Journal of Mathematical Chemistry 11(1992)223–258

THE DISTANCE MATRIX IN CHEMISTRY*

Zlatko MIHALIĆ, Darko VELJAN
The Faculty of Science and Mathematics, The University of Zagreb, 41000 Zagreb, The Republic of Croatia
Dragan AMIĆ
The Faculty of Agriculture, The University of Osijek, 54001 Osijek, The Republic of Croatia
Sonja NIKOLIĆ, Dejan PLAVŠIĆ and Nenad TRINAJSTIĆ
The Rugjer Bošković Institute, 41001 Zagreb, The Republic of Croatia

> *"... the progress of chemistry depends more than ever on applications of mathematics."*
>
> Vladimir Prelog (1986) [1]

Abstract

The graph-theoretical (topological) distance matrix and the geometric (topographic) distance matrix and their invariants (polynomials, spectra, determinants and Wiener numbers) are presented. Methods of computing these quantities are discussed. The uses of the distance matrix in both forms and the related invariants in chemistry are surveyed. Special attention is paid to the 2D and 3D Wiener numbers, defined respectively as one half of the sum of entries in the topological distance matrix and the topographic distance matrix. These numbers appear to be very valuable molecular descriptors in the structure–property correlations.

1. Introduction

The distance matrix is a mathematical object which is being increasingly used in both graph-theoretical (topological) [2–4] and geometric (topographic) [5–7] versions in chemistry [8–14]. The distance matrix has also found a considerable use in other areas much less mathematical than applied mathematics [2–4,7,15–22], physics [5,6,23] or chemistry such as [10] anthropology, archeology, genetics, geology, history, ornithology, philology, psychology, sociology, etc. The origins of the distance matrix may be traced back to the very first paper of Cayley [24]. However, this matrix was first introduced in rudimentary form by Brunel [25] in 1895.

---

*Dedicated to Professor Vladimir Prelog (Zurich) on the happy occasion of his 85th birthday.

The distance matrix is used in chemistry in explicit and implicit forms [10]. The earliest explicit use of the distance matrix in chemistry is work by Clark and Kettle [26] in 1975, although in biochemistry it was used in disguise somewhat earlier (in 1971) [27]. Clark and Kettle have used the distance matrix for studying the permutational isomers of stereochemically nonrigid molecules. These authors have distinguished between the various interconversion mechanisms for pairs of permutational isomers by determining the shortest path sequence necessary to effect the rearrangement. These shortest paths were used in the construction of an appropriate distance matrix. Clark and Kettle have also mentioned in their work that to their knowledge, there are no details on the distance matrix in the chemical literature. (This was not quite so because of the paper by Hosoya [28] which appeared in 1971 and in which the author has defined the distance matrix.) Therefore, the purpose of their work was also to remedy this omission and to introduce the distance matrix to the chemical community-at-large. Today, of course, the situation is quite different after the well-written review articles by Rouvray [8,10,29] and books on chemical graph theory edited or authored by Balaban [30], King [31], and Trinajstić [9,14,32]. However, there is still a lack of a single source of information in the literature on the topological and topographic distance matrices, their invariants and applications. The present authors hope to remedy this situation with this article. This was indeed the main motivation for preparing the article.

The distance matrix in explicit form is also employed to generate the distance polynomials and the distance spectra of various molecular structures [13,33–35].

The earliest implicit use of the distance matrix in chemistry was made, albeit unknowingly, by Wiener [36] in 1947. He wanted to develop a structure–property model for predicting physical parameters of alkanes. In order to do that, Wiener very cleverly introduced the path number as a numerical characteristic of a molecule. The path number he defined as the sum of the distances between any two carbon atoms in alkane in terms of the carbon–carbon bonds. Wiener also introduced the polarity number, which is defined as the number of pairs of carbon atoms separated by three carbon–carbon bonds. By using a linear combination of the path number and the polarity number, Wiener was able to obtain a fair prediction of alkane boiling points [36]. In view of the pioneering contribution of Wiener in recognizing the significance of the number of paths in a molecular skeleton, the term "the Wiener index" (e.g. ref. [37]) or "the Wiener number" (e.g. ref. [38]) has been adopted for the number of distances in all chemical structures. The connection between the distance matrix and the Wiener number was first pointed out by Hosoya [28]: The Wiener number is simply equal to the half-sum of the elements of the distance matrix.

The distance matrix in both topological and topographic forms has been used continuously as a source for deriving novel topological and topographic indices [39–54]. This appears to date to be one of the most important uses of this matrix in chemistry [8,10,29,55–57].

This article is structured as follows. In section 2, the graph-theoretical distance matrix and the related topological invariants (the distance polynomial, the distance spectrum, the determinant and the two-dimensional Wiener number) and some of their uses in chemistry are presented. In section 3, the geometric-distance matrix and the corresponding topographic invariants (the geometric-distance polynomial, the geometric-distance spectrum, the determinant and the three-dimensional Wiener number) and some of their applications in chemistry are discussed. The article ends with concluding remarks in section 4.

