Detour-Cluj Matrix and Derived Invariants

Mircea V. Diudea,*,† Bazil Parv,‡ and Ivan Gutman§

Department of Chemistry and Department of Mathematics, “Babes-Bolyai” University, 3400 Cluj, Romania, and Faculty of Science, University of Kragujevac, 34000 Kragujevac, Yugoslavia

Received June 9, 1997⊗

Detour-variant, $\mathbf{CJ\Delta}_u$, of the recently proposed Cluj matrix, $\mathbf{CJD}_u$, is defined and exemplified. Cluj indices built up on the two variants of the Cluj matrix are evaluated and compared, for selected sets of graphs. A new graph-theoretical local property is defined, namely, the “*internal ending* of all longest paths joining a vertex i, with $\deg_i > 1$, with all the remaining vertices in the graph”, as the property of the vertex i (called an *internal endpoint*) to have all its entries in $\mathbf{CJ\Delta}_u$ equal to 1. Classes of graphs possessing a minimal $CJ\Delta_p$ value (*i.e.*, in which all vertices are internal endpoints) have been identified. Correlating tests of the invariants calculated on the two variants of Cluj matrices are also performed.

INTRODUCTION

Two vertices i and j in a connected graph G are always joined by a path $p(i,j)$ (*i.e.*, a continuous sequence of vertices/edges, with the property that all are distinct and any two subsequent vertices/edges are adjacent).[1,2] The length of the path is given by the cardinality $|p(i,j)|$ of $p(i,j)$ viewed as a set of edges.

In acyclic graphs, the path having its endpoints as the vertices i and j is unique. In cycle-containing graphs, more than one path may exist, so that the symbol $p_k(i,j)$ will be used for specifying the kth path.

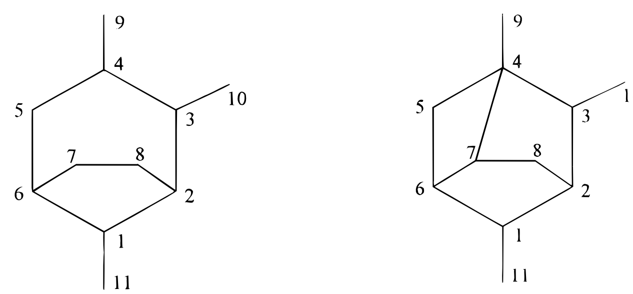
If the path is the shortest one among any possible paths joining i and j, its edge cardinality represents just the topological distance, $D_{i,j}$. If the path is the longest one, then the number of edges which separate i and j is called the detour (or maximum) distance, Δ_{ij}. The two quantities can be collected in square arrays, called the *distance*,[1] $\mathbf{D}_e$, and the *detour*,[3−5] $\mathbf{\Delta}_e$, matrices, respectively

$$[\mathbf{D}_e]_{ij} = \begin{cases} N_{e,p(i,j)}; |p(i,j)| = \min, & \text{if } i \neq j \\ 0, & \text{if } i = j \end{cases} \quad (1)$$

$$[\mathbf{\Delta}_e]_{ij} = \begin{cases} N_{e,p(i,j)}; |p(i,j)| = \max, & \text{if } i \neq j \\ 0, & \text{if } i = j \end{cases} \quad (2)$$

where $N_{e,p(i,j)} = |p(i,j)|$ is the number of edges which separate the vertices i and j on the shortest/longest path $p(i,j)$. The subscript e in the symbol of the two matrices means that they are edge-defined (*i.e.*, their entries count edges on the path $p(i,j)$). Matrices $\mathbf{D}_e$ and $\mathbf{\Delta}_e$ for the graph G_1 (the skeleton graph of ecgonine, a tropanic alkaloid, in which the nitrogen atom in position 1 is not distinguished from the carbon atoms) are illustrated in Chart 1.

Another pair of matrices can be constructed when paths of length $1 \leq |p| \leq |p(i,j)|$ that are counted in the path $p(i,j)$ are considered



$\mathbf{G_1}$ $\mathbf{G_2}$

$$[\mathbf{D}_p]_{ij} = \begin{cases} N_{p,p(i,j)}; |p(i,j)| = \min, & \text{if } i \neq j \\ 0, & \text{if } i = j \end{cases} \quad (3)$$

$$[\mathbf{\Delta}_p]_{ij} = \begin{cases} N_{p,p(i,j)}; |p(i,j)| = \max, & \text{if } i \neq j \\ 0, & \text{if } i = j \end{cases} \quad (4)$$

They are path-defined matrices and the number of paths, $N_{p,p(i,j)}$, is obtained from the entries $[\mathbf{M}_e]_{ij}$, $\mathbf{M}_e = \mathbf{D}_e, \mathbf{\Delta}_e$, by[6,7]

$$N_{p,p(i,j)} = \binom{[\mathbf{M}_e]_{ij} + 1}{2} = \{([\mathbf{M}_e]_{ij})^2 + [\mathbf{M}_e]_{ij}\}/2 \quad (5)$$

Matrices $\mathbf{D}_p$ and $\mathbf{\Delta}_p$ for the graph G_1 are illustrated in Chart 1.

On the above specified matrices, the structure descriptors, I, (*i.e.*, topological indices) can be calculated as half-sums of their entries

$$I_{e/p} = (1/2)\sum_i \sum_j [\mathbf{M}_{e/p}]_{ij} \quad (6)$$

where the edge-defined index, I_e, is the Wiener,[8] W, and the detour,[4,5,9,10] w, while the path-defined one, I_p, is the hyper-Wiener,[11] WW, and the hyperdetour,[9] ww, respectively. Values of these indices for the graph G_1 are presented in Chart 1.

In simple cycles, the above indices can be calculated by simple analytical relations[9,12]

† Department of Chemistry, “Babes-Bolyai” University. E-mail address: diudea@chem.ubbcluj.ro.
‡ Department of Mathematics, “Babes-Bolyai” University.
§ University of Kragujevac.
⊗ Abstract published in *Advance ACS Abstracts,* October 1, 1997.

S0095-2338(97)00041-3 CCC: $14.00

Chart 1. Distance and Detour Matrices of the Graph G_1

$\mathbf{D_e}$

	1	2	3	4	5	6	7	8	9	10	11
1	0	1	2	3	2	1	2	2	4	3	1
2	1	0	1	2	3	2	2	1	3	2	2
3	2	1	0	1	2	3	3	2	2	1	3
4	3	2	1	0	1	2	3	3	1	2	4
5	2	3	2	1	0	1	2	3	2	3	3
6	1	2	3	2	1	0	1	2	3	4	2
7	2	2	3	3	2	1	0	1	4	4	3
8	2	1	2	3	3	2	1	0	4	3	3
9	4	3	2	1	2	3	4	4	0	3	5
10	3	2	1	2	3	4	4	3	3	0	4
11	1	2	3	4	3	2	3	3	5	4	0
	21	19	20	22	22	21	25	24	31	29	30

$$W = 1/2 \sum_i \sum_j [\mathbf{D_e}]_{ij} = 132$$

Δ_e

	1	2	3	4	5	6	7	8	9	10
1	0	5	7	6	7	5	7	7	7	8
2	5	0	6	5	4	4	5	6	6	7
3	7	6	0	6	5	4	7	6	7	1
4	6	5	6	0	6	5	6	6	1	7
5	7	4	5	6	0	6	6	7	7	6
6	5	4	4	5	6	0	6	5	6	5
7	7	5	7	6	6	6	0	6	7	8
8	7	6	6	6	7	5	6	0	7	7
9	7	6	7	1	7	6	7	7	0	8
10	8	7	1	7	6	5	8	7	8	0
11	1	6	8	7	8	6	8	8	8	9
	60	54	57	55	62	52	66	65	64	66

$$w = 1/2 \sum_i \sum_j [\Delta_e]_{ij} = 335$$

$\mathbf{D_p}$

	1	2	3	4	5	6	7	8	9	10	11
1	0	1	3	6	3	1	3	3	10	6	1
2	1	0	1	3	6	3	3	1	6	3	3
3	3	1	0	1	3	6	6	3	3	1	6
4	6	3	1	0	1	3	6	6	1	3	10
5	3	6	3	1	0	1	3	6	3	6	6
6	1	3	6	3	1	0	1	3	6	10	3
7	3	3	6	6	3	1	0	1	10	10	6
8	3	1	3	6	6	3	1	0	10	6	6
9	10	6	3	1	3	6	10	10	0	6	15
10	6	3	1	3	6	10	10	6	6	0	10
11	1	3	6	10	6	3	6	6	15	10	0
	37	30	33	40	38	37	49	45	70	61	66

$$\mathrm{WW} = 1/2 \sum_i \sum_j [\mathbf{D_p}]_{ij} = 253$$

Δ_p

	1	2	3	4	5	6	7	8	9	10
1	0	15	28	21	28	15	28	28	28	36
2	15	0	21	15	10	10	15	21	21	28
3	28	21	0	21	15	10	28	21	28	1
4	21	15	21	0	21	15	21	21	1	28
5	28	10	15	21	0	21	21	28	28	21
6	15	10	10	15	21	0	21	15	21	15
7	28	15	28	21	21	21	0	21	28	36
8	28	21	21	21	28	15	21	0	28	28
9	28	21	28	1	28	21	28	28	0	36
10	36	28	1	28	21	15	36	28	36	0
11	1	21	36	28	36	21	36	36	36	45
	228	177	209	192	229	164	255	247	255	274

$$\mathrm{ww} = 1/2 \sum_i \sum_j [\Delta_p]_{ij} = 1263$$

$$W = N(N^2 - z)/8 \tag{7}$$

$$\mathrm{WW} = (N - z)(N - z + 1)(N - z + 2)(N + 3z)/48 \tag{8}$$

$$w = N(3N^2 - 4N + z)/8 \tag{9}$$

$$\mathrm{ww} = N(7N^3 - 3N^2 - 10N + 3z(N + 1))/48 \tag{10}$$

$$z = N \bmod 2$$

The Wiener index is the first molecular-graph-based structure descriptor introduced for the purpose of QSPR studies.[8] It is one of the most studied indices, from the point of view of applications and theory. The other distance-based indices have been introduced more recently, but they have already promoted interesting research.[9,10] Some recent reviews are available on this matter.[13−15]

DEFINITION OF THE CLUJ MATRIX

The unsymmetric Cluj matrix, $\mathbf{CJ}_u$, has been recently proposed by Diudea.[16−18] It is defined[19] by using the minimum-path concept (*i.e.*, the $D_{i,j}$), so that in the following the $\mathbf{CJD}_u$ symbol is used. A variant of this matrix, $\mathbf{CJ\Delta}_u$, which involves the maximum-path concept (*i.e.*, the $\mathbf{\Delta}_{i,j}$), is added here. The nondiagonal entries, $[\mathbf{M}_u]_{ij}$, $\mathbf{M}_u = \mathbf{CJD}_u$, $\mathbf{CJ\Delta}_u$, in the two Cluj matrices are defined as

$$[\mathbf{M}_u]_{ij} = N_{i,p_k(i,j)} = \max|V_{i,p_k(i,j)}| \tag{11}$$

where $|V_{i,p_k(i,j)}|$ is the number of elements of the set $V_{i,p_k(i,j)}$, where the maximum is taken over all paths $p_k(i,j)$ and where

$$V_{i,p_k(i,j)} = \{v | v \in V(G); D_{iv} < D_{jv}; p_h(i,v) \cap p_k(i,j) = \{i\}; |p_k(i,j)| = \min\} \tag{12}$$

or

$$V_{i,p_k(i,j)} = \{v | v \in V(G); D_{iv} < D_{jv}; p_h(i,v) \cap p_k(i,j) = \{i\}; |p_k(i,j)| = \max\} \tag{13}$$

$$k = 1, 2, ...; h = 1, 2, ...$$

The set $V_{i,p_k(i,j)}$ (eqs 12 and 13) consists of the vertices lying *closer* to the vertex i and *external* with respect to the path $p_k(i,j)$ (condition $p_h(i,v) \cap p_k(i,j) = \{i\}$). Since, in cycle-containing structures, various shortest paths, $p_k(i,j)$, could supply various sets $V_{i,p_k(i,j)}$, by definition, the (ij)-entries in the Cluj matrices are taken as max $|V_{i,p_k(i,j)}|$. The diagonal entries are zero. The above definitions (eqs 11−13) hold for any connected graph.

Chart 2. (a) Paths (2,5) and (b) Distance-Cluj, $\mathbf{CJD}_u$, and Detour-Cluj, $\mathbf{CJ\Delta}_u$, Matrices for the graphs G_1 and G_2

(a) Paths (2, 5) for G_1: [2, 8, 7, 6, 5], [2, 1, 6, 5], [2, 3, 4, 5]

Paths (2, 5) for G_2:

[2, 1, 6, 7, 4, 5], [2, 3, 4, 7, 6, 5], [2, 8, 7, 4, 5], [2, 8, 7, 6, 5][2, 1, 6, 5], [2, 3, 4, 5]

(b)

$\mathbf{CJD}_u(G_1)$

	1	2	3	4	5	6	7	8	9	10	11
1	0	4	4	5	4	5	5	4	5	6	10
2	6	0	6	5	4	4	7	7	6	6	7
3	4	5	0	6	4	4	4	5	7	10	6
4	4	3	4	0	5	4	4	3	10	5	4
5	3	3	3	5	0	5	5	3	6	5	6
6	5	3	4	5	6	0	6	6	6	5	6
7	2	3	2	4	3	2	0	3	5	5	5
8	2	2	3	4	2	4	4	0	5	5	6
9	1	1	1	1	1	1	1	1	0	1	1
10	1	1	1	1	1	1	1	1	1	0	1
11	1	1	1	1	1	1	1	1	1	1	0

$$\mathrm{CJD_p} = \sum\nolimits_p [\mathbf{CJD}_u]_{ij}[\mathbf{CJD}_u]_{ji} = 616$$
$$\mathrm{CJD_e} = \sum\nolimits_e [\mathbf{CJD}_u]_{ij}[\mathbf{CJD}_u]_{ji} = 226$$

$\mathbf{CJ\Delta}_u(G_1)$

	1	2	3	4	5	6	7	8	9	10	11
1	0	2	2	2	2	2	2	2	2	2	10
2	2	0	3	3	5	2	3	3	5	3	3
3	2	2	0	2	2	4	2	2	2	10	2
4	4	2	2	0	2	2	4	3	10	2	4
5	1	3	1	1	0	1	1	1	1	3	1
6	2	2	4	3	3	0	3	3	4	4	3
7	1	1	1	1	2	1	0	1	1	1	1
8	1	1	2	1	1	1	1	0	1	2	1
9	1	1	1	1	1	1	1	1	0	1	1
10	1	1	1	1	1	1	1	1	1	0	1
11	1	1	1	1	1	1	1	1	1	1	0

$$\mathrm{CJ\Delta_p} = \sum\nolimits_p [\mathbf{CJ\Delta}_u]_{ij}[\mathbf{CJ\Delta}_u]_{ji} = 200$$
$$\mathrm{CJ\Delta_e} = \sum\nolimits_e [\mathbf{CJ\Delta}_u]_{ij}[\mathbf{CJ\Delta}_u]_{ji} = 60$$

$\mathbf{CJD}_u(G_2)$

	1	2	3	4	5	6	7	8	9	10	11
1	0	4	3	3	4	5	3	4	4	6	10
2	6	0	5	4	4	4	5	5	5	6	7
3	4	5	0	5	4	4	3	5	6	10	7
4	5	4	5	0	7	4	5	5	10	6	6
5	3	3	2	4	0	5	1	3	4	5	6
6	5	3	4	3	6	0	4	4	6	5	6
7	4	5	3	5	3	6	0	5	6	6	7
8	4	2	3	4	2	4	2	0	5	5	6
9	1	1	1	1	1	1	1	1	0	1	1
10	1	1	1	1	1	1	1	1	1	0	1
11	1	1	1	1	1	1	1	1	1	1	0

$$\mathrm{CJD_p} = \sum\nolimits_p [\mathbf{CJD}_u]_{ij}[\mathbf{CJD}_u]_{ji} = 616$$
$$\mathrm{CJD_e} = \sum\nolimits_e [\mathbf{CJD}_u]_{ij}[\mathbf{CJD}_u]_{ji} = 256$$

$\mathbf{CJ\Delta}_u(G_2)$

	1	2	3	4	5	6	7	8	9	10	11
1	0	2	2	2	2	2	2	2	2	2	10
2	3	0	3	3	4	3	3	3	5	3	3
3	2	2	0	2	2	2	2	2	2	10	2
4	4	2	2	0	4	4	4	3	10	2	4
5	1	1	1	1	0	1	1	1	1	1	1
6	1	1	3	3	3	0	3	2	4	3	1
7	1	1	1	1	2	1	0	1	1	1	1
8	1	1	1	1	1	1	1	0	1	1	1
9	1	1	1	1	1	1	1	1	0	1	1
10	1	1	1	1	1	1	1	1	1	0	1
11	1	1	1	1	1	1	1	1	1	1	0

$$\mathrm{CJ\Delta_p} = \sum\nolimits_p [\mathbf{CJ\Delta}_u]_{ij}[\mathbf{CJ\Delta}_u]_{ji} = 177$$
$$\mathrm{CJ\Delta_e} = \sum\nolimits_e [\mathbf{CJ\Delta}_u]_{ij}[\mathbf{CJ\Delta}_u]_{ji} = 66$$

Cluj matrices are square arrays of dimension $N \times N$ and are, in general, unsymmetric with respect to the main diagonal. They are illustrated for the graphs G_1 and G_2 in Chart 2. Note the changing in the path counting when a supplementary bond (4,7) is added in G_1 to give G_2. More details in defining the Cluj matrices can be found in ref 19.

In what follows we call a vertex *external* if it has just one first neighbor (*i.e.*, its degree is equal to 1) and *internal* if it has several first neighbors (*i.e.*, its degree is greater than 1).

As can be seen from Chart 2, some non-external vertices in G_2 (*e.g.*, numbers 5 and 8) happen to have all their entries in $\mathbf{CJ\Delta}_u$ equal to 1. The same entries show the external vertices, numbers 9−11 (which are at the same time endpoints of all paths which contain them). We call this property *the internal ending of all paths* (particularly, longest paths) joining the vertex i and the remaining vertices in G. Further, the vertex i, like numbers 5 and 8, is said to be an *internal endpoint*. Distance-Cluj matrix, $\mathbf{CJD}_u$ has all entries equal to 1, for an internal vertex, only accidentally, in cases when the shortest path coincides with the longest path (see below). Such a property did not appear, to our knowledge, in any other topological index.

The two Cluj matrices, $\mathbf{M}_u$, allow the construction of the corresponding symmetric matrices, $\mathbf{M}_p$ (defined on paths) and $\mathbf{M}_e$ (defined on edges) by

$$\mathbf{M}_p = \mathbf{M}_u \cdot (\mathbf{M}_u)^T \tag{14}$$

$$\mathbf{M}_e = \mathbf{M}_p \cdot \mathbf{A} \tag{15}$$

where $\mathbf{A}$ is the adjacency matrix (having the nondiagonal entries equal to 1 if the vertices i and j are adjacent and zero otherwise). The symbol · indicates the Hadamard (pairwise) matrix product[20] (*i.e.*, $[\mathbf{M}_a \cdot \mathbf{M}_b]_{ij} = [\mathbf{M}_a]_{ij}[\mathbf{M}_b]_{ij}$).

In acyclic structures, the two variants of Cluj matrices coincide, as a consequence of the uniqueness of the path $p(i,j)$. The symmetric matrices, edge-defined and path-defined ones, in both variants, are identical to the Wiener matrices, $\mathbf{W}_e$ (edge-defined) and $\mathbf{W}_p$ (path-defined), respectively (see below). The relations between $\mathbf{CJD}_u$, $\mathbf{W}_e$, and $\mathbf{D}_e$ were detailed elsewhere.[16,17]

Recall that the Wiener index, in acyclic structures, can be calculated[8] as

$$W = \sum_{e} N_{i,p(i,j)} N_{j,p(i,j)} \tag{16}$$

where $N_{i,p(i,j)}$ and $N_{j,p(i,j)}$ have the same meaning (in such graphs) as the quantity $N_{i,p_k(i,j)}$ in eq 11. The summation runs over all edges $e = (i,j)$, which means that $p(i,j)$ is here just an edge (*i.e.*, a path of length 1). The product $N_{i,p(i,j)}N_{j,p(i,j)}$ is the (i,j)-entry of the Wiener matrix,[21,22] $\mathbf{W}_e$, from which W can be calculated as

$$W = (1/2)\sum_{i}\sum_{j}[\mathbf{W}_e]_{ij} \tag{17}$$

Note that the vertices i and j must be adjacent, otherwise the nondiagonal entries in $\mathbf{W}_e$ are zero. When $p(i,j)$ represents a path of arbitrary length, then a relation similar to (16) will define the *hyper-Wiener* index,[11] WW

$$\mathrm{WW} = \sum_{p} N_{i,p(i,j)} N_{j,p(i,j)} \tag{18}$$

and subsequently, the product $N_{i,p(i,j)}N_{j,p(i,j)}$ is the (i,j)-entry of the Wiener matrix,[21,22] $\mathbf{W}_p$, from which WW can be calculated as the half-sum of its entries.

Relations 16 and 18 say that the index (particularly, the edge/path contributions to the global index) results by multiplying the number of vertices on the left- and on the right-hand sides of the edge or path having the endpoints i and j. Such local products count all the paths which include the edge/path (i,j) as a subgraph and their endpoints are *external* with respect to (i,j).

In cycle-containing graphs, the Wiener matrices are not defined. Wiener indices are calculated by means of the distance-type matrices (see the previous section). Cluj matrices try to fill the lack of the Wiener matrices, in the same manner (*i.e.*, imposing the externality for the paths they count with respect to the path (i,j)). In such graphs, the two versions of Cluj matrices differ from each other.

Several indices can be calculated on the Cluj matrices,[17,18] either as a half-sum of entries in the corresponding symmetric matrices (a relation of the type (17)) or directly from the unsymmetric matrices, by

$$I_{e/p} = \sum_{e/p} [\mathbf{M}_u]_{ij}[\mathbf{M}_u]_{ji} \tag{19}$$

When defined on edges, I_e is an *index* (*e.g.*, $\mathrm{CJD_e}$); when defined on paths, I_p is a *hyperindex* (*e.g.*, $\mathrm{CJ\Delta_p}$). The values of some Cluj indices, calculated for a set of 24 planar benzenoids[23,24] of Figures 1 and 2, are reported in Table 1. For comparison, values of the Wiener- and detour-type indices are also given.

The ability of these indices to explain two physicochemical properties, boiling points, BP, and chromatographic retention index, I_{CHR}, were evaluated from data sets of 22 and respectively 16 compounds (Figures 1 and 2). The regression equations, of the form $Y = a + \sum_i b_i X_i$, and the accompanying statistics are given in Table 2.

The best single variable regression against BP was found with ln $\mathrm{CJD_p}$: $r = 0.9887$; $s = 20.150$; $F = 866.3$. Here r is the coefficient of the correlation, s is the standard error, and F is the Fischer ratio. The equation is quite satisfactory,

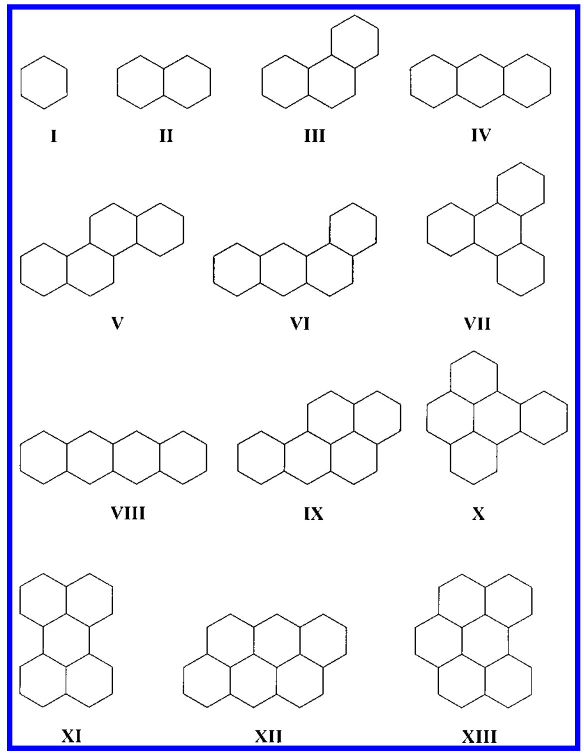


Figure 1. Benzenoid hydrocarbons studied: I, benzene; II, naphtalene; III, phenanthrene; IV, anthracene; V, chrysene; VI, benzanthracene; VII, triphenylene; VIII, tetracene; XIX, benzo(*a*)-pyrene; X, benzo(*e*)pyrene; XI, perylene; XII, anthanthrene; XIII, benzoperylene.

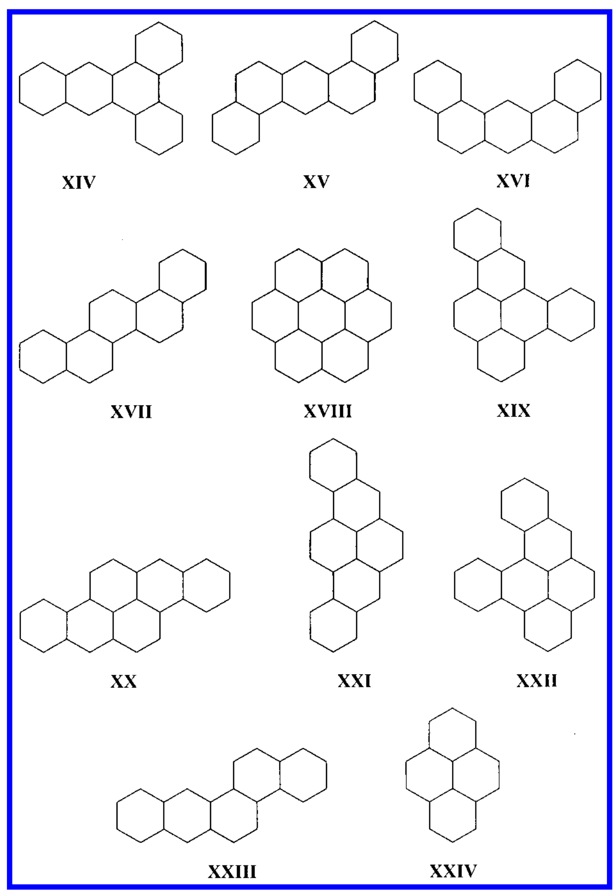


Figure 2. Benzenoid hydrocarbons studied: XIV, dibenzo(*a*,*c*)-anthracene; XV, dibenzo(*a*,*h*)anthracene; XVI, dibenzo(*a*,*i*)anthracene; XVII, picene; XVIII, coronene; XIX, dibenzo(*a*,*k*)pyrene; XX, dibenzo(*a*,*h*)pyrene; XXI, dibenzo(*a*,*g*)pyrene; XXII, dibenzo-(*a*,*d*)pyrene; XXIII, benzo(*b*)chrysene; XXIV, pyrene.

Table 1. Boiling Points, BP, Chromatographic Retention Indices, I_{CHR}, and Topological Indices for the Benzenoids of Figures 1 and 2

graph	BP	I_{CHR}	W	WW	w	ww	CJD_p	CJD_e	$CJ\Delta_p$	$CJ\Delta_e$
I	80.1		27	42	63	168	90	54	24	6
II	218	1784	109	215	345	1 521	591	243	129	19
II	338	2874	271	636	1019	6 314	2 150	632	326	32
IV	340	2800	279	680	1015	6 272	2 240	656	390	32
V	431	4198	545	1513	2275	18 286	5 917	1301	638	85
VI	425	4169	557	1585	2265	18 135	6 031	1352	671	71
VII	429	4017	513	1305	2229	17 679	5 307	1269	753	45
VIII	440		569	1661	2257	18 017	6 225	1381	709	61
IX	496	5488	680	1868	3166	28 199	9 253	1887	758	56
X	493	4650	652	1684	3190	28 589	8 876	1852	724	62
XI	497	4739	654	1697	3146	27 915	9 081	1858	781	50
XII	547	5439	839	2308	4341	43 139	14 229	2613	865	52
XIII	542	5262	815	2148	4433	44 945	13 890	2571	909	67
XIV	535	5099	911	2710	4215	41 168	12 194	2321	1264	74
XV	536	5325	975	3190	4261	41 935	13 654	2377	1300	82
XVI	531		973	3171	4265	42 011	13 505	2396	1267	82
XVII	519		963	3118	4299	42 630	13 522	2330	1126	82
XVIII	590	5821	1002	2697	5790	63 855	20 547	3438	1125	66
XIX	592		1082	3163	5640	60 914	17 915	3095	1229	76
XX	596		1142	3619	5644	60 938	19 147	3174	1173	69
XXI	594		1142	3619	5642	60 893	19 147	3174	1193	69
XXII	595		1066	3051	5640	60 926	17 295	3023	1267	76
XXIII		5488	975	3190	4281	42 287	13 584	2385	1161	86
XXIV		3301	362	845	1564	11 066	3 867	1008	427	43

Table 2. Regression Equations, $Y = a + \sum_i b_i X_i$, and the Corresponding Statistics

no.	Y	X_1 (b_1)	X_2 (b_2)	X_3 (b_3)	a	r	s	F
1	BP	ln CJD_p (99.432)			−408.280	0.9887	20.150	866.3
2		ln $CJ\Delta_p$ (80.449)	CJD_e (0.063)		−176.141	0.9955	13.041	1048.6
3		ln CJD_p (122.796)	$1/CJ\Delta_e$ (1044.669)		−641.263	0.9963	11.836	1274.9
4		ln CJD_p (74.310)	w (0.021)		−257.042	0.9982	8.149	2700.1
5		ln CJD_p (78.617)	w (0.021)	$CJ\Delta_e$ (−0.318)	−276.613	0.9986	7.420	2172.9
6	I_{CHR}	ln CJD_p (1233.053)			−6489.431	0.9877	186.334	558.9
7		ln CJD_p (1062.066)	$CJ\Delta_p$ (0.519)		−5381.288	0.9901	173.701	323.2
8		ln CJD_p (859.713)	w (0.240)		−3916.317	0.9937	138.337	513.3
9		ln CJD_p (784.905)	W (1.568)		−3539.182	0.9949	124.291	637.4
10		ln CJD_p (697.965)	W (2.948)	$CJ\Delta_p$ (−0.933)	−2935.399	0.9970	99.397	667.2

as reflected by the high correlation coefficient and the standard error value, which is less than 5% of the mean experimental value.

In two-variable regression, the correlation was improved by adding the detour and detour-type Cluj indices (Table 2, entries 2−4). The best regression equation was obtained with ln CJD_p and w ($r = 0.9982$, $s = 8.149$, and $F = 2700.1$; for cross-validation, $r = 0.9978$ and $s = 8.943$). The standard error of estimate significantly diminishes when a third variable is added (entry 5): ln CJD_p and w and $CJ\Delta_e$ ($r = 0.9986$, $s = 7.420$, and $F = 2172.9$; for cross validation, $r = 0.9977$ and $s = 9.053$).

In prediction, the three-variable regression equation does not surpass the two-variable one (see above and Table 2, entries 4 and 5). Nevertheless, the standard error of prediction is less than 2% of the mean experimental value. The predicting ability of the regression equations was established by a "leave one out" *cross-validation* procedure.

The above discussed equations are superior to the best regression equations reported by Randić[23] by using ${}^kD$ descriptors.

The chromatographic retention index, I_{CHR}, was the second property tested. As can be seen from Table 2, satisfactory correlations have been obtained in two- and three-variable regressions and again ln CJD_p was the best single-variable descriptor ($r = 0.9877$; $s = 186.334$; $F = 558.9$). The best second variable to be added in regression equations was the Wiener index, W (Table 2, entry 9): $r = 0.9949$; $s = 124.291$; $F = 637.4$. The standard error of estimate is now less than 3% of the mean experimental value. A third variable is justified by the significant drop in the standard error of estimate but also by the high values of the coefficient of the correlation and Fischer ratio: ln CJD_p and W and $CJ\Delta_p$; $r = 0.9970$, $s = 99.397$, and $F = 667.2$; for cross-validation, $r = 0.9928$ and $s = 142.731$. The standard error of prediction is less than 3.3%. The results are comparable with those reported by Randić.[24]

In simple cycles, the distance-Cluj indices can be calculated by means of the analytical relations[12]

$$CJD_e = N(N - z)^2/4 \tag{20}$$

$$CJD_p = N(72z^2N - 9y^2N - 16z^3 - 3y^2 - 12yz^2 + 12yzN + 12y^2z - 3N^2 - 2y^3 + 7N^3 + 8N + 8y - 108zN + 12z^2 + 18yN - 12yz + 16z)/96 \tag{21}$$

$$y = N \bmod 4; \quad z = N \bmod 2$$

The detour-Cluj indices can be calculated by similar relations, but do not depend on z

$$CJ\Delta_e = N \tag{22}$$

$$CJ\Delta_p = (k + 1)N(4k^2 + 3yk + 2k + 3y)/6 \tag{23}$$

$$k = [(N - 1)/4]; \quad y = (N - 1) \bmod 4$$

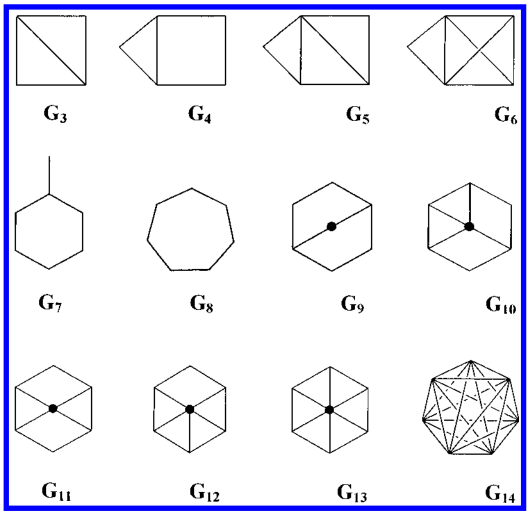


Figure 3. Selected simple graphs in searching for a minimal $CJ\Delta_p$ value.

The edge-defined index is just equal to N (here, the number of vertices, or edges, M, or the parameter girth, g). The path-defined index shows a similar mod 4 dependency, but the relation is simpler than that for calculating CJD_p.

GRAPHS WITH MINIMAL $CJ\Delta_P$ VALUE

As in the case of the Wiener index, Cluj hyperindices, $CJ\Delta_p$ and CJD_p, attain their minimal values for the complete graphs, K_N. Both in distance and detour versions, the Cluj matrices have all their entries (out of the diagonal) equal to 1. Thus, in complete graphs, the value of the Cluj indices (cf. eq 19) equals the number of edges, $N(N-1)/2$.

Simple cycles of girth 3−5 also have all entries equal to 1, but only in the detour version, $\mathbf{CJ\Delta_u}$. Since only the triangle is a complete graph, it appears that there exist graphs, other than the complete graphs, for which $CJ\Delta_p$ conforms to the formula for the complete graphs

$$CJ\Delta_p = N(N-1)/2 \tag{24}$$

First, we investigated structures having small cycles, condensed in various ways. The first row in Figure 3 shows such structures with minimal $CJ\Delta_p$.

Observe that only triangles and squares can be fused to give a periphery with $g < 6$. The second row collects structures with $N = 7$ and decreasing values of $CJ\Delta_p$. In the third row, all structures have the minimal value, given by eq 24. In fact, in structures with a minimal $CJ\Delta_p$ value, *all of the vertices are internal endpoints*. Note that the star-triangulane, G_{13} (see also refs 25 and 26), also shows values of detour indices, ww and w, identical with those given by the complete graph, G_{14}, with the same number of vertices, $N = 7$ (Table 3). Formulas for calculating the detour-type indices in such graphs are shown in Table 4. Also note that, for all graphs showing a minimal $CJ\Delta_p$ value, the edge-defined index, $CJ\Delta_e$, equals their number of edges, M (see Table 3). It is true also for all simple cycles.

Second, we found that strips (like prism, cube, etc.) with odd g and Möbius strips (irrespective of g) show the minimal value; cf. eq 24. For such strips, all of the detour-type indices are identical. Conversely, the strips with even g do not have a minimal value for $CJ\Delta_p$. In a pair of strip and Möbius strip, the last one shows the minimal value of $CJ\Delta_p$ and the maximal value for the other detour-type indices. Analytical

Table 3. Indices of Some Small Cyclic Graphs

graph	$CJ\Delta_p$	$CJ\Delta_e$	CJD_P	CJD_e	ww	w	WW	W
3	6	5	10	9	33	17	8	7
4	10	6	36	24	88	37	18	14
5	10	7	26	18	92	38	16	13
6	10	8	20	16	96	39	14	12
7	48	14	142	78	71	42	251	90
8	42	7	154	63	322	105	70	42
9	40	14	135	62	312	103	53	36
10	26	11	137	82	411	121	49	34
11	21	10	126	64	429	124	46	33
12	21	11	118	62	435	125	44	32
13	21	12	99	48	441	126	39	30
14	21	21	21	21	441	126	21	21

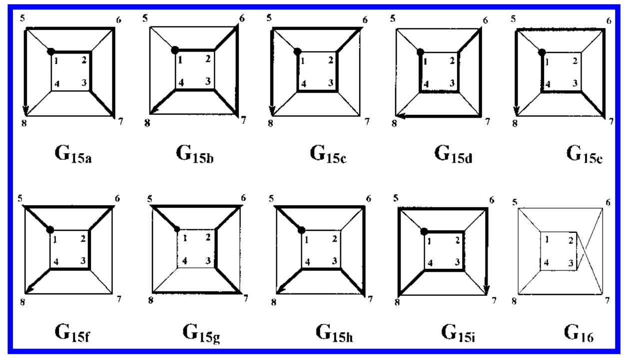


Figure 4. Cube (various longest paths), G_{15a-i}, and mobius cube, G_{16}.

formulas for calculating the detour-type indices (as a function of g and as a function of $CJ\Delta_p$) are given in Table 4.

Figure 4 illustrates the calculation of (i,j)-entries in the detour-Cluj matrix of the cube (G_{15a}−G_{15h}). The longest paths joining vertices 1 and 8 in the graphs G_{15a}, G_{15b}, G_{15d}, and G_{15g} lead to $|V_{1,p_k(1,8)}| = 1$ since (cf. eq 13) vertices equidistant to i and j are not counted. The longest paths in the graphs G_{15e} and G_{15h} give $|V_{1,p_k(1,8)}| = 2$. According to eq 13, the entry $[\mathbf{CJ\Delta_u}]_{1,8} = \max|V_{1,p_k(1,8)}| = 2$.

In the case of the entry (1,7), (cf. G_{15i}) the longest path obeys the relation

$$|p_{\Delta k}(i,j)| = N - 1 \tag{25}$$

which is the condition for a Hamiltonian path[1,2,7] (*i.e.*, a path which visits all of the vertices in a graph). In such cases the (i,j)-entry takes the minimal value (*e.g.*, $[\mathbf{CJ\Delta_u}]_{1,7} = 1$). Condition 25 is sufficient for the minimal value of $[\mathbf{CJ\Delta_u}]_{ij}$ or minimal value of $CJ\Delta_p$, but it is, however, not necessary: eq 13 says that vertices equidistant to i and j are not counted so that the longest path can be less than $N - 1$.

The Möbius cube, whose Schlegel projection is given in G_{16} satisfies condition 25 for all pairs (i,j), thus showing a minimal $CJ\Delta_p$ value. This is true for all Möbius strips which are full Hamiltonian graphs (*i.e.*, the longest paths between all vertex pairs (i,j) are Hamiltonian ones). This is also true for the star-triangulanes. Values of the above discussed topological indices in strips and Möbius strips are shown in Table 5.

Third, dipyramids of any girth show minimal $CJ\Delta_p$ value, as the Möbius strips. Between the two classes of graphs a simple relation can be established: Dipyramids(g) = Möbius strips ($g/2 + 1$; g-even; $g \geq 4$). Note that dipyramids are also full Hamiltonian graphs. Analytical relations for calculating the detour-type indices in dipyramids are given

Table 4. Formulas for Detour-Type Indices in Particular Classes of Graphs (g = Girth)

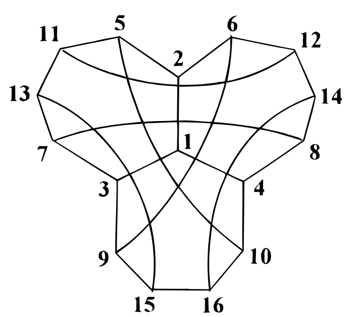
graphs	N	M	$CJ\Delta_p$	$CJ\Delta_e$	ww	w
complete graphs	g	$g(g-1)/2$	$g(g-1)/2$	$g(g-1)/2$	$(CJ\Delta_p)^2$	$(g-1)CJ\Delta_p$
triangulanes	$g+1$	$2g$	$g(g+1)/2$	$2g$	$(CJ\Delta_p)^2$	$g(CJ\Delta_p)$
strips (g-even)	$2g$	$3g$	$g(5g-4)$	$3g$	$g(2g-1)(2g^2-2g+1)$	$g(4g^2-5g+2)$
strips (g-odd) and Möbius strips	$2g$	$3g$	$g(2g-1)$	$3g$	$(CJ\Delta_p)^2$	$(2g-1)CJ\Delta_p$
dipyramids	$g+2$	$3g$	$(g+2)(g+1)/2$	$3g$	$(CJ\Delta_p)^2$	$(g+1)CJ\Delta_p$

Table 5. Indices of Strips and Möbius (mö) Strips of Girth g from 3 to 10

g	$CJ\Delta_p$	$CJ\Delta_e$	CJD_p	CJD_e	ww	w	WW	W
3	15	9	75	51	225	75	27	21
3 (mö)	15	9	75	51	225	75	27	21
4	64	12	276	192	700	184	72	48
4 (mö)	28	12	212	108	784	196	60	44
5	45	15	625	285	2025	405	135	85
5 (mö)	45	15	561	251	2025	405	123	81
6	156	18	1452	648	4026	696	258	144
6 (mö)	66	18	1226	416	4356	726	222	134
7	91	21	2625	847	8281	1183	413	217
7 (mö)	91	21	2455	733	8281	1183	377	207
8	288	24	4824	1536	13560	1744	672	320
8 (mö)	120	24	4450	1062	14400	1800	596	302
9	153	27	7659	1881	23409	2601	981	441
9 (mö)	153	27	7421	1621	23409	2601	905	423
10	460	30	12260	3000	34390	3520	1450	600
10 (mö)	190	30	11580	2170	36100	3610	1314	572

in Table 4. Values of the discussed indices in dipyramids with g between 3 and 10 are listed in Table 6.

A supplementary example of graphs having a minimal $CJ\Delta_p$ value is offered by the cage depicted in the graph G_{17}. It was built up[27] in analogy to the Petersen and Heawood graphs.[1,2,7] It is a cubic graph (*i.e.*, a regular graph having



G_{17}

all valencies equal to 3) which shows degenerate rearrangements, for some selected permutations. Note that G_{17} is a full Hamiltonian graph, whereas the Petersen and Heawood graphs are not.

CONCLUSIONS

The detour-Cluj matrix, $\mathbf{CJ\Delta}_u$, counts vertices lying closer to the focused vertex i and external to the longest path which joins the vertices i and j in a graph. It is a proper extension of the Wiener matrix in cycle-containing structures. Indices derived on it show relatively small values and interesting regularities. In a correlating test, the boiling points and the chromatographic retention indices of a set of planar benzenoid molecules was quite well estimated and predicted by the Cluj indices.

Table 6. Indices of Dipyramids of Girth g from 3 to 10

g	$CJ\Delta_p$	$CJ\Delta_e$	CJD_p	CJD_e	ww	w	WW	W
3	10	9	16	15	100	40	12	11
4	15	12	51	48	225	75	21	18
5	21	15	101	80	441	126	33	27
6	28	18	172	120	784	196	48	38
7	36	21	260	168	1296	288	66	51
8	45	24	365	224	2025	405	87	66
9	55	27	487	288	3025	550	111	83
10	66	30	626	360	4356	726	138	102

A new graph-theoretical property was found: *the internal ending of all longest paths* joining the vertex i (with $\deg_i > 1$) and the remaining vertices in G. It results in all off-diagonal entries in $\mathbf{CJ\Delta}_u$ equal to 1. Then the vertex i is called an *internal endpoint*. Classes of graphs, having all vertices internal points and thus supplying a minimal value of $CJ\Delta_p$ are identified.

The detour-variant of the Cluj matrix and derived indices appears to be promising tools in the investigation of the chemical structure. In a future paper,[28] the chemical nature of vertices as well as the 3D distances between them will be considered.

ACKNOWLEDGMENT

This work is supported in part by the Romanian Ministry of Education under Grant 339 T74 and by the Romanian Academy under Grant 2627.

REFERENCES AND NOTES

(1) Harary, F. *Graph Theory*; Addison-Wesley: Reading, MA, 1969.
(2) Trinajstić, N. *Chemical Graph Theory*; CRC Press: Boca Raton, FL, 1983.
(3) Buckley, F.; Harary, F. *Distance in Graphs*; Addison-Wesley: Reading, MA, 1971; p.203.
(4) Amić, D.; Trinajstić, N. On the Detour Matrix. *Croat. Chem. Acta* **1995**, *68*, 53–62.
(5) Ivanciuc, O.; Balaban, A. T. Design of topological indices. Part. 8. Path matrices and derived molecular graph invariants. *Commun. Math. Comput. Chem.(MATCH)* **1994**, *30*, 141–152.
(6) Diudea, M. V. Walk Numbers $^{e}W_M$: Wiener-Type Numbers of Higher Rank. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 535–540.
(7) Klein, D. J.; Lukovits, I.; Gutman, I. On the Definition of the Hyper-Wiener Index for Cycle-Containing Structures. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 50–52.
(8) Wiener, H. Structural Determination of Paraffin Boiling points. *J. Am. Chem. Soc.* **1947**, *69*, 17–20.
(9) Lukovits, I. The Detour Index. *Croat. Chem. Acta* **1996**, *69*, 873–882.
(10) Lukovits, I.; Razinger, M. On Calculation of the Detour Index. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 283–286.
(11) Randić, M. Novel Molecular Descriptor for Structure-Property Studies. *Chem. Phys. Lett.* **1993**, *211*, 478–483.
(12) Diudea, M. V. Indices of Reciprocal Property or Harary Indices. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 292–299.
(13) Gutman, I.; Yeh, Y. N.; Lee, S. L.; Luo, Y. L. Some Recent Results in the Theory of the Wiener Number. *Indian J. Chem.* **1993**, *32A*, 651–661.
(14) Nikolić, S.; Trinajstić, N.; Mihalić, Z. The Wiener Index: Development and Applications. *Croat. Chem. Acta* **1995**, *68*, 105–129.
(15) Diudea, M. V.; Gutman, I. Wiener-Type Topological Indices *Croat. Chem. Acta*, submitted for publication.

(16) Diudea, M. V. Wiener and Hyper-Wiener Numbers in a Single Matrix. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 833−836.
(17) Diudea, M. V. Cluj Matrix, CJ_u: Source of Various Graph Descriptors. *Commun. Math. Comput. Chem*. (*MATCH*) **1997**, *35*, 169−183.
(18) Diudea, M. V. Cluj Matrix Invariants. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 300−305.
(19) Gutman, I.; Diudea, M. V. Defining Cluj Matrices and Cluj Matrix Invariants. *Commun. Math. Comput. Chem.* (*MATCH)*, submitted for publication.
(20) Horn, R. A.; Johnson, C. R. *Matrix Analysis*; Cambridge University Press: Cambridge, U.K., 1985.
(21) Randić, M.; Guo, X.; Oxley, T.; Krishnapriyan, H. Wiener Matrix: Source of Novel Graph Invariants. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 700−716.
(22) Randić, M.; Guo, X.; Oxley, T.; Krishnapriyan, H.; Naylor, L. Wiener Matrix Invariants. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 361−367.
(23) Randić, M. Quantitative Structure-Property Relationship. Boiling Points of Planar Benzenoids. *New J. Chem.* **1996**, *20*, 1001−1009.
(24) Randić, M. Molecular Profiles. Novel Geometry-Dependent Molecular Descriptors. *New J. Chem.* **1995**, *19*, 781−791.
(25) Zefirov, N. S.; Kozhushkov, S. I.; Kuznstsova, T. S.; Kokoreva, O. V.; Lukin, O.V.; Ugrak, B. I.; Tratch, S. S. Triangulanes: Stereoisomerism and General Method of Synthesis. *J. Am. Chem. Soc*. **1990**, *112*, 7702−7707.
(26) John, P. E. On the Calculation of Wiener-Index for Selected Δ-dimensional Lattice Structures. *Commun. Math. Comput. Chem.* (*MATCH)* **1995**, *32*, 207−219.
(27) Diudea, M. V. Multiple Line Derivatives. Lecture at MATH/CHEM/COMP’94, Dubrovnik, Croatia.
(28) Diudea, M. V.; B. Parv, B. Fragmental Cluj Indices. Manuscript in preparation.

CI970041N