

The Wiener Polynomial Derivatives and Other Topological Indices in Chemical Research

Elena V. Konstantinova^{a,*} and Mircea V. Diudea^{b,*}

^a*Institute of Mathematics, Siberian Branch of the Russian Academy of Science, Novosibirsk 630090, Russia*

^b*Department of Chemistry, Babes-Bolyai University, 3400 Cluj, Romania*

Received February 18, 1999; revised July 1, 1999; accepted July 19, 1999

Wiener polynomial derivatives and some other information and topological indices are investigated with respect to their discriminat- ing power and property correlating ability.

Key words: information indices, topological indices, Wiener polynomial

INTRODUCTION

A single number, representing a chemical structure, in graph-theoretical terms, is called a topological descriptor. It must be a structural invariant, *i.e.*, it does not depend on the labeling or the pictorial representation of a graph. If such a descriptor correlates with a molecular property, it can be denominated as a topological index (*TI*). Despite the considerable loss of information by the »projection« of a structure to a single number, topological indices have found broad applications in the correlation (estimation and prediction) of several molecular properties.^{1–6} Another representation that has gained particular attention, both from the theoretical point of view and applications, is by polynomials.

* Author to whom correspondence should be addressed.

POLYNOMIALS

Various graph-theoretical polynomials are applied to many different areas of chemistry. The spectra of the characteristic polynomial of graphs were studied to obtain the molecular orbitals.⁷ Analysis of the characteristic polynomial was important for understanding the global electronic structure of unsaturated hydrocarbon molecules.^{8,9} A reference polynomial was proposed to discuss the aromatic character of polycyclic hydrocarbons.¹⁰ Independently, the same polynomial was proposed for the same purpose under the name of acyclic polynomial.¹¹ Further, Farrell extended the idea of these polynomials and introduced the matching polynomial.¹² The rotational polynomial¹³ was used for analysis of the entropy of chain hydrocarbon molecules. The sextet polynomial¹⁴ was important to analyze what kinds of aromatic sextets can stabilize the whole molecule. Hosoya¹⁵ introduced the distance polynomial and conjectured that it could be used for the unique characterization of a molecular graph. Later, he defined¹⁶ the Wiener polynomial $H_G(x)$ of G as

$$H_G(x) = \sum_{k=1}^d d_k x^k, \quad (1)$$

where d is the diameter of graph G and d_k is the number of vertex pairs at a distance k from each other. In 1947, Harry Wiener proposed two structural parameters, W and p , which give a good correlation with the thermodynamic properties of saturated hydrocarbon molecules.¹⁷ The Wiener number W and the polarity number p are defined as the half sum of the off-diagonal elements of the distance matrix and as d_3 , respectively. Using the following expressions for the first and third derivatives of $H_G(x)$

$$H'_G(x) = \sum_{k=1}^d k d_k x^{k-1} \quad (2)$$

$$H'''_G(x) = \sum_{k=3}^d k(k-1)(k-2) d_k x^{k-3} \quad (3)$$

we can obtain the relations for W and p :

$$W = \sum_{k=1}^d k d_k = H'_G(1) \quad (4)$$

$$p = d_3 = H'''_G(1)/6 \quad (5)$$

or, generally,

$$d_k = H_G^{(k)}(1)/k! . \quad (6)$$

We study the all derivatives of $H_G(x)$ as the Wiener-like graph topological indices. The discrimination power of indices and the structure-activity relationships are investigated below.

The same Wiener-like topological indices termed Extended Wiener indices were considered in Ref 18.

INFORMATION DISTANCE INDEX AND CLUJ-TYPE INDICES – POWERFUL DISCRIMINATORS

We will consider the information index based on distances within a graph. The following well-known principle¹⁹ is generally used for the construction of information indices. Let X be a set consisting of n elements. Let us assume that by some equivalence criterium the elements are divided into N equivalence classes X_i , where $n = \sum_{i=1}^N n_i$ and n_i is the number of elements in subset X_i . Then, $p_i = n_i / n$ is the probability of a single element belonging to the i -th subset, and estimating quantitatively the information that corresponds to one element of the set. One can use the distribution entropy for the set elements defined by the following formula of Shannon:²⁰

$$H = - \sum_{i=1}^p p_i \log p_i . \quad (7)$$

Information indices of molecular graphs were constructed for various matrices of graphs and also for some topological indices.¹ The Information Distance Index was introduced in Ref. 21 and it is defined on the basis of the distance matrix, $\mathbf{D} = \|d_{i,j}\|$, $i, j = 1 \dots p$, where $d_{i,j}$ is the topological distance between vertices i and j (*i.e.*, the number of edges on the shortest path joining vertices i and j) in graph G . Let define the information distance index, IDI , of vertex i as follows

$$IDI(i) = - \sum_{j=1}^p \frac{d_{i,j}}{d(i)} \log \frac{d_{i,j}}{d(i)} \quad (8)$$

where $d(i) = \sum_{j=1}^p d_{i,j}$ and $p_{i,j} = \frac{d_{i,j}}{d(i)}$ is the probability of an accidentally chosen vertex being at a distance $d_{i,j}$ from vertex i . Then, the Information Distance Index of graph vertices takes the form

$$IDI = IDI(G) = \sum_{i=1}^p IDI(i) . \quad (9)$$

This index has a high discriminating ability for molecular graphs. The discrimination power of the *IDI* index was investigated on graphs of unbranched hexagonal systems.²² The results are shown below.

Two unsymmetric Cluj matrices, \mathbf{CJD}_u (distance-Cluj), and \mathbf{CJA}_u (detour-Cluj), have been recently proposed by Diudea,²³⁻²⁵ using the geodesic concept (*i.e.*, the shortest path joining vertices i and j , evaluated as topological distance, $d_{i,j}$) and the elongation concept (*i.e.*, the longest path, evaluated as detour-distance, $\Delta_{i,j}$), respectively.