2. The graph-theoretical distance matrix and the related invariants

2.1. THE DISTANCE MATRIX

The distance matrix $\mathbf{D} = \mathbf{D}(G)$ of a labeled connected graph G is a real symmetric $N \times N$ matrix whose elements $(\mathbf{D})_{ij}$ are defined as [2,9,14]

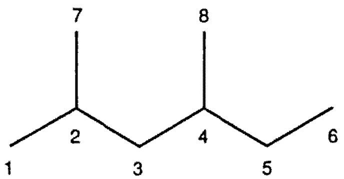
$$(\mathbf{D})_{ij} = \begin{cases} l_{ij} & \text{if } i \neq j \\ 0 & \text{if } i = j, \end{cases} \tag{1}$$

where l_{ij} is the length of the shortest path, i.e. the minimum number of edges, between the vertices i and j in G. The length l_{ij} is also called the distance [58–60] between the vertices i and j in G, hence the name "distance martix". The term related to the length has been also used. For example, in 1956 Kruskal used the term "matrix of lengths" [61]. Certainly, this latter term is more poetic, but less handy. Therefore, the pragmatic mathematicians adopted the former term for their use. The graph-theoretical distance matrix will be denoted in this article by 2-**D**. The meaning of 2 is that the matrix in question is associated with the two-dimensional realization of molecular graphs. As an example, the distance matrix of a labeled tree T corresponding to the carbon skeleton of 2,4-dimethylhexane is given in fig. 1.

In the case of edge-weighted graphs G_{EW} [62–64], the distance matrix entries are edge-weighted [15,16,22,60]:

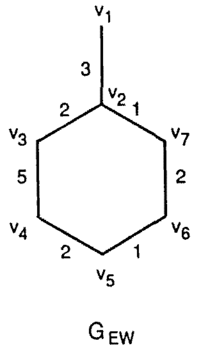
$$(\mathbf{D})_{ij} = \begin{cases} w_{ij} & \text{if } i \neq j \\ 0 & \text{if } i = j, \end{cases} \tag{2}$$

where w_{ij} is the minimum sum of edge-weights along the path between the vertices i and j, which is not necessarily the shortest possible path between these two vertices in G_{EW} as it would be in terms of just unweighted edges. Hence, in the case of the weighted distance matrix, the entry $(\mathbf{D})_{ij}$ is the minimum path-weight between the vertices i and j in G_{EW} [16]. The weighted distance matrices, even with negative entries, have been studied [60]. If the edge-weights in G_{EW} are all unity, then (1) and (2) are clearly identical. The distance matrix of an edge-weighted graph G_{EW} is given in fig. 2.



$$2\text{-}\mathbf{D}\,(T) = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 2 & 4 \\ 1 & 0 & 1 & 2 & 3 & 4 & 1 & 3 \\ 2 & 1 & 0 & 1 & 2 & 3 & 2 & 2 \\ 3 & 2 & 1 & 0 & 1 & 2 & 3 & 1 \\ 4 & 3 & 2 & 1 & 0 & 1 & 4 & 2 \\ 5 & 4 & 3 & 2 & 1 & 0 & 5 & 3 \\ 2 & 1 & 2 & 3 & 4 & 5 & 0 & 4 \\ 4 & 3 & 2 & 1 & 2 & 3 & 4 & 0 \end{bmatrix}$$

Fig. 1. The graph-theoretical distance matrix of a (labeled) tree T corresponding to the carbon skeleton of 2,4-dimethylhexane.



$$2\text{-}\mathbf{D}\,(G_{EW}) = \begin{bmatrix} 0 & 3 & 5 & 9 & 7 & 6 & 4 \\ 3 & 0 & 2 & 6 & 4 & 3 & 1 \\ 5 & 2 & 0 & 5 & 6 & 5 & 3 \\ 9 & 6 & 5 & 0 & 2 & 3 & 5 \\ 7 & 4 & 6 & 2 & 0 & 1 & 3 \\ 6 & 3 & 5 & 3 & 1 & 0 & 2 \\ 4 & 1 & 3 & 5 & 3 & 2 & 0 \end{bmatrix}$$

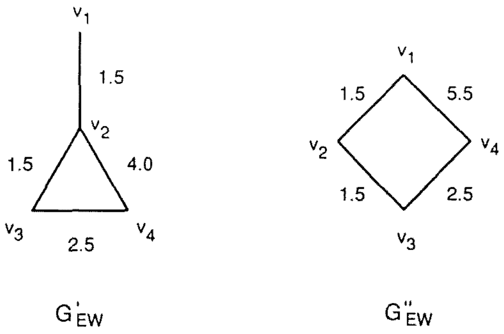
Fig. 2. The distance matrix of a labeled edge-weighted graph G_{EW}.

