our centric index $c(\mathbf{LW}^{(e)})_i$ can say about the vertex centricity more than the simple walk count does.[11] Thus vertex no. 4 is less central than vertex no. 8. This fact is supported by the Bonchev's 1D–3D centric criteria[8] (minimal vertex eccentricity; minimal vertex distance sum; minimal number of occurrences of the



(a) from layer matrix of walk degrees, $\mathbf{LW}^{(e)}$ (only the rows for the vertices no. 4 and no. 8 which show identical walk degree sequence: 2, 5, 9, 22, 39, 95, 168, ...), at elongation e = 1–3:

no. 4:
$\mathbf{LW}^{(1)}$: 2, 5, 5, 4, 3, 2, 2, 2, 1 $c(\mathbf{LW}^{(1)})_i = 0.106042$
$\mathbf{LW}^{(2)}$: 5, 9, 12, 7, 8, 4, 4, 3, 2 $c(\mathbf{LW}^{(2)})_i = 0.091148$
$\mathbf{LW}^{(3)}$: 9, 22, 21, 17, 14, 9, 7, 6, 3 $c(\mathbf{LW}^{(3)})_i = 0.078761$
no. 8:
$\mathbf{LW}^{(1)}$: 2, 5, 5, 4, 3, 3, 3, 1, 0 $c(\mathbf{LW}^{(1)})_i = 0.118459$
$\mathbf{LW}^{(2)}$: 5, 9, 12, 7, 7, 5, 7, 2, 0 $c(\mathbf{LW}^{(2)})_i = 0.1022951$
$\mathbf{LW}^{(3)}$: 9, 22, 21, 16, 12, 12, 12, 4, 0 $c(\mathbf{LW}^{(3)})_i = 0.088687$

(b) from spectral data:
first eigenvector: 7, 6, (3;5), (4;8), 9, (2;14), (10;13), (1;11), 12

(c) from layer matrix of distance sums, **LDS** (supported by the Bonchev's centric criteria:[8] minimal eccentricity (min ecc_i); minimal distance sum (min DS_i) in squared brackets):

$c(\mathbf{LDS})_i$:	7	6	5	8	14	4	9	3	10	2	13	11	1	12
[min ecc_i:	6	6	7	7	7	8	8	9	9	10	10	10	11	11]
[min DS_i:	42	42	44	46	54	48	52	54	60	64	66	70	76	82]

LDS (G_5): missing entries are zeroes

1	76	64	54	114	44	42	42	100	52	60	70	82
2	64	130	114	44	42	42	100	52	60	70	82	
3	54	178	120	42	42	100	52	60	70	82		
4	48	98	172	118	100	52	60	70	82			
5	44	90	96	230	128	60	70	82				
6	42	86	148	106	190	146	82					
7	42	142	96	108	124	212	76					
8	46	94	156	114	130	54	130	76				
9	52	106	112	178	44	48	54	130	76			
10	60	122	128	42	96	44	48	54	130	76		
11	70	142	52	46	42	96	44	48	54	130	76	
12	82	70	60	52	46	42	96	44	48	54	130	76
13	66	54	112	120	42	42	100	52	60	70	82	
14	54	42	88	96	108	124	212	76				

Figure 5. Centricities in G_5.

largest distance) and by another centric index, $c(\mathbf{LDS})_i$, constructed by us on the layer matrix of distance sums,[1,6] **LDS**, and tested earlier[6] in this respect (see Figure 5). For comparison, the principal eigenvector coefficients are also given.

Next we tested a set of ten trees (G_6–G_{15}), constructed by Ivanciuc/Balaban so that they show pairwise identical distance degree sequence[24] and degenerate indices based on information on distances (Figure 6).

The minimal elongation which gave distinct global walk count, $GW^{(e)}$, values was 8, and for a connectivity index (see ref 25), $\chi W^{(e)}$, it was 4 (see Table 1). Elongation 4 (or higher) was needed for separating these graphs by means of ordered walk degree sequences (of selected elongation—see Table 2). In contrast, the centric index $C(\mathbf{LW}^{(e)})$ gave distinct values just from $e = 2$ (see Table 1).

The topological symmetry perception by three walk related parameters (the walk degree $w_i^{(e)}$, the connectivity per vertex index $\chi W_i^{(e)}$, and the centric index $c(\mathbf{LW}^{(e)})_i$) we tested on a set of cuneane congeners having double bonds and nitrogen atoms, respectively (see Figure 7).

The orbits of automorphism are given in Table 3, at the minimal separating elongation, e. One can see that the correct orbits were not found (at that stage) by the first two parameters. Table 4 shows the values for the three parameters in the case of the graphs G_{20} and G_{21} at $e = 3$. The layer

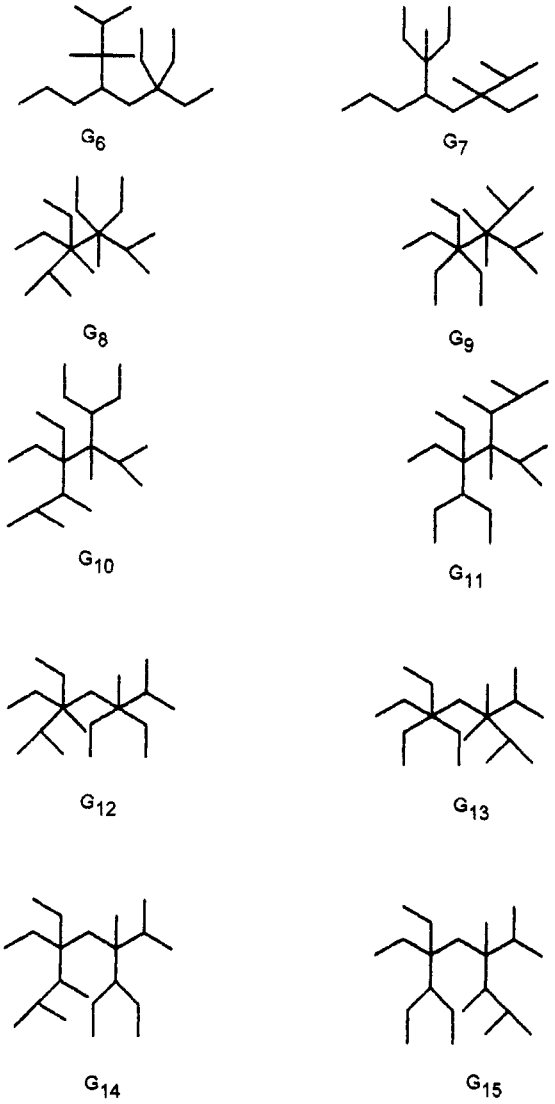


Figure 6. Pairs of graphs with degenerate indices based on information on distances (ref 24).

Table 1. Topological Walk Related Parameters for the Graphs G_6–G_{15} (at the Discriminating Elongation, e; dsp = 20)

graph	$C(\mathbf{LW}^{(2)})$	$\chi W^{(4)}$	e	$GW^{(e)}$	$e = n$	$GW^{(e)}$
G_6	2.018 812	1.377 73	8	7379	17	19 263 438
G_7	1.997 183	1.377 51	8	7375	17	19 178 006
G_8	2.892 750	0.837 20	8	18641	17	152 345 669
G_9	2.895 929	0.837 24	8	18645	17	152 412 977
G_{10}	2.356 520	1.289 62	9	31726	19	378 665 905
G_{11}	2.356 392	1.289 63	9	31730	19	378 892 113
G_{12}	2.424 271	1.013 98	8	14672	18	208 060 496
G_{13}	2.426 273	1.014 10	8	14676	18	208 385 984
G_{14}	2.048 772	1.458 12	9	25834	20	544 335 808
G_{15}	2.048 973	1.458 13	9	25838	20	545 025 778

matrix $\mathbf{LW}^{(3)}(G_{21})$ built by using weighted $tw_i^{(e)}$ values is given in the following:

matrix $\mathbf{LW}^{(3)}(G_{21})$; with $tw_i^{(e)}$ values (missing entries are zeroes)

1	29.0624	83.5876	82.8529	26.9136
2	26.9136	85.0047	110.5012	
3	27.6116	80.7408	86.4525	27.6116
4	27.6116	86.4525	80.7406	27.6116
5	28.3277	81.4388	112.6500	
6	26.9136	82.8529	83.5876	29.0624
7	29.0624	83.5876	109.7665	
8	26.9136	83.5876	111.9153	

DISCUSSION

Layer matrices, $\mathbf{LM}^{(e)}$, introduced in ref 1, has proved to be important tools in molecular graph description, namely in vertex and graph ordering, in checking the graph isomorphism or in perception of topological symmetry of a graph. Some topological indices built on layer matrices were successfully tested[1,26] in QSAR/QSPR studies.

Among $\mathbf{LM}^{(e)}$, the $\mathbf{LW}^{(e)}$ matrices are of particular interest. They represent the extension of **LK** (layer matrix of degrees, $k = w_i^{(1)}$) to those based on walk degrees of various elongation e, $w_i^{(e)}$. The simple summation algorithm proposed here enables the rapid calculation (manually or automatically performed) of walk degrees, $w_i^{(e)}$, and hence the total number of walks, $GW^{(e)}$, in the graph. Moreover, it accounts both for

Table 2. Ordered Walk Degree Sequences (at the Minimal Discriminating Elongation, e) in the Graphs G_6–G_{15}, Having Pairwise Degenerate Indices Based on Information on Distances—See Ref 24

graph	e	
G_6	4	$7^{(1)}$; $10^{(3)}$; $14^{(3)}$; $22^{(2)}$; $27^{(4)}$; $34^{(1)}$; $42^{(1)}$; $46^{(1)}$; $49^{(1)}$; $50^{(1)}$
G_7	4	$7^{(1)}$; $10^{(3)}$; $14^{(3)}$; $22^{(2)}$; $27^{(4)}$; $34^{(1)}$; $42^{(1)}$; $48^{(2)}$; $49^{(1)}$
G_8	4	$15^{(4)}$; $19^{(4)}$; $37^{(2)}$; $43^{(4)}$; $51^{(2)}$; $99^{(2)}$
G_9	4	$15^{(4)}$; $19^{(4)}$; $37^{(2)}$; $43^{(4)}$; $51^{(2)}$; $97^{(1)}$; $101^{(1)}$
G_{10}	5	$23^{(2)}$; $29^{(2)}$; $34^{(2)}$; $41^{(2)}$; $51^{(1)}$; $59^{(2)}$; $76^{(1)}$; $79^{(1)}$; $87^{(2)}$; $110^{(1)}$; $122^{(2)}$; $193^{(1)}$; $195^{(1)}$
G_{11}	5	$23^{(2)}$; $29^{(2)}$; $34^{(2)}$; $41^{(2)}$; $51^{(1)}$; $59^{(2)}$; $76^{(1)}$; $79^{(1)}$; $87^{(2)}$; $110^{(1)}$; $120^{(1)}$; $124^{(1)}$; $193^{(1)}$; $195^{(1)}$
G_{12}	4	$12^{(4)}$; $16^{(4)}$; $34^{(2)}$; $40^{(4)}$; $48^{(2)}$; $68^{(1)}$; $70^{(2)}$
G_{13}	4	$12^{(4)}$; $16^{(4)}$; $34^{(2)}$; $40^{(4)}$; $48^{(2)}$; $68^{(2)}$; $72^{(1)}$
G_{14}	5	$21^{(2)}$; $27^{(2)}$; $31^{(2)}$; $38^{(2)}$; $48^{(1)}$; $56^{(2)}$; $59^{(1)}$; $68^{(2)}$; $76^{(1)}$; $89^{(1)}$; $101^{(2)}$; $116^{(1)}$; $162^{(2)}$
G_{15}	5	$21^{(2)}$; $27^{(2)}$; $31^{(2)}$; $38^{(2)}$; $48^{(1)}$; $56^{(2)}$; $59^{(1)}$; $68^{(2)}$; $76^{(1)}$; $89^{(1)}$; $99^{(1)}$; $103^{(1)}$; $116^{(1)}$; $160^{(1)}$; $164^{(1)}$

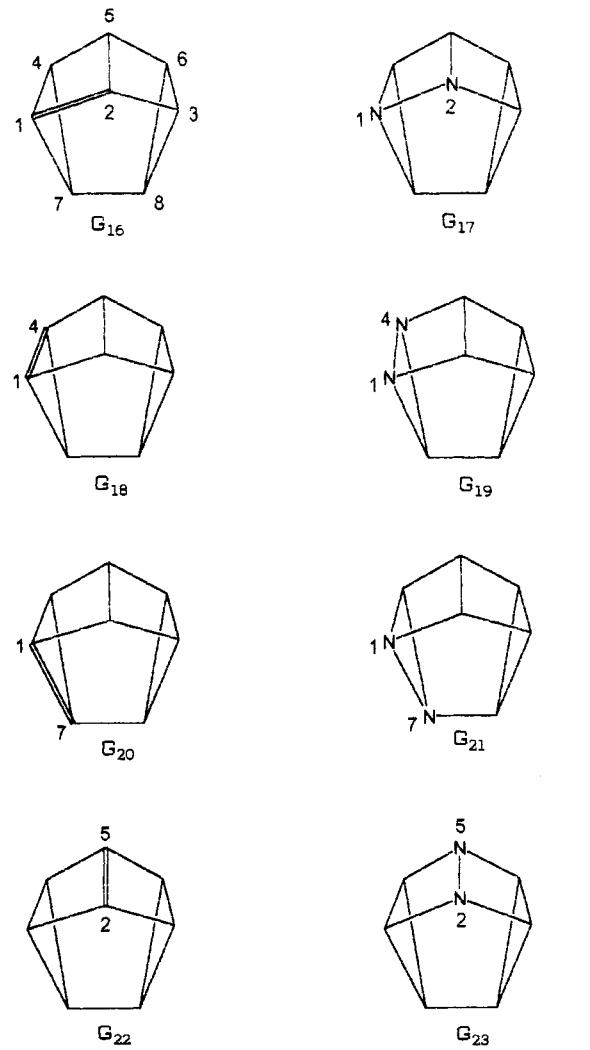


Figure 7. Congeners of cuneane.

Table 3. Topological Equivalence Perception in G_{16}–G_{23} by Three Walk Related Invariants (at the Minimal Discriminating Elongation, e)

G_{16}	$w_i^{(4)}$	{1}; {2}; {3}; {4}; {5,7}; {6,8}
	$\chi w_i^{(4)}$	{1}; {2}; {3}; {4}; {5}; {6,8}; {7}
	$c(\mathbf{LW}^{(4)})_i$	{1}; {2}; {3}; {4}; {5}; {6}; {7}; {8}
G_{17}	$w_i^{(4)}$	{1,5,7}; {2}; {3,4}; {6,8}
	$\chi w_i^{(4)}$	{1}; {2}; {3}; {4}; {5}; {6,8}; {7}
	$c(\mathbf{LW}^{(4)})_i$	{1}; {2}; {3}; {4}; {5}; {6}; {7}; {8}
G_{18}	$w_i^{(1)}$	{1,4}; {2,3,5,6,7,8}
	$\chi w_i^{(1)}$	{1,4}; {2,5}; {3,6,8}; {7}
	$c(\mathbf{LW}^{(1)})_i$	{1,4}; {2,5}; {3,6}; {7}; {8}
G_{19}	$w_i^{(1)}$	{1,4}; {2,3,5,6,7,8}
	$\chi w_i^{(1)}$	{1,4}; {2,5}; {3,6,8}; {7}
	$c(\mathbf{LW}^{(1)})_i$	{1,4}; {2,5}; {3,6}; {7}; {8}
G_{20}	$w_i^{(3)}$	{1,7}; {2,8}; {3}; {4}; {5}; {6}
	$\chi w_i^{(3)}$	{1,7}; {2}; {3}; {4}; {5}; {6}; {8}
	$c(\mathbf{LW}^{(3)})_i$	{1}; {2}; {3}; {4}; {5}; {6}; {7}; {8}
G_{21}	$w_i^{(3)}$	{1,7}; {2,6,8}; {3,4}; {5}
	$\chi w_i^{(3)}$	{1,7}; {2}; {3}; {4}; {5}; {6}; {8}
	$c(\mathbf{LW}^{(3)})_i$	{1}; {2}; {3}; {4}; {5}; {6}; {7}; {8}
G_{22}	$w_i^{(1)}$	{1,3,4,6,7,8}; {2,5}
	$\chi w_i^{(1)}$	{1,3,4,6}; {2,5}; {7,8}
	$c(\mathbf{LW}^{(1)})_i$	{1,3,4,6}; {2,5}; {7,8}
G_{23}	$w_i^{(1)}$	{1,3,4,6,7,8}; {2,5}
	$\chi w_i^{(1)}$	{1,3,4,6}; {2,5}; {7,8}
	$c(\mathbf{LW}^{(1)})_i$	{1,3,4,6}; {2,5}; {7,8}

Table 4. Values for the Walk Related Parameters Shown in Table 3 for the Graphs G_{20} and G_{21} (at the Minimal Separating Elongation, e)

	G_{20}			G_{21}		
i	$w_i^{(3)}$	$\chi w_i^{(3)}$	$c(\mathbf{LW}^{(3)})_i$	$w_i^{(3)}$	$\chi w_i^{(3)}$	$c(\mathbf{LW}^{(3)})_i$
1	49	0.068 281	0.147 196	29.0624	0.105 465	0.150 123
2	32	0.090 654	0.232 206	26.9136	0.108 655	0.242 606
3	29	0.101 776	0.139 565	27.6116	0.110 049	0.149 317
4	37	0.076 225	0.146 178	27.6116	0.106 358	0.149 820
5	30	0.096 987	0.230 623	28.3277	0.108 189	0.242 416
6	28	0.104 147	0.135 491	26.9136	0.110 055	0.148 662
7	49	0.068 281	0.237 407	29.0624	0.105 465	0.242 961
8	32	0.092 168	0.231 968	26.9136	0.109 595	0.231 968

multiple bond and heteroatom (or different chemical nature of a vertex in molecular graphs) thus offering "true" atomic walk counts, $tw_i^{(e)}$, as exemplified in Figure 2.

The centricity of vertices is reflected in the $w_i^{(e)}$ values: the closer they are located to the central region of a graph, the larger are walk degrees, $w_i^{(e)}$. However, the endospectral vertices (like no. 4 and no. 8 in G_5) show an identical walk degree sequence but different centricity (see Figure 5). Our layer matrices $\mathbf{LW}^{(e)}$ (by means of the centric index $c(\mathbf{LW}^{(e)})_i$) give more information (though sometimes not in agreement with the Bonchev's 1D–3D criteria[8]) about the vertex centricity.

As could be seen from Figure 4 and Tables 1 and 3, our descriptors $c(\mathbf{LW}^{(e)})_i/C(\mathbf{LW}^{(e)})$ discriminate vertices/graphs within the tested structures at lower e values than the other walk related parameters or sequences do. Thus they can be used in QSPR/QSAR studies.

It was well established by Cvetkovic's and Gutman[27] that $GW^{(e)}$ is directly related to the largest eigenvalue, λ_1, of the characteristic polynomial

$$GW^{(e)} \simeq n(\lambda_1)^e \qquad (34)$$

as well as to the molecular branching. Very recently, Rucker and Rucker[28] demonstrated a similar relation between $GW^{(e)}$ and the coefficients of the first eigenvector.

In regular graphs, the approximation 34 becomes equality. In such graphs the walk degrees $w_i^{(e)}$ fail in discriminating the topologically nonequivalent vertices. In this respect several procedures[17,22,29] take into account (explicitly or not) information on rings including the checked vertices. We solved the problem by means of the iterative line derivatives,[30] L_n, and local invariants constructed on layer matrices (i.e., **LDS**, see Figure 5) as proposed within the MOLORD algorithm.[21]

COMPUTER PROGRAM

A computer program, TOP-W (in TURBO PASCAL, and operable on IBM PC compatible computers), is developed. It enables the calculation of walk degrees, $w_i^{(e)}$ according to the proposed algorithm, and also the $\mathbf{LW}^{(e)}$ matrices. The output file gives the walk degree sequences, the global number of walks in graph, $GW^{(e)}$, the matrix $\mathbf{LW}^{(e)}$ (at request), and the proposed index $c(\mathbf{LW}^{(e)})_i/C(\mathbf{LW}^{(e)})$, either as a full listing or global results, after an interactive dialogue.

CONCLUSIONS

The summation algorithm for counting walk degrees, $w_i^{(e)}$, when iteration occurs on the connectivity matrix, takes into account the multiple bonds in molecular graphs. The chemical nature of vertices (i.e., heteroatoms) is considered by a valance electronegativity parameter, thus counting weighted walk degrees, $tw_i^{(e)}$. These parameters were used to construct the layer matrices $\mathbf{LW}^{(e)}$ which by means the centric index $C(\mathbf{LW}^{(e)})$ enabled the discrimination of vertices/graphs at an e value lower than other walk related parameters or sequences. It was used in topological symmetry perception and is promising in QSPR/QSAR studies. Formulas for counting walk degrees in some particular graphs were given.

ACKNOWLEDGMENT

We are very indebted to Dr. C. Rücker (Institut fur Organische Chemie und Biochemie, Universitat Freiburg, Germany) and Dr. O. Ivanciuc (Polytechnic University, Bucharest, Romania) for helpful comments and reprints.

Appendix: Valence EC electronegativities[16] used in the calculus of $tw_i^{(e)}$.

>C<	1.0000	–H	0.9175
>C=	1.0747	–N<	1.2234
–C≡	1.1476	=N–	1.3147
=C=	1.1581	≡N	1.5288
–CH<	0.9716	–NH–	1.1021
=CH–	1.0441	=NH	1.2474
≡CH	1.2142	$-NH_2$	1.0644
$-CH_2-$	0.9622	$-NHCH_3$	1.0379
$=CH_2$	1.0891	$-N(CH_3)_2$	1.0292
$-CH_3$	0.9575	–NO	1.4063
$-CH_2Br$	1.0110	$-NO_2$	1.4861
$-CH_2Cl$	1.0305	–O–	1.4634
$-CH_2F$	1.0674	=O	1.6564
$-CH_2I$	0.9744	–OH	1.2325
$-CH_2OH$	1.0228	$-OCH_3$	1.1248
$-CH_2SH$	0.9804	–S–	1.1064
$-CHBr_2$	1.0672	=S	1.2523
–CHCl2	1.1089	$-SCH_3$	1.0073
$-CHF_2$	1.1897	–P<	0.8988
$-CHI_2$	0.9914	=P–	0.9658
$-CBr_3$	1.1266	–PH–	0.9124
$-CCl_3$	1.1932	$-PH_2$	0.9170
$-CF_3$	1.3260	–F	1.6514
$-CI_3$	1.0088	–Cl	1.3717
>C=O	1.2397	–Br	1.2447
–CH=O	1.1596	–I	1.0262
–COOH	1.2220	–C≡N	1.2377

REFERENCES AND NOTES

(1) Part 16 of the series Molecular Topology: Diudea, M. V. Layer Matrices in Molecular Graphs. *J. Chem. Inf. Comput. Sci.*, preceding paper in this issue.

(2) Morgan, H. The Generation of a Unique Machine Description for Chemical Structures. A Technique Developed at Chemical Abstracts Service. *J. Che. Doc.* **1965**, *5*, 107–113.

(3) Skorobogatov, V. A.; Dobrynin, A. A. Metric Analysis of Graphs. *MATCH* **1988**, *23*, 105–151.

(4) Diudea, M. V.; Parv, B. A New Centric Connectivity Index (CCI). *MATCH* **1988**, *23*, 65–87.

(5) Diudea, M. V.; Minailiuc, O. M.; Balaban, A. T. Regressive Vertex Degrees (New Graph Invariants) and Derived Topological Indices. *J. Comput. Chem.* **1991**, *12*, 527–535.

(6) Balaban, A. T.; Diudea, M. V. Real Number Vertex Invariants: Regressive Distance Sums and Related Topological Indices. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 421–428.

(7) Randic', M.; DeAlba, L.; Guo, X.; Kleiner, A. F.; Krishnapriyan, H.; Naylor, L.; Oxley, T. Matrices Associated with Graphs. *J. Chem. Inf. Comput. Sci.*, submitted for publication.

(8) Bonchev, D.; Balaban, A. T.; Randic', M. The Graph Center Concept for Polycyclic Graphs. *Int. J. Quantum Chem.* **1981**, *19*, 61–82.

(9) Trinajstic', N. *Chemical Graph Theory*; CRC Press, Inc.: Boca Raton, FL, 1983; Vol. 1, p 8.

(10) Randic', M.; Woodworth, W. L.; Graovac, A. Unusual Random Walks. *Int. J. Quantum Chem.* **1983**, *24*, 435–452.

(11) Rücker, G.; Rücker, C. Counts of All Walks as Atomic and Molecular Descriptors. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 683–695.

(12) Ivanciuc, O.; Balaban, A. T. Nonisomorphic Graphs with Identical Atomic Counts of Self-Returning Walks: Isocodal Graphs. *J. Math. Chem.* **1992**, *11*, 155–167.

(13) Figueras, J. Morgan Revisited. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 717–718.

(14) Harary, F. *Graph Theory*; Addison-Wesley: Reading, MA, 1969; p 151.

(15) Razinger, M. Extended Connectivity in Chemical Graphs. *Theor. Chim. Acta* **1982**, *61*, 581–586.

(16) Diudea, M. V.; Kacso, I. E.; Topan, M. I. A QSPR/QSAR Study by Using New Valence Group Carbon-Related Electronegativities. *Rev. Roumaine Chim.*, in press.

(17) Mekenyan, O.; Balaban, A. T.; Bonchev, D. Unique Description of Chemical Structure Based on Hierarchically Ordered Extended Connectivities. Condensed Benzenoid Hydrocarbons and Their ^{1}H NMR Chemical Shifts. *J. Magn. Reson.* **1985**, *63*, 1–13.

(18) Bonchev, D.; Mekenyan, O.; Balaban, A. T. Iterative Procedure for the Generalized Graph Center in Polycyclic Graphs. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 91–97.

(19) Rücker, G.; Rücker, C. Computer Perception of Constitutional (Topological) Symmetry: TOPSYM, a Fast Algorithm for Partitioning Atoms and Pairwise Relations among Atoms into Equivalence Classes. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 187–191.

(20) 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.

(21) Diudea, M. V.; Horvath, D.; Bonchev, D. MOLORD Algorithm and Real Number Subgraph Invariants. *Croat. Chem. Acta*, in press.

(22) Balaban, A. T.; Mekenyan, O.; Bonchev, D. Unique Description of Chemical Structures Based on Hierarchically Ordered Extended Connectivities (HOC Procedures). Algorithms for Finding Graph Orbits and Canonical Numbering of Atoms. *J. Comput. Chem.* **1985**, *6*, 538–551.

(23) Barysz, M.; Trinajstic', N. A Novel Approach to the Characterization of Chemical Structures. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1984**, *18*, 661–673.

(24) Ivanciuc, O.; Balaban, T. S.; Balaban, A. T. Chemical Graphs with Degenerate Topological Indices Based on Information on Distances. *J. Math. Chem.* **1993**, *14*, 21–33.

(25) Razinger, M. Discrimination and Ordering of Chemical Structures by the Number of Walks. *Theor. Chim. Acta* **1986**, *70*, 365–378.

(26) Diudea, M. V.; Bal, L. Recursive Relationships for Computing Y-Indices in Some Particular Graphs. *Studia Univ. Babes-Bolyai* **1990**, *35*, 17–28.

(27) Cvetkovic', D. M.; Gutman, I. Note on Branching. *Croat. Chem. Acta* **1977**, *49*, 115–121.

(28) Rücker, G.; Rücker, C. Mathematical Relation between Extended Connectivity and Eigenvector Coefficients. *J. Math. Chem.*, in press.

(29) Bonchev, D.; Kier, L. B.; Mekenyan, O. Self-Returning Walks and Fractional Electronic Charges of Atoms in Molecules. *Int. J. Quantum Chem.* **1993**, *46*, 635–649.

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