The non-diagonal entries, $[\mathbf{M}_u]_{i,j}$, $\mathbf{M}_u = \mathbf{CJD}_u$; \mathbf{CJA}_u , in the two Cluj matrices are defined as follows

$$[\mathbf{M}_u]_{i,j} = N_{i,(i,j)_k} = \max |V_{i,(i,j)_k}| \quad (10)$$

where $|V_{i,(i,j)_k}|$ is the cardinality of the set $V_{i,(i,j)_k}$, where the maximum is taken over all paths $(i,j)_k$ between i and j , and where

$$V_{i,(i,j)_k} = \{v | v \in V(G); d_{i,v} < d_{j,v}; (i,v)_h \cap (i,j)_k = \{i\}; (i,j)_k - \text{is a geodesic}\} \quad (11)$$

or

$$V_{i,(i,j)_k} = \{v | v \in V(G); d_{i,v} < d_{j,v}; (i,v)_h \cap (i,j)_k = \{i\}; (i,j)_k - \text{is an elongation}\} \quad (12)$$

$k = 1, 2, \dots; h = 1, 2, \dots$

The set $V_{i,(i,j)_k}$, Eqs. (11) and (12), consists of the vertices lying *closer* to vertex i , and *external* with respect to the path $(i,j)_k$ (condition $(i,v)_h \cap (i,j)_k = \{i\}$). Since in cycle-containing structures, more than one geodesic, $(i,j)_k$, could supply various sets $V_{i,(i,j)_k}$, by definition, the (i,j) - entries in the Cluj matrices are taken as $\max |V_{i,(i,j)_k}|$. The diagonal entries are zero. For paths $(i,v)_h$, no restriction is imposed. The above definitions, Eqs. (10) – (12), hold for any connected graph.

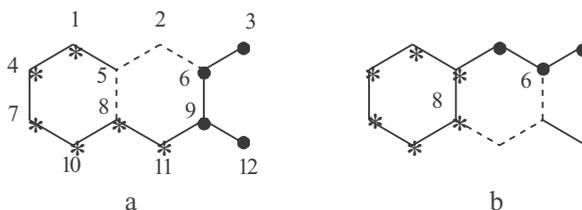
Cluj matrices are square arrays of dimension $N \times N$ and are, in general, unsymmetric with respect to the main diagonal. They are illustrated on the graph of 2,3-dimethyldecaline, in Figures 1 and 2.

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

$$\mathbf{M}_p = \mathbf{M}_u \bullet (\mathbf{M}_u)^T \quad (13)$$

$$\mathbf{M}_e = \mathbf{M}_p \bullet \mathbf{A} \quad (14)$$

where \mathbf{A} is the adjacency matrix. The symbol \bullet indicates the Hadamard (pairwise) matrix product²⁶ (*i.e.*, $[\mathbf{M}_a \bullet \mathbf{M}_b]_{i,j} = [\mathbf{M}_a]_{i,j} [\mathbf{M}_b]_{i,j}$).



Cluj-Distance matrix, CJD_u ; path (6,8):

(a) (6, 8) [6, 2, 5, 8] { 3, 6, 9, 12 }; entry-(6,8) = $\max |V_{6,(6,8)k}| = 4$.

(b) (6, 8) [6, 9, 11, 8] { 2, 3, 6 }

(a) (8, 6) [8, 5, 2, 6] { 1, 4, 7, 8, 10, 11 }; entry-(8,6) = $\max |V_{8,(6,8)k}| = 6$.

(b) (8, 6) [8, 11, 9, 6] { 1, 4, 5, 7, 8, 10 }

Cluj-Distance Matrix, CJD_u

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

$CJD_p = 1024$

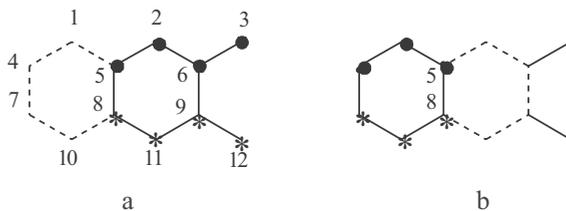
$CJD_e = 378$

Figure 1. Illustration of the Cluj-Distance matrix, CJD_u .

In acyclic structures, the two variants of Cluj matrices coincide, as a consequence of the uniqueness of the path (i,j) . The symmetric matrices, edge-defined and path-defined ones, in both variants, are identical to the Wiener matrices,^{27,28} W_e (edge-defined) and W_p (path-defined), respectively. Several indices can be calculated from Cluj matrices,²⁵ either as the half sum of entries in the corresponding symmetric matrices or directly from the unsymmetric matrices, by

$$I_e = \sum_{\text{all } (i,j) \in E(G)} [M_u]_{i,j} [M_u]_{j,i} \tag{15}$$

$$I_p = \sum_{\text{all } (i,j) \in P(G)} [M_u]_{i,j} [M_u]_{j,i} . \tag{16}$$



Cluj Detour matrix, $CJ\Delta_u$; path (5,8):

(a) $(5, 8) [5, 1, 4, 7, 10, 8] \{ 2, 3, 5, 6 \}$; entry-(5,8) = $\max |V_{5,(5,8)k}| = 4$.

(b) $(5, 8) [5, 2, 6, 9, 11, 8] \{ 1, 4, 5 \}$

(a) $(8, 5) [8, 10, 7, 4, 1, 5] \{ 8, 9, 11, 12 \}$; entry-(8,5) = $\max |V_{8,(5,8)k}| = 4$.

(b) $(8, 5) [8, 11, 9, 6, 2, 5] \{ 7, 8, 10 \}$

Cluj-Detour Matrix, $CJ\Delta_u$

0	1	2	1	1	2	1	2	1	2	1	1
1	0	1	2	1	1	1	3	1	1	3	3
1	1	0	1	1	1	1	1	1	1	1	1
1	3	3	0	1	3	1	2	2	1	2	2
1	1	2	1	0	1	2	4	2	2	2	2
5	2	11	5	2	0	3	4	2	3	2	2
1	2	2	1	2	2	0	1	3	1	3	3
2	2	2	2	4	2	1	0	1	1	1	2
3	2	2	3	4	2	5	2	0	5	2	11
2	1	1	1	2	1	1	1	2	0	1	2
1	3	3	1	3	1	2	1	1	1	0	1
1	1	1	1	1	1	1	1	1	1	1	0

$CJ\Delta_p = 247$

$CJ\Delta_e = 53$

Figure 2. Illustration of the Cluj-Detour matrix, $CJ\Delta_u$.

When defined on edges, I_e is an index (e.g., CJD_e); when defined on paths, I_p is a hyper-index (e.g. $CJ\Delta_p$).

Tratch *et al.*²⁹ have proposed an *extended distance matrix*, E , whose entries are the product of entries in the distance matrix D and a multiplier, $m_{i,j}$, which is the number of paths in a graph having the path (i,j) as a sub-path. In acyclic structures, $m_{i,j}$ equals the entries in the W_p matrix, so that E is further referred to as $D-W_p$ matrix

$$[D-W_p]_{i,j} = [D]_{i,j} m_{i,j} = [D]_{i,j} [W_p]_{i,j} = [D \bullet W_p]_{i,j}. \tag{17}$$

Thus, $D \cdot W_p$ matrix results as the Hadamard product (denoted by \bullet), $D \bullet W_p$. It is a square symmetric matrix of dimensions $N \times N$, having the diagonal entries zero. The half sum of its entries gives an expanded Wiener number.

In full analogy, Diudea^{23,24,30} has proposed distance-extended Cluj un-symmetric matrices

$$D_CJD_u = D \bullet CJD_u \quad (18)$$

$$\Delta_CJD\Delta_u = \Delta \bullet CJ\Delta_u \quad (19)$$

which offered new distance-extended indices

$$\begin{aligned} D^2_CJD_p &= \sum_p [D_CJD_u]_{i,j} [D_CJD_u]_{j,i} = \\ &= (1/2) \sum_i \sum_j [(D_CJD_u)(D_CJD_u)^T]_{i,j} \end{aligned} \quad (20)$$

$$\begin{aligned} \Delta^2_CJ\Delta_p &= \sum_p [\Delta_CJ\Delta_u]_{i,j} [\Delta_CJ\Delta_u]_{j,i} = \\ &= (1/2) \sum_i \sum_j [(\Delta_CJ\Delta_u)(\Delta_CJ\Delta_u)^T]_{i,j} \end{aligned} \quad (21)$$

where summation goes over all paths in G: $p = \text{all } (i,j) \in P(G)$. Note that index $D^2_CJ_p$ involves squared distances (see the superscript number).

In chemical graph theory, the distance matrix accounts for the »through bond« interactions of atoms in molecules. However, these interactions decrease as the distance between atoms increases. This is the reason why the »reciprocal distance« matrix, RD , was recently introduced.³¹⁻³³ The non-diagonal entries in this matrix are defined by

$$[RD]_{i,j} = 1/[D]_{i,j} \quad (22)$$

RD matrix allows the calculation of a Wiener index analogue, as the half sum of its entries

$$H_D = H_D(G) = (1/2) \sum_i \sum_j [RD]_{i,j} \quad (23)$$

The resulting number was called³³ the »Harary index«, in honor of Frank Harary. Diudea³⁴ has recently extended the use of »reciprocal (topological) property« matrices in defining novel Harary-type indices, H_M .

$$H_{M_{e/p}} = (1/2) \sum_i \sum_j 1/[M]_{i,j} = (1/2) \sum_i \sum_j [RM]_{i,j} \quad (24)$$

with $1/[M]_{i,j} = 0$ if $[M]_{i,j} = 0$. Subscript M is the identifier for a square matrix M , which collects some topological property, while the sub-subscript e/p specifies that the matrix (and the corresponding index) is defined either on edges or on paths. Within this work, only the hyper-indices H_{CJDp} and H_{CJAp} are considered.

DISCRIMINATING TESTS

Two basic characteristics of a topological index, I are: the correlating ability with a molecular property and the discrimination power (in the process of molecular structure classification).^{35,36} It is assumed that molecules with similar structures (or close values of a topological index, as a measure of structure similarity) have similar properties. However, the two characteristics are not necessarily correlated.

The discriminating sensitivity of I is a measure of its ability to distinguish among nonisomorphic graphs by distinct numerical values. The theoretical evaluation of index sensitivity, S , on a fixed set, M , of nonisomorphic graphs can be achieved by the formula

$$S = (N - N_I) / N, \quad (25)$$

where $N = |M|$ and N_I is the number of degeneracies of index I within set M .

The discrimination powers of all the Wiener polynomial derivatives for some classes of molecular structures, as well as of some Cluj-type indices, are examined below.

Planar Hexagonal Graphs

Graphs of this class are subgraphs of the regular hexagonal lattice and represent the cata-condensed and peri-condensed benzenoid hydrocarbons.³⁷ Let us consider all the derivatives of $H_G(x)$ as topological indices and calculate their discrimination powers evaluated for $x = 1$ on the class of planar hexagonal graphs with $h = 3 \dots 7$ rings in accordance with Eq. (25). The results are given in Table I.

Note that the actual number of graphs having higher derivatives is less than the real number of graphs in the considered set. For example, for the set M of planar hexagonal graphs G_1, G_2, G_3 presented in Figure 3, the Wiener polynomials are the following:

$$H_{G_1}(x) = 279x + 802x^2 + 2202x^3 + 5376x^4 + 10560x^5 + 14400x^6 + 10080x^7$$

$$H_{G_2}(x) = 271x + 7302x^2 + 1806x^3 + 3864x^4 + 6600x^5 + 7920x^6 + 5040x^7$$

$$H_{G_3}(x) = 210x + 468x^2 + 846x^3 + 1080x^4 + 720x^5$$

TABLE I

The discrimination power of all Wiener polynomial derivatives for the planar hexagonal graphs consisting of $h = 3 \dots 7$ rings

Number of derivative	$h = 3$ $N = 3$	$h = 4$ $N = 7$	$h = 5$ $N = 22$	$h = 6$ $N = 81$	$h = 7$ $N = 331$
1	1.000	1.000	0.909	0.556	0.338
2	1.000	1.000	0.909	0.778	0.677
3	1.000	1.000	0.909	0.802	0.752
4	1.000	1.000	0.909	0.802	0.749
5	1.000	1.000	0.909	0.802	0.755
6	1.000	1.000	0.909	0.802	0.746
7	1.000*	0.714	0.909	0.802	0.728
8		0.500*	0.474*	0.405*	0.403*
9		0.333*	0.222*	0.167*	0.161*
10			0.111*	0.100*	0.065*
11			0.167*	0.082*	0.038*
12				0.050*	0.027*
13				0.100*	0.032*
14					0.022*
15					0.050*

Within Table I, N is the number of graphs in the class. The values of discrimination power are marked by asterisks if the number of graphs having the corresponding derivative is less than the real number of graphs.

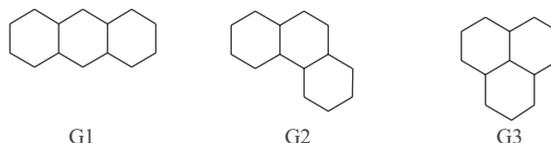


Figure 3. The planar benzenoid graphs with $h = 3$.

As one can see from these polynomials, it is possible to calculate the seven non-zero derivatives of the Wiener polynomials for graphs G_1, G_2 and only the five non-zero derivatives for graph G_3 . Thus, $N = 3$ (in Eq. (25)) to calculate the discrimination power for the first, second, third, fourth and fifth derivatives while $N = 2$ when the discrimination power for the sixth and seventh derivatives of Wiener polynomials are calculated.

It is seen from Table I that for $h \geq 6$ the discrimination powers of the second and third derivatives of the Wiener polynomial are greater than for the Wiener number.

Within this set of graphs, IDI yields $S = 1$.

Subgraphs of the Square Planar Lattice

Similar results (Table II) were recorded for the subgraphs of the regular quadrangular lattice with $h = 3 \dots 7$ rings.³⁸ A pair of graphs of this class are shown in Figure 4. One can see from Table II that for $h \geq 6$ the discrimination power of the second and third derivatives of the Wiener polynomial are greater than for the first derivative.



Figure 4. Subgraphs of the square planar lattice with $h = 3$.

TABLE II

The discrimination power of all Wiener polynomial derivatives for subgraphs of the regular quadrangular lattice consisting of $h = 3 \dots 7$ rings

Number of derivative	$h = 3$ $N = 2$	$h = 4$ $N = 4$	$h = 5$ $N = 12$	$h = 6$ $N = 35$	$h = 7$ $N = 107$
1	1.000	1.000	0.667	0.257	0.168
2	1.000	1.000	0.667	0.371	0.364
3	1.000	1.000	0.667	0.371	0.336
4	1.000	1.000	0.667	0.371	0.383
5		0.333*	0.273*	0.114	0.159
6			0.161*	0.071*	0.057*
7				0.091*	0.029*
8					0.050*

The notation is the same as in Table I.

IDI yields $S = 1$ for all graphs with $h = 3, 4$ and 5 ; for the graphs with $h = 6$, $S = 0.943$ while for those with $h = 7$, $S = 0.981$. For the whole set of graphs with $3 \leq h \leq 7$ ($N = 163$), $S = 0.975$.

Unbranched Hexagonal Graphs

Graphs of this class model molecular structures of unbranched cata-condensed benzenoid hydrocarbons.³⁷ The set of these graphs breaks up into two subsets of graphs: those which are embeddable and those which are not embeddable in the hexagonal lattice. To obtain these graphs, we used the algorithm of the fast generation of graphs of the unbranched hexagonal sys-

tems.³⁹ The discrimination powers of all $H_G(x)$ derivatives, evaluated for $x = 1$, on the class of all unbranched cata-condensed benzenoid graphs with three to ten rings, are given in Table III. Note that for graphs consisting of 9 rings the discrimination power of the second and third derivatives are, respectively, 12 and 15 times higher than for the first one.

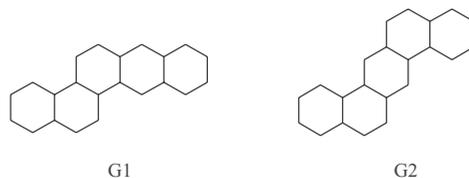


Figure 5. Graphs with the same Wiener polynomials.

TABLE III

The discrimination power of all Wiener polynomial derivatives for unbranched hexagonal graphs consisting of $h = 3 \dots 10$ rings

Number of derivative	$h = 3$ $N = 2$	$h = 4$ $N = 4$	$h = 5$ $N = 10$	$h = 6$ $N = 25$	$h = 7$ $N = 70$	$h = 8$ $N = 196$	$h = 9$ $N = 574$	$h = 10$ $N = 1681$
1	1.000	1.000	0.800	0.480	0.200	0.112	0.028	0.012
2	1.000	1.000	0.800	0.680	0.600	0.480	0.350	0.252
3	1.000	1.000	0.800	0.680	0.600	0.500	0.429	0.365
4	1.000	1.000	0.800	0.680	0.600	0.500	0.429	0.367
5	1.000	1.000	0.800	0.680	0.600	0.500	0.429	0.371
6	1.000	1.000	0.800	0.680	0.600	0.500	0.429	0.372
7	1.000	1.000	0.800	0.680	0.600	0.500	0.429	0.372
8		0.500	0.400	0.280	0.200	0.158	0.124	0.099
9		0.333*	0.400	0.240	0.200	0.148	0.113	0.090
10			0.111*	0.160	0.100	0.056	0.042	0.031
11			0.167*	0.125*	0.071	0.046	0.024	0.018
12				0.050*	0.043*	0.036	0.017	0.009
13				0.100*	0.047*	0.021*	0.016	0.007
14					0.022*	0.016*	0.010*	0.006
15					0.050*	0.019*	0.007*	0.005*
16						0.010*	0.006*	0.004*
17						0.028*	0.007*	0.002*
18							0.004*	0.002*
19							0.014*	0.003*
20								0.002*
21								0.007*

The notation is the same as in Table I.

For graphs consisting of 5 rings there are two graphs, G_1 and G_2 , (Figure 5) with the same Wiener polynomial:

$$H_{G_1}(x) = H_{G_2}(x) = 26x^1 + 76x^2 + 123x^3 + 136x^4 + 140x^5 + 132x^6 + 119x^7 + 96x^8 + 72x^9 + 40x^{10} + 11x^{11}.$$

Within this set, *IDI* yields $S = 1$ for all the graphs with $h = 3, \dots, 10$.

On the same set, the discriminating power of the Cluj-type indices was tested in comparison with the Wiener index, W (actually the first derivative Wiener polynomial – see Table III). The results are listed in Table IV.

TABLE IV

The discrimination power of some Cluj-type indices (see text) on the set of all cata-condensed benzenoid graphs with $h = 3 \dots 10$

Number of derivative	$h = 3$ $N = 2$	$h = 4$ $N = 4$	$h = 5$ $N = 10$	$h = 6$ $N = 25$	$h = 7$ $N = 70$	$h = 8$ $N = 196$	$h = 9$ $N = 574$	$h = 10$ $N = 1681$
W	1.000	1.000	0.800	0.480	0.200	0.112	0.028	0.012
CJD_p	1.000	1.000	1.000	0.920	0.971	0.929	0.955	0.902
CJA_p	1.000	1.000	1.000	0.920	0.886	0.694	0.606	0.137
$D^2_CJD_p$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\Delta^2_CJA_p$	1.000	1.000	1.000	1.000	1.000	1.000	0.997	0.985
H_{CJD_p}	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.988
H_{CJA_p}	1.000	1.000	1.000	1.000	1.000	1.000	0.979	0.963

One can see that the discriminating power of the Cluj-type indices is far higher than that of Wiener polynomials. Among the tested indices, special mention may be made of the distance-extended index, D^2_CJD , which, in the considered set, showed no degeneracy (*i.e.*, $S = 1$), neither within each individual subset nor the whole set of 2562 structures. The next most powerful index was H_{CJD_p} ($S = 0.988$). This index showed the maximum discriminating power within the set of all graphs with ten vertices and three to eight membered cycles (see below).

The Cycle-Containing Molecular Graphs

The set of these graphs consists of 437 structures with ten vertices and three to eight-membered cycles ($C = 3, \dots, 8$).²³ The results of the discriminating tests are given in Tables V and VI.

Note that in Ref. 23, H_{CJD_p} was reported to have no degeneracy (*cf.* one pair of structures within the subset $C = 3$ is degenerated), due to a printing

TABLE V

The discrimination power of all Wiener polynomial derivatives for cycle-containing graphs consisting of ten vertices and $C = 3...8$

Number of derivative	$C = 3$ $N = 168$	$C = 4$ $N = 140$	$C = 5$ $N = 70$	$C = 6$ $N = 40$	$C = 7$ $N = 13$	$C = 8$ $N = 6$
1	0.060	0.086	0.100	0.275	0.462	1.000
2	0.268	0.321	0.500	0.575	0.846	1.000
3	0.375	0.414	0.614	0.625	0.846	1.000
4	0.196	0.307	0.329	0.525	0.769	0.667
5	0.072*	0.105*	0.151*	0.171*	0.400*	0.500
6	0.065*	0.054*	0.250*	0.071*	1.000*	0.000*
7	0.133*	0.042*	0.333*	0.000*		
8	0.250*	0.000*				

TABLE VI

The discrimination power of some Cluj-type indices on the set of all cycloalkanes with ten vertices and $C = 3...8$

Number of derivative	$C = 3$ $N = 168$	$C = 4$ $N = 140$	$C = 5$ $N = 70$	$C = 6$ $N = 40$	$C = 7$ $N = 13$	$C = 8$ $N = 6$
W	0.036	0.079	0.086	0.250	0.462	1.000
CJD_p	0.464	0.507	0.557	0.950	0.846	1.000
CJA_p	0.363	0.293	0.686	0.625	0.846	1.000
$D^2_CJD_p$	0.952	0.957	0.971	0.925	1.000	1.000
$\Delta^2_CJA_p$	0.917	0.943	0.914	1.000	1.000	1.000
H_{CJD_p}	0.988	1.000	1.000	1.000	1.000	1.000
H_{CJA_p}	0.810	0.971	1.000	0.950	1.000	0.667

error of the ordering program. In the same ref. the other values of S are different from the actually reported ones, because of different interpretation of Eq. (25).

The same conclusion about the discriminating ability of the discussed descriptors: Cluj-type indices, as a variant of the Wiener index, are superior to the classical Wiener index and all Wiener polynomial derivatives.

Within this set of graphs, IDI showed $S = 1$ for all graphs with $C = 3, \dots, 8$.

PROPERTY CORRELATING TESTS

The usefulness of a topological index in exploring the molecular properties is directly measured by its correlating ability with such properties.

TABLE VII
Boiling Points and Topological Indices for Some Cycloalkanes

No	Graph*	BP °C	I^{st} Deriv.	I^{2nd} Deriv.	IDI	CJ_e	CJ_p	H_{De}	H_{CJp}
1	C4	13.1	8	4	6.000	16	18	5.000	4.667
2	11MC3	21	15	10	9.737	15	24	7.500	6.667
3	EC3	35.9	17	18	9.416	17	32	7.167	6.333
4	MC4	40.5	16	14	9.507	28	37	7.333	6.500
5	C5	49.3	15	10	9.591	20	40	7.500	6.667
6	112MC3	56.5	26	26	13.462	26	49	10.167	8.667
7	123MC3	66	27	30	13.402	27	54	10.000	8.500
8	EC4	70.7	29	40	13.145	45	73	9.750	8.267
9	MC5	71.8	26	26	13.362	33	71	10.167	8.667
10	C6	80.7	27	30	13.183	54	90	10.000	8.500
11	PC4	110	48	92	16.957	68	132	12.283	10.033
12	11MC5	88.9	39	44	17.477	48	105	13.333	11.000
13	12MC5	91.9**	40	48	17.435	49	109	13.167	10.833
14	13MC5	91.7**	41	54	17.324	51	119	13.083	10.767
15	MC6	100.9	42	58	17.213	78	142	12.917	10.600
16	C7	117	42	56	17.214	63	154	12.833	10.500
17	112MC5	114	56	72	21.762	67	150	15.567	12.367
18	113MC5	105	58	84	21.599	71	170	15.783	12.533
19	123MC5	115	58	82	21.664	70	164	16.000	12.700
20	1M2EC5	124	61	98	21.534	72	178	16.667	13.333
21	1M3EC5	121	63	112	21.378	76	199	16.500	13.200
22	PC5	131	67	136	21.201	78	215	16.417	13.100
23	IPC5	126.4	62	104	21.445	73	186	16.083	12.800
24	11MC6	119.5	59	88	21.521	104	197	15.950	12.700
25	12MC6	123.4**	60	92	21.509	106	202	16.000	12.733
26	13MC6	124.5**	61	98	21.434	108	211	16.333	13.033
27	14MC6	120	62	106	21.351	110	220	16.167	12.867
28	EC6	131.8	64	116	21.330	109	226	16.083	12.800
29	MC7	134	61	96	21.445	88	225	16.033	12.767
30	C8	146	64	112	21.245	128	288	15.667	12.400
31	1123MC5	132.7	78	118	26.131	93	222	20.167	15.700
32	113MC6	136.6	82	140	25.898	140	285	19.750	15.333
33	124MC6	136	84	152	25.838	144	296	19.533	15.133
34	135MC6	138.5	84	150	25.853	144	291	19.500	15.100
35	1M2EC6	151	86	162	25.846	142	300	19.283	14.900
36	1M3EC6	149	88	176	25.728	146	322	19.150	14.800
37	PC6	154	94	218	25.532	148	352	18.683	14.414
38	IPC6	146	88	176	25.731	142	313	19.150	14.800
39	EC7	163.5	88	174	25.748	121	337	19.067	14.700
40	C9	170	90	180	25.618	144	450	18.750	14.400
41	1M2IPC6	171	114	234	30.412	180	401	22.900	17.267
42	1M3IPC6	167.5	117	256	30.272	186	436	22.717	17.133
43	13EC6	170.5	121	284	30.197	192	467	22.383	16.848
44	PC7	183.5	124	306	30.112	163	503	22.133	16.629
45	C10	201	125	300	29.991	250	705	21.833	16.333

* M = methyl; E = ethyl; P = propyl; IP = isopropyl; C_n = n -membered cycle

** values for the *trans*-isomer

TABLE VIII
 Statistics of Multivariable Regression ($Y_{\text{calc}} = a + \sum_i b_i X_i$) for compounds
 of Table VII

No.	X_i	b_i	a	r	s	$cv(\%)$	F
single variable regression							
1	lw_1^*	1.313	34.784	0.9439	14.568	12.577	351.32
2	$\ln lw_1$	65.747	-144.322	0.9688	10.940	9.445	656.169
3	lw_2	0.486	61.974	0.8911	20.018	17.282	165.84
4	$\ln lw_2$	42.475	-69.199	0.9594	12.440	10.740	497.76
5	IDI	6.625	-20.158	0.9644	11.659	10.065	572.68
6	$\ln IDI$	111.933	-215.799	0.9560	12.940	11.171	456.82
7	CJ_p	0.273	56.644	0.9200	17.289	14.926	236.98
8	$\ln CJ_p$	50.707	-142.612	0.9909	5.934	5.123	2333.73
9	H_{CJ_p}	6.709	77.041	0.322	41.765	36.058	4.976
Two variable regression							
10	lw_1	2.470	14.684	0.9590	12.657	10.927	240.21
	lw_2	-0.463					
11	$\ln CJ_p$	45.483	-130.884	0.9913	5.880	5.076	1189.35
	IDI	0.726					
12	$\ln CJ_p$	42.759	-116.448	0.9928	5.339	4.609	1447.08
	lw_1	0.232					
13	$\ln CJ_p$	44.793	-120.539	0.9932	5.209	4.497	1521.20
	lw_2	0.073					
14	$\ln CJ_p$	42.893	-113.779	0.9939	4.934	4.260	1697.54
	CJ_p	0.501					
Three variable regression							
15	lw_1	-2.049	-36.044	0.9664	11.611	10.024	193.26
	lw_2	0.480					
	IDI	10.968					
16	lw_1	0.222	-36.862	0.9716	10.690	9.229	230.45
	$\ln lw_2$	19.371					
	IDI	2.659					
17	lw_1	-0.058	-12.637	0.9813	8.692	7.504	355.55
	H_{De}	-27.426					
	IDI	26.990					
18	$\ln CJ_p$	28.972	-84.513	0.9944	4.793	4.138	1200.63
	H_{Dp}	-2.798					
	IDI	3.354					
19	$\ln CJ_p$	44.871	-119.444	0.9948	4.619	3.987	1294.06
	CJ_p	0.092					
	CJ_e	-0.141					
20	$\ln CJ_p$	39.969	-93.266	0.9948	4.592	3.964	1309.31
	H_{Dp}	-13.249					
	IDI	8.150					

* $lw_1 = W$ (Wiener Index); $lw_2 =$ the 2nd derivative; $IDI =$ Information Distance Index

Both diagnostic and prognostic dimensions have to be considered in such studies.

The most studied molecular property was the boiling point due to its accessibility and direct relation to the chemical structure. Table VII collects the boiling points and some of the herein discussed topological indices for a set of 45 cycloalkanes.²³ Only the first and the second derivatives of the polynomial are calculated on this set of graphs. The discrimination powers for these indices are 0.511 and 0.689, respectively. Statistics of multilinear regression are listed in Table VIII.

In single variable regression, the only satisfactory results were offered by the natural logarithm of the index values (entries 2, 4, 8 – Table VIII);

TABLE IX
Structural formula for some cycloalkanes

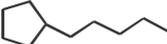
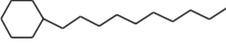
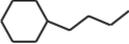
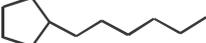
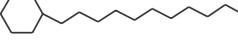
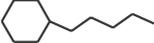
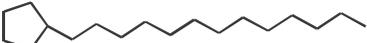
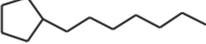
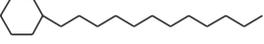
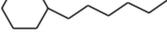
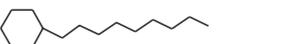
No.	Structural formula	No.	Structural formula
1		14	
2		15	
3		16	
4		17	
5		18	
6		19	
7		20	
8		21	
9		22	
10		23	
11		24	
12		25	
13			

TABLE X

Wiener and Cluj with their corresponding Harary-type indices and viscosity ($\log \eta$) in cycloalkanes of Table IX

No	W	CJD_p	CJD_e	CJA_p	CJA_e	H_{CJD_p}	H_{CJD_e}	H_{CJA_p}	H_{CJA_e}	$\log \eta$
1	140	520	155	330	110	5.19972	0.83861	14.88107	3.63623	0.061
2	133	529	196	285	84	5.05179	0.54861	16.94385	4.66290	0.117
3	191	756	208	510	157	5.60567	0.79960	15.58544	3.58532	0.274
4	182	767	254	449	126	5.40179	0.51627	17.69940	4.59960	0.235
5	253	1062	272	752	215	5.99900	0.76634	16.23670	3.54412	0.373
6	242	1077	323	673	179	5.75428	0.48777	18.39293	4.55126	0.344
7	327	1450	348	1068	285	6.37898	0.73752	16.84366	3.50974	0.464
8	314	1471	404	969	244	6.10231	0.46252	19.03540	4.51252	0.447
9	414	1933	437	1471	368	6.74571	0.71223	17.41296	3.48041	0.550
10	399	1962	498	1350	322	6.44291	0.44001	19.63506	4.48041	0.544
11	515	2525	540	1975	465	7.09966	0.68981	17.94974	3.45496	0.550
12	498	2564	606	1830	414	6.77482	0.41981	20.19818	4.45315	0.634
13	631	3241	658	2595	577	7.44146	0.66977	18.45809	3.43259	0.631
14	612	3292	729	2424	521	7.09757	0.40159	20.72968	4.42956	0.719
15	763	4097	792	3347	705	7.77178	0.65172	18.94131	3.41271	0.708
16	742	4162	868	3148	644	7.41116	0.38505	21.23347	4.40886	0.801
17	912	5110	943	4248	850	8.09129	0.63535	19.40212	3.39488	0.781
18	889	5191	1024	4019	784	7.71575	0.36997	21.71274	4.39048	0.876
19	1079	6298	1112	5316	1013	8.40065	0.62044	19.84280	3.37877	0.852
20	1054	6397	1198	5055	942	8.01167	0.35615	22.17011	4.37401	0.949
21	1265	7680	1300	6570	1195	8.70047	0.60678	20.26526	3.36413	0.919
22	1471	9276	1508	8030	1397	8.99133	0.59421	20.67116	3.35075	0.982
23	67	215	78	113	45	4.36508	0.94286	13.25952	3.79286	-0.167
24	43	127	52	57	25	3.96667	1.01667	12.30000	3.93333	-0.248
25	94	352	148	170	52	4.72262	0.58532	16.11151	4.75198	0.001

exception, $\ln IDI$, whose correlation coefficient, r , was lower than that given by IDI .

In double variable regression, the correlation was improved, as indicated by the drop in s (the standard error of estimate) – to around 5 °C, which is lower than 5% in the coefficient of variance, v). However, the two derivatives of the Wiener polynomials are rather unsuitable for this purpose. No major changes appeared in the correlation adding the third variable. IDI can be used in association with other indices.

A second property in the correlating test was the viscosity (as $\log \eta$) of a set of 25 cycloalkanes⁴⁰ (Table IX). Topological indices and viscosity are

TABLE XI

MLR data: $Y = a + \sum_i b_i X_i$; $Y =$ viscosity of structures in Table IX

No.	X_i	b_i	a	r	s	$cv(\%)$	F
1	W	0.0008	0.0789	0.9034	0.1532	30.903	102.031
2	CJD_p	0.0001	0.1489	0.8760	0.1721	34.746	75.897
3	CJD_e	0.0008	0.0387	0.9213	0.1389	28.020	129.074
4	$CJ\Delta_p$	0.0001	0.1872	0.8530	0.1864	37.598	61.459
5	$CJ\Delta_e$	0.0008	0.1178	0.8887	0.1638	33.037	86.392
6	$\ln W$	0.3651	-1.6642	0.9932	0.0416	8.387	1674.472
7	$\ln CJD_p$	0.3018	-1.7589	0.9944	0.0377	7.600	2044.310
8	$\ln CJD_e$	0.3905	-1.8722	0.9922	0.0446	9.002	1450.474
9	$\ln CJ\Delta_p$	0.2619	-1.3651	0.9920	0.0451	9.094	1420.770
10	$\ln CJ\Delta_e$	0.3172	-1.3142	0.9887	0.0536	10.824	996.154
11	H_{CJD_p}	0.2467	-1.1345	0.9737	0.0813	16.403	420.770
12	H_{CJD_e}	-1.2242	1.2428	0.6225	0.2796	56.387	14.551
13	$H_{CJ\Delta_p}$	0.1266	-1.8127	0.9205	0.1396	28.157	127.597
14	$H_{CJ\Delta_e}$	-0.1510	1.0933	0.2260	0.3480	70.185	1.238
15	H_{CJD_p} $H_{CJ\Delta_p}$	0.1717 0.0487	-1.5276	0.9930	0.0430	8.678	781.792
16	$\ln CJD_p$ $H_{CJ\Delta_p}$	0.2703 0.0158	-1.8114	0.9956	0.0341	6.883	1249.072
17	$\ln W$ $H_{CJ\Delta_p}$	0.3115 0.0224	-1.7563	0.9959	0.0331	6.668	1331.656
18	W H_{CJD_p} $H_{CJ\Delta_p}$	-0.0002 0.2204 0.0468	-1.7274	0.9947	0.0386	7.780	650.453
19	H_{CJD_p} $H_{CJ\Delta_p}$ $H_{CJ\Delta_e}$	-0.0047 0.1376 -0.2651	-0.9343	0.9964	0.0318	6.408	962.190
20	$\ln CJD_p$ CJD_p $CJ\Delta_p$	0.2684 0.0002 -0.0002	-1.5958	0.9969	0.0292	5.879	1137.362
21	$\ln CJ\Delta_p$ CJD_p $CJ\Delta_p$	0.2169 0.0003 -0.0003	-1.1728	0.9970	0.0289	5.833	1162.764
22	$\ln W$ CJD_p $CJ\Delta_p$	0.3160 0.0003 -0.0003	-1.4796	0.9971	0.0286	5.767	1189.530

listed in Table X, while the statistics of multilinear regression (MLR) appear in Table XI.

The logarithm of the values of both Wiener and Cluj-type indices led to good correlation coefficients (over 0.99, already in single variable regression) and coefficients of variance less than 10%.

A cross-validation procedure (leave one out – loo) indicated a good predicting ability of our indices: $\ln CJD_p$ (entry 7 – Table XI), $r_{(loo)} = 0.9933$; $s = 0.0414$; $v\% = 8.349$; $\ln W \& H_{CJ\Delta p}$ (entry 17), $r_{(loo)} = 0.9946$; $s = 0.0371$; $v\% = 7.487$; $\ln CJD_p \& CJ\Delta_p$ (entry 20), $r_{(loo)} = 0.9957$; $s = 0.0330$; $v\% = 6.661$; $\ln W \& CJ\Delta_p$ (entry 22), $r_{(loo)} = 0.9957$; $s = 0.0329$; $v\% = 6.638$.

CONCLUSIONS

The present study proposes the Wiener polynomial derivatives as topological descriptors. Their discriminating power and correlating ability were tested on selected sets of molecular graphs, in comparison with the Information Distance Index, *IDI*, and some Cluj-type indices. The data obtained on large sets of isomeric, nonisomorphic structures, including polyhexes and cycle-containing structures with side-chains, indicate that these »extended Wiener« indices are not especially good descriptors, though their discriminating ability increases with the order of derivative. But *IDI* and some of the Cluj indices showed a much more discriminating power.

In the property correlating tests, the Wiener-type descriptors, *IDI* and Cluj indices complement each other in modeling some physico-chemical properties such as the boiling point and viscosity of alkanes and cycloalkanes.

Acknowledgements. – The authors wish to thank Dr. E. Estrada and Dr. I. Gutman for valuable discussions.

REFERENCES

1. D. H. Rouvray, in: R. B. King (Ed.), *Chemical Applications of Topology and Graph Theory*, Elsevier, Amsterdam, 1983, pp. 159–177.
2. D. H. Rouvray, *J. Mol. Structure (Theochem)* **185** (1989) 187–201.
3. N. Trinajstić, *Chemical Graph Theory*, CRC Press, Boca Raton, FL, 1983, Vol. 2, Chapter 4, p. 105.
4. I. Gutman and O. Polansky, *Mathematical Concepts in Organic Chemistry*, Springer, Berlin, 1986.
5. A. T. Balaban, *J. Chem. Inf. Comput. Sci.* **25** (1985) 334–343.
6. D. Bonchev, *Information Theoretic Indices for Characterization of Chemical Structure*, Research Studies Press, Chichester, 1983.
7. H. Hosoya, *Theor. Chim. Acta (Berlin)* **25** (1972) 215–222.
8. H. Hosoya and K. Hosoi, *J. Chem. Phys.* **64** (1976) 1065–1073.

9. H. Hosoya, K. Hosoi, and I. Gutman, *Theor. Chim. Acta (Berlin)* **38** (1975) 37–47.
10. J. Aihara, *J. Am. Chem. Soc.* **98** (1976) 2750–2758.
11. I. Gutman, M. Milun, and N. Trinajstić, *J. Am. Chem. Soc.* **99** (1977) 1692–1704.
12. E. J. Farrell, *J. Combin. Theory Ser. B* **27** (1979) 60–86.
13. H. Hosoya, Y. Kisu, and H. Marumi, *Rotational Polynomial, Proceedings of the Symposium on Molecular Structure*, Sapporo, 1977.
14. N. Ohkami, A. Motoyama, T. Yamaguchi, H. Hosoya, and I. Gutman, *Tetrahedron*, **37** (1981) 1113–1122.
15. H. Hosoya, M. Murakami, and M. Gotoh, *Natl. Sci. Rept. Ochanomizu Univ.* **24** (1973) 27–34.
16. H. Hosoya, *Discrete Appl. Math.* **19** (1988) 239–257.
17. H. Wiener, *J. Am. Chem. Soc.* **69** (1947) 17–20.
18. E. Estrada, O. Ivanciuc, I. Gutman, A. Gutierrez, and L. Rodrigues, *New J. Chem.*, in press.
19. D. Bonchev, *Commun. Math. Comput. Chem. (MATCH)* **7** (1979) 65–113.
20. C. E. Shannon and W. Weaver, *Mathematical Theory of Communications*, University of Illinois, Urbana, 1949.
21. E. V. Konstantinova and A. A. Paleev, *Vyshisl. Systemy* **136** (1990) 38–48 (in Russian).
22. E. V. Konstantinova, *J. Chem. Inf. Comput. Sci.* **36** (1996) 54–57.
23. M. V. Diudea, *J. Chem. Inf. Comp. Sci.* **37** (1997) 300–305.
24. M. V. Diudea, B. Parv, and I. Gutman, *J. Chem. Inf. Comput. Sci.* **37** (1997) 1101–1108.
25. M. V. Diudea and I. Gutman, *Croat. Chem. Acta*, **71** (1998) 21–51.
26. R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1985.
27. M. Randić, X. Guo, T. Oxley, and H. Krishnapriyan, *J. Chem. Inf. Comput. Sci.* **33** (1993) 709–716.
28. M. Randić, X. Guo, T. Oxley, H. Krishnapriyan, and L. Naylor, *J. Chem. Inf. Comput. Sci.* **34** (1994) 361–367.
29. S. S. Tratch, M. I. Stankevitch, and N. S. Zefirov, *J. Comput. Chem.* **11** (1990) 899–908.
30. M. V. Diudea, *Commun. Math. Comput. Chem. (MATCH)* **35** (1997) 169–183.
31. O. Ivanciuc, T. S. Balaban, and A. T. Balaban, *J. Math. Chem.* **12** (1993) 309–318.
32. Z. Mihalić, S. Nikolić, and N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **32** (1992) 28–37.
33. D. Plavšić, S. Nikolić, N. Trinajstić, and Z. Mihalić, *J. Math. Chem.* **12** (1993) 235–250.
34. M. V. Diudea, *J. Chem. Inf. Comput. Sci.* **37** (1997) 292–299.
35. D. Bonchev, O. Mekenyan, and N. Trinajstić, *J. Comput. Chem.* **2** (1981) 127–148.
36. M. Razinger, J. R. Chretien, and J. K. Dubois, *J. Chem. Inf. Comput. Sci.* **25** (1985) 23–27.
37. N. Trinajstić, *J. Math. Chem.* **5** (1990) 171–175.
38. E. V. Mzhel'skaya, *Vyshisl. Systemy* **127** (1988) 143–151 (in Russian).
39. A. A. Dobrynin, *Vyshisl. Systemy* **122** (1983) 3–38 (in Russian).
40. V. Pachaiyappan, S. H. Ibrahim, and N. R. Kuloor, *Chem. Eng.* **74** (1967) 193–196.

SAŽETAK**Wienerov polinom i drugi topologijski indeksi u kemijskom istraživanju**

Elena V. Konstantinova i Mircea V. Diudea

Proučavani su Wienerovi polinomi i neki drugi informacijski i topologijski indeksi s obzirom na njihova diskriminacijska i korelacijska svojstva.