The diagonal elements of the distance matrices belonging to the vertex- and edge-weighted graphs G_{VEW} [64–66] no longer vanish [67]:

$$(\mathbf{D})_{ij} = \begin{cases} w_{ij} & \text{if } i \neq j \\ w_{ii} & \text{if } i = j, \end{cases} \tag{3}$$

where w_{ii} is the weight of the vertex i in G_{VEW}. Since in chemical graph theory the vertex- and edge-weighted graphs serve to depict molecules containing heteroatoms [64], there are a number of proposals in the chemical literature to define the vertex- and edge-weights [41,55,67–71].

The construction of the distance matrix for large complex graphs is a nontrivial task and for its completion, a computer must be used. Most commonly, the distance matrix of a graph G has been generated using powers of the corresponding adjacency matrix of G [72,73]. However, there are algorithms and computer programs in the literature which are much faster than the matrix power method [35,74–80]. For example, Bersohn's algorithm [75] is about thirty times faster than the method based on powers of the adjacency matrix when employed to construct the distance matrices for steroids on an IBM 3033 computer. The weak point of the matrix power method is the number of matrix multiplications required by this procedure, since computers multiply much more slowly than they store, fetch, add and subtract. The method of Bersohn does not involve any multiplications at all. One of the fastest methods to date for the construction of the distance matrix appears to be [49] the computational method developed by Müller et al. [78].



$$2\text{-}\mathbf{D}\,(G'_{\mathrm{EW}}) = 2\text{-}\mathbf{D}\,(G''_{\mathrm{EW}}) = \begin{bmatrix} 0.0 & 1.5 & 3.0 & 5.5 \\ 1.5 & 0.0 & 1.5 & 4.0 \\ 3.0 & 1.5 & 0.0 & 2.5 \\ 5.5 & 4.0 & 2.5 & 0.0 \end{bmatrix}$$

Fig. 3. Two edge-weighted graphs G'_{EW} and G''_{EW} with identical distance matrices.

There are several interesting observations concerning the distance matrix. For example, the adjacency matrix and the distance matrix are identical for complete graphs. Similarly, non-isomorphic weighted graphs with identical distance matrices can be constructed (e.g. ref. [16]). One such example is shown in fig. 3. The same is not possible in the case of the adjacency matrix: Two non-isomorphic (unweighted or weighted) graphs cannot possess identical adjacency matrices.

The distance matrix can also be of help [75] in partitioning the atoms of the molecule into equivalence classes [81–84]. For instance, if there is a unique atom, atoms which have distances of different lengths from the unique atom must be in different equivalence classes. In general, the set of distances to distinguishable kinds of atoms could be made the basis for further partitions, and so on, recursively [75]. A similar procedure was used for detecting the graph center in polycyclic graphs [85–87].

2.2. THE DISTANCE POLYNOMIAL AND THE DISTANCE SPECTRUM

The distance polynomial $\delta(G,x)$ of a graph G is defined as [18,19,33,34)

$$\delta(G;x) = \det|x\mathbf{I} - \mathbf{D}|, \tag{4}$$

where $\mathbf{I}$ is the $N \times N$ unit matrix. The coefficient form of the distance polynomial is given by

$$\delta(G;x) = c_0 x^N - \sum_{n=1}^{N} c_n x^{N-n}. \tag{5}$$

The expansion of the determinant (4) produces the coefficients of the polynomial:

$$\delta(G;x) = c_0 x^N - c_1 x^{N-1} - c_2 x^{N-2} - \ldots - c_{N-1} x - c_N\,. \tag{6}$$

However, this is a time-consuming procedure which can be replaced by faster computer-oriented methods such as the modified forms of the Le Verrier–Faddeev–Frame (LVFF) method [13,35,88–94]. The modified form of the LVFF method, which uses diagonal matrices, is employed in this work [95]. The distance matrix can be conveniently diagonalized by means of the Householder tridiagonalization-QL algorithm (e.g. ref. [96]). Once the matrix is brought to a tridiagonal form, the eigenvalues are directly read. The collection of eigenvalues of the distance matrix is called the distance spectrum of G. For example, the distance matrix and distance spectrum of benzene are given below:

$$2\text{-}\mathbf{D}\ (\text{benzene}) = \begin{bmatrix} 0 & 1 & 2 & 3 & 2 & 1 \\ 1 & 0 & 1 & 2 & 3 & 2 \\ 2 & 1 & 0 & 1 & 2 & 3 \\ 3 & 2 & 1 & 0 & 1 & 2 \\ 2 & 3 & 2 & 1 & 0 & 1 \\ 1 & 2 & 3 & 2 & 1 & 0 \end{bmatrix}, \tag{7}$$

$$\text{distance spectrum: } \{9, 0, 0, -1, -4, -4\}. \tag{8}$$

The coefficients c_n in the LVFF method can be computed by

$$c_n = \frac{1}{n} \operatorname{tr} \mathbf{D}_n , \tag{9}$$

where the matrix $\mathbf{D}_n$ is given by

$$\mathbf{D}_n = \mathbf{D}\mathbf{B}_n . \tag{10}$$

In this expression, $\mathbf{D}$ is the distance matrix and $\mathbf{B}_n$ is an auxiliary matrix defined as

$$\mathbf{B}_n = \mathbf{D}_n - c_n \mathbf{I}. \tag{11}$$

The above is an iterative procedure which ends when the $\mathbf{B}$-matrix vanishes, i.e. when $n = N$:

$$\mathbf{B}_N = \mathbf{D}_N - c_N \mathbf{I} = \mathbf{O}. \tag{12}$$

Since in the modified LVFF method the distance matrix enters in the diagonal form, instead of $\mathbf{D}_n$ and $\mathbf{B}_n$ matrices one uses their diagonal forms. Therefore, instead of multiplying matrices, one multiplies sets of numbers representing eigenvalues of these matrices. Therefore, the above formulae (9)–(12) may be given in the different form:

$$c_n = \frac{1}{n} \sum_{i=1}^{N} (\mathbf{D}_n)_{ii} , \tag{13}$$

$$(\mathbf{D}_n)_{ii} = (\mathbf{D})_{ii} (\mathbf{B}_n)_{ii} , \tag{14}$$

$$(\mathbf{B}_n)_{ii} = (\mathbf{D}_n)_{ii} - (c_n \mathbf{I})_{ii} , \tag{15}$$

$$(\mathbf{B}_N)_{ii} = (\mathbf{D}_N)_{ii} - (c_N \mathbf{I})_{ii} = 0. \tag{16}$$

This modification of the LVFF method may be schematically presented as follows:

$$\begin{aligned} G \to \mathbf{D} \to \{x_i\} \to \sum_i (\mathbf{D}_1)_{ii} \to c_1 \to \sum_i (\mathbf{B}_1)_{ii} \to \sum_i (\mathbf{D}_2)_{ii} \to c_2 \to \ldots \to \\ c_{N-1} \to \sum_i (\mathbf{B}_{N-1})_{ii} \to \sum_i (\mathbf{D}_N)_{ii} \to c_N \to \sum_i (\mathbf{B}_N)_{ii} \to \mathbf{O}. \end{aligned} \tag{17}$$

As an example, the distance spectrum of benzene (8) will be used for the computation of the distance polynomial of benzene via the modified LVFF method. This is shown in table 1.

Table 1

The computation of the distance polynomial of benzene via the modified Le Verrier–Faddeev–Frame method.

(i) Distance spectrum of benzene

$\{9, 0, 0, -1, -4, -4\}$

(ii) $\sum_i (\mathbf{D})_{ii} = \sum_i (\mathbf{D}_1)_{ii} = 0$

$c_1 = 0$

(iii) $\{(\mathbf{B}_1)_{ii} = (\mathbf{D}_1)_{ii} - (c_1\mathbf{I})_{ii}\}_{i=1,\dots,6} = \{9, 0, 0, -1, -4, -4\}$

$\{(\mathbf{D}_2)_{ii} = (\mathbf{D})_{ii}\,(\mathbf{B}_1)_{ii}\}_{i=1,\dots,6} = \{81, 0, 0, 1, 16, 16\}$

$$c_2 = \frac{1}{2}\sum_i (\mathbf{D}_2)_{ii} = 57$$

(iv) $\{(\mathbf{B}_2)_{ii} = (\mathbf{D}_2)_{ii} - (c_2\mathbf{I})_{ii}\}_{i=1,\dots,6} = \{24, 0, 0, -56, -41, -41\}$

$\{(\mathbf{D}_3)_{ii} = (\mathbf{D})_{ii}\,(\mathbf{B}_2)_{ii}\}_{i=1,\dots,6} = \{216, 0, 0, 56, 164, 164\}$

$$c_3 = \frac{1}{3}\sum_i (\mathbf{D}_3)_{ii} = 200$$

(v) $\{(\mathbf{B}_3)_{ii} = (\mathbf{D}_3)_{ii} - (c_3\mathbf{I})_{ii}\}_{i=1,\dots,6} = \{16, 0, 0, -144, -36, -36\}$

$\{(\mathbf{D}_4)_{ii} = (\mathbf{D})_{ii}\,(\mathbf{B}_3)_{ii}\}_{i=1,\dots,6} = \{144, 0, 0, 144, 144, 144\}$

$$c_4 = \frac{1}{4}\sum_i (\mathbf{D}_4)_{ii} = 144$$

(vi) $\{(\mathbf{B}_4)_{ii} = (\mathbf{D}_4)_{ii} - (c_4\mathbf{I})_{ii}\}_{i=1,\dots,6} = \{0, 0, 0, 0, 0, 0\}$

(vii) Distance polynomial of benzene

$$\delta(\text{benzene}; x) = x^6 - 57x^4 - 200x^3 - 144x^2$$

It is interesting to point out that the distance polynomials of isospectral molecules (graphs) are generally different. A similar finding was also reported by Balasubramanian [35]. Isospectral molecules are non-isomorphic molecules (graphs) which possess identical spectra of their adjacency matrices [97–105]. The smallest pair of isospectral alkanes are two C_{10} alkanes [106]: 2,3-dimethyloctane (**1**) and 3,5-dimethyloctane (**2**). Trees corresponding to their carbon skeletons are shown in fig. 4. Their characteristic polynomials are identical: $\alpha(\mathbf{1}; x) \equiv \alpha(\mathbf{2}, x) = x^{10} - 9x^8 + 26x^6 - 27x^4 + 8x^2$, and, consequently, their eigenvalue spectra are also identical: spectrum of **1** ≡ spectrum of **2** = $\{\pm 2.089, \pm 1.681, \pm 1.149, \pm 0.701, 0, 0\}$. However, their distance polynomials are different:

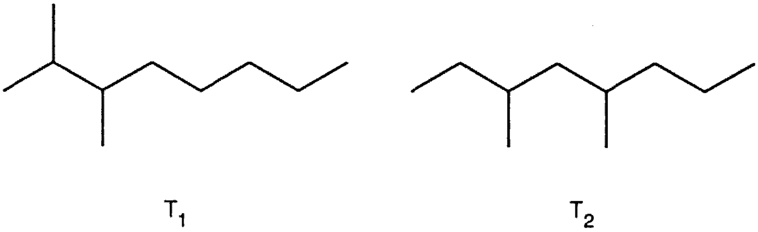


Fig. 4. Trees T_1 and T_2 depicting, respectively, the carbon skeletons of 2,3-dimethyloctane and 3,5-dimethyloctane.

$$\delta(\mathbf{1};x) = x^{10} - 593x^8 - 7256x^7 - 36624x^6 - 97568x^5 - 149536x^4 - 135424x^3 - 71168x^2 - 19968x - 2304, \tag{18}$$

$$\delta(\mathbf{2};x) = x^{10} - 540x^8 - 6800x^7 - 35264x^6 - 95744x^5 - 148416x^4 - 135168x^3 - 71168x^2 - 19968x - 2304, \tag{19}$$

and their distance spectra also differ:

$$\text{distance spectrum of } \mathbf{1} = \{29.709, -0.431, -0.547, -0.659, -0.894, -1.265, -2.000, -2.802, -4.839, -16.272\}, \tag{20}$$

$$\text{distance spectrum of } \mathbf{2} = \{28.705, -0.460, -0.501, -0.661, -0.874, -1.272, -1.933, -3.185, -5.301, -14.518\}. \tag{21}$$

An additional example is the classic pair of isospectral molecules [107]: 1,4-divinylbenzene (**3**) and 2-phenylbutadiene (**4**). Their graphs are given in fig. 5.

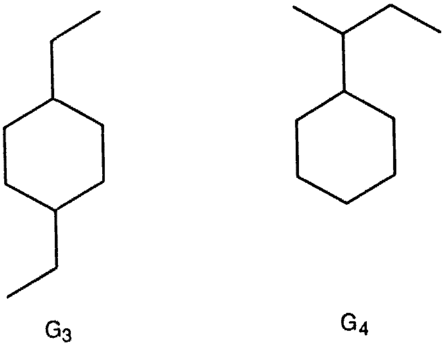


Fig. 5. Graphs G_3 and G_4 representing the carbon skeletons of 1,4-divinylbenzene and 2-phenylbutadiene.

Their characteristic polynomials and spectra are identical: $\alpha(\mathbf{3}; x) \equiv \alpha(\mathbf{4}; x) = x^{10} - 10x^8 + 33x^6 - 44x^4 + 24x^2 - 4$ and spectrum of **3** ≡ spectrum of **4** = {±2.214, ±1.675, ±1.000, ±1.000, ±0.539}. However, the distance polynomials of these two molecules are different:

$$\delta(\mathbf{3}; x) = x^{10} - 449x^8 - 5032x^7 - 22360x^6 - 48000x^5 - 51856x^4 - 27008x^3 - 5376x^2, \tag{22}$$

$$\delta(\mathbf{4}; x) = x^{10} - 405x^8 - 4540x^7 - 20588x^6 - 45808x^5 - 51280x^4 - 27264x^3 - 5376x^2, \tag{23}$$

and consequently their distance spectra are also different:

$$\text{spectrum of } \mathbf{3} = \{26.039, 0.000, 0.000, -0.597, -0.752, -1.000, -2.000, -4.000, -4.287, -13.403\}, \tag{24}$$

$$\text{spectrum of } \mathbf{4} = \{24.960, 0.000, 0.000, -0.493, -0.839, -1.121, -2.433, -3.935, -4.000, -12.140\}. \tag{25}$$

The above appears to be very encouraging in that the distance polynomials could be different and more discriminating compared to characteristic polynomials of graphs in general. However, for trees it has been shown that non-isomorphic trees can have the same distance polynomial [108, 109]. The smallest such pair, discovered by McKay [108] is shown in fig. 6. They possess the following distance polynomial [35]:

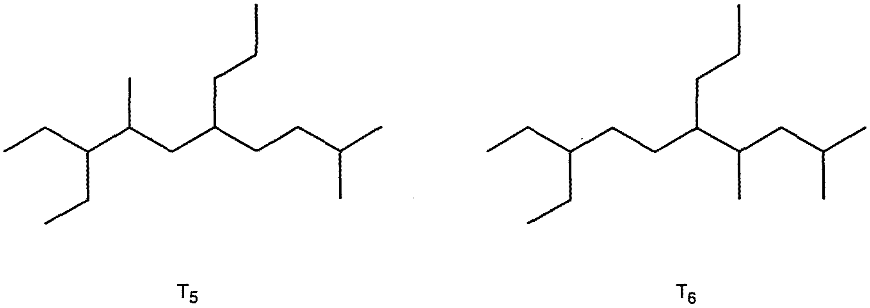


Fig. 6. Trees T_5 and T_6 depicting, respectively, the carbon skeletons of 8-ethyl-2,7-dimethyl-5-propyldecane and 8-ethyl-2,4-dimethyl-5-propyldecane.

$$\begin{aligned}\delta(T_5 \text{ or } T_6; x) = {} & x^{17} - 3176x^{15} - 110380x^{14} - 1683216x^{13} - 14561424x^{12} \\ & - 79661344x^{11} - 293089536x^{10} - 753990272x^9 - 1390462464x^8 \\ & - 1865146368x^7 - 1830330368x^6 - 1310025728x^5 - 674648064x^4 \\ & - 242974720x^3 - 57982976x^2 - 8224768x - 524288 \end{aligned} \tag{26}$$

and the distance spectrum:

$$\{71.385, -0.417, -0.452, -0.504, -0.592, -0.650, -0.764, -0.827, -1.198, \\ -1.406, -1.687, -2.000, -3.444, -5.236, -6.918, -14.565, -30.724\}. \tag{27}$$

Hence, in general, there could be two non-isomorphic trees with the same distance polynomial. Consequently, it is not evident if the distance polynomials are more discriminating than the characteristic polynomials.

The coefficients of the distance polynomials and the distance spectra of undirected and simple graphs possess a number of interesting properties [18, 19, 33–35, 109, 110]. In order that the reader can more easily follow at least some of the arguments given below, the distance polynomials of the lowest alkanes are given in table 2 and their distance spectra in table 3.

Some of the properties of the graph-theoretical distance polynomials and their eigenvalue spectra are summarized as follows:

(1) The coefficients c_0 and c_1 at x^N and x^{N-1} are, respectively, always unity and zero. The c_1 coefficient is equal to zero because of the relationship:

$$c_1 = \sum_{i=1}^{N} x_i = \operatorname{tr} \mathbf{D}, \tag{28}$$

where x_i are the elements of the distance spectrum. However, the sum of elements of the distance spectrum is always zero.

(2) The coefficient at x^{N-2} is the half-sum of the squares of elements of the distance matrix:

$$c_2 = \frac{1}{2} \sum_i \sum_j (\mathbf{D}^2)_{ij} . \tag{29}$$

(3) The last coefficient c_N of the distance polynomial of a tree T with N vertices is given by

$$c_N = (-1)^{N+1} (-2)^{N-2} (N-1). \tag{30}$$

This is so because the last coefficient of the (distance) polynomial is equal to the determinant of the (distance) matrix:

$$c_N = (-1)^N \det |\mathbf{D}| \tag{31}$$

Table 2

The graph-theoretical distance polynomials of the lowest alkanes.

Alkane	Coefficients of the polynomial								
	c_0	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8
propane	1	0	-6	-4					
2-methylpropane	1	0	-15	-28	-12				
butane	1	0	-20	-32	-12				
2,2-dimethylpropane	1	0	-28	-88	-96	-32			
2-methylbutane	1	0	-38	-116	-112	-32			
pentane	1	0	-50	-140	-120	-32			
2,2-dimethylbutane	1	0	-60	-272	-468	-336	-80		
2,3-dimethylbutane	1	0	-65	-296	-504	-352	-80		
2-methylpentane	1	0	-84	-368	-580	-368	-80		
3-methylpentane	1	0	-77	-356	-572	-368	-80		
hexane	1	0	-105	-448	-648	-384	-80		
2,2,3-trimethlybutane	1	0	-96	-584	-1464	-1776	-992	-192	
2,2-dimethylpentane	1	0	-122	-732	-1752	-1984	-1024	-192	
3,3-dimethylpentane	1	0	-108	-680	-1688	-1952	-1024	-192	
2,3-dimethylpentane	1	0	-120	-752	-1840	-2080	-1056	-192	
2,4-dimethylpentane	1	0	-134	-804	-1904	-2112	-1056	-192	
2-methylhexane	1	0	-164	-976	-2208	-2288	-1088	-192	
3-methylhexane	1	0	-148	-928	-2160	-2272	-1088	-192	
3-ethylpentane	1	0	-132	-876	-2112	-2256	-1088	-192	
heptane	1	0	-196	-1176	-2520	-2464	-1120	-192	
2,2,3,3-tetramethylbutane	1	0	-136	-1040	-3468	-6112	-5792	-2688	-448
2,2,3-trimethylpentane	1	0	-167	-1320	-4384	-7440	-6592	-2816	-448
2,3,3-trimethylpentane	1	0	-160	-1280	-4300	-7360	-6560	-2816	-448
2,2,4-trimethylpentane	1	0	-188	-1440	-4636	-7680	-6688	-2816	-448
2,2-dimethylhexane	1	0	-227	-1740	-5460	-8640	-7120	-2880	-448
3,3-dimethylhexane	1	0	-195	-1580	-5172	-8416	-7056	-2880	-448
3-ethyl-3-methylpentane	1	0	-172	-1440	-4908	-8208	-6992	-2880	-448
2,3,4-trimethylpentane	1	0	-179	-1428	-4740	-7968	-6928	-2880	-448
2,3-dimethylhexane	1	0	-216	-1744	-5644	-9024	-7392	-2944	-448
3-ethyl-2-methylpentane	1	0	-191	-1616	-5424	-8880	-7360	-2944	-448
3,4-dimethylhexane	1	0	-200	-1648	-5472	-8896	-7360	-2944	-448
2,4-dimethylhexane	1	0	-223	-1784	-5728	-9104	-7424	-2944	-448
2,5-dimethylhexane	1	0	-248	-1904	-5932	-9248	-7456	-2944	-448
2-methylheptane	1	0	-291	-2268	-6884	-10240	-7888	-3008	-448
3-methylheptane	1	0	-264	-2136	-6688	-10112	-7856	-3008	-448
4-methylheptane	1	0	-255	-2112	-6648	-10096	-7856	-3008	-448
3-ethylhexane	1	0	-228	-1968	-6444	-9968	-7824	-3008	-448
octane	1	0	-336	-2688	-7920	-11264	-8320	-3072	-448

Table 3

The eigenvalue spectra corresponding to the alkane graph-theoretical distance polynomials given in table 2.

Alkane	Eigenvalue spectrum							
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
propane	2.7321	– 0.7321	– 2.0000					
2-methylpropane	4.6458	– 0.6458	– 2.0000	– 2.0000				
butane	5.1623	– 0.5858	– 1.1623	– 3.4142				
2,2-dimethlypropane	6.6056	– 0.6056	– 2.0000	– 2.0000	– 2.0000			
2-methylbutane	7.4593	– 0.5120	– 1.0846	– 2.0000	– 3.8627			
pentane	8.2882	– 0.5578	– 0.7639	– 1.7304	– 5.2361			
2,2-dimethylbutane	9.6702	– 0.4727	– 1.0566	– 2.0000	– 2.0000	– 4.1409		
2,3-dimethylbutane	10.0000	– 0.4384	– 1.0000	– 2.0000	– 2.0000	– 4.5616		
2-methylpentane	11.0588	– 0.5115	– 0.6730	– 1.7026	– 2.0000	– 6.1717		
3-methylpentane	10.7424	– 0.4754	– 0.7639	– 1.3363	– 2.9307	– 5.2361		
hexane	12.1093	– 0.6359	– 0.6798	– 1.0000	– 2.4295	– 7.4641		
2,2,3,trimethylbutane	12.3945	– 0.3973	– 0.9692	– 2.0000	– 2.0000	– 2.0000	– 5.0279	
2,2-dimethylpentane	13.6353	– 0.4703	– 0.6481	– 1.6923	– 2.0000	– 2.0000	– 6.8246	
3,3-dimethlypentane	13.0698	– 0.4307	– 0.7639	– 1.2627	– 2.0000	– 3.3764	– 5.2361	
2,3-dimethylpentane	13.6346	– 0.4325	– 0.6651	– 1.3089	– 2.0000	– 3.0055	– 6.2228	
2,4-dimethylpentane	14.1760	– 0.5073	– 0.5359	– 1.6687	– 2.0000	– 2.0000	– 7.4641	
2-methylhexane	15.4048	– 0.4943	– 0.6242	– 0.9174	– 2.0000	– 2.4757	– 8.8932	
3-methylhexane	14.8636	– 0.4749	– 0.6461	– 0.9171	– 1.7796	– 3.3529	– 7.6929	
3-ethylpentane	14.2969	– 0.4559	– 0.7639	– 0.7639	– 1.8410	– 5.2361	– 5.2361	
heptane	16.6254	– 0.5272	– 0.6160	– 0.8405	– 1.2862	– 3.2576	– 10.0978	
2,2,3,3-tetramethylbutane	14.9373	– 0.3542	– 0.9373	– 2.0000	– 2.0000	– 2.0000	– 2.0000	– 5.6458
2,2,3-trimethylpentane	16.3152	– 0.3971	– 0.6305	– 1.2999	– 2.0000	– 2.0000	– 3.0325	– 6.9552
2,3,3-trimethylpentane	16.0683	– 0.3869	– 0.6625	– 1.2355	– 2.0000	– 2.0000	– 3.5063	– 6.2770
2,2,4-trimethylpentane	17.0338	– 0.4593	– 0.5144	– 1.6559	– 2.0000	– 2.0000	– 2.0000	– 8.4043
2,2-dimethylhexane	18.4133	– 0.4525	– 0.6053	– 0.8933	– 2.0000	– 2.0000	– 2.4942	– 9.9680
3,3-dimethylhexane	17.4426	– 0.4294	– 0.6348	– 0.8925	– 1.7432	– 2.0000	– 3.8238	– 7.9190
3-ethyl-3-methylpentane	16.6705	– 0.4074	– 0.7639	– 0.7639	– 1.4828	– 2.7803	– 5.2361	– 5.2361
2,3,4-trimethylpentane	16.8079	– 0.4197	– 0.5359	– 1.2747	– 2.0000	– 2.0000	– 3.1134	– 7.4641
2,3-dimethylhexane	18.1815	– 0.4310	– 0.6207	– 0.8046	– 1.7789	– 2.0000	– 3.5950	– 8.9514
3-ethy-2-methylpentane	17.4187	– 0.4275	– 0.6288	– 0.7639	– 1.8313	– 2.0000	– 5.2361	– 6.5311
3,4-dimethylhexane	17.6759	– 0.4268	– 0.6000	– 0.8566	– 1.4606	– 2.7447	– 3.6153	– 7.9719
2,4-dimethylhexane	18.3964	– 0.4731	– 0.5298	– 0.8643	– 1.7623	– 2.0000	– 3.4071	– 9.3598
2,5-dimethylhexane	19.1115	– 0.4605	– 0.5784	– 0.8095	– 2.0000	– 2.0000	– 2.5331	– 10.7300
2-methylheptane	20.4792	– 0.4941	– 0.5647	– 0.7852	– 1.2238	– 2.0000	– 3.3947	– 12.0167
3-methylheptane	19.7628	– 0.4694	– 0.5911	– 0.8405	– 1.0393	– 2.3198	– 3.7491	– 10.7535
4-methylheptane	19.5420	– 0.4746	– 0.6160	– 0.7079	– 1.2862	– 1.9227	– 4.4369	– 10.0978
3-ethylhexane	18.7788	– 0.4551	– 0.6409	– 0.7639	– 1.0251	– 2.4223	– 5.2361	– 8.2354
octane	21.8364	– 0.5198	– 0.5890	– 0.7232	– 1.0332	– 1.6199	– 4.2142	– 13-1371

and it has been shown [18] that the determinant of the distance matrix for a tree T is independent of the structure of T and that can be given only in terms of the number of vertices of T. This also explains why the c_N coefficients of distance polynomials belonging to isomeric alkanes are all the same.

(4) The distance spectrum consists of one positive and $N-1$ nonpositive elements. This is a result of a particular structure of the distance polynomial, i.e. all coefficients but the very first having the negative sign.

(5) The sum of the squares of the elements in the distance spectrum is equal to the trace of $\mathbf{D}^2$.

(6) For complete bipartite graphs K_{N_1,N_2} $(N_1+N_2=N)$, the elements in the distance spectrum are $-2(N_1+N_2-2)$ times and the remaining two elements are given by the following relationships:

$$x_1 + x_2 = 2(N_1 + N_2 - 2), \tag{32}$$

$$x_1^2 + x_2^2 = 2(2N_1^2 - 4N_1 + 2N_2^2 - 4N_2^2 + N_1N_2 + 4). \tag{33}$$

(7) If G is an even cycle, then at least one element in the distance spectrum is zero (for more details on the distance spectrum of a cycle, see Graovac et al. [34]).

(8) If i is the number of zero elements in the distance spectrum, then c_j, the coefficient of x^j, is given by

$$c_j = 0 \text{ if } j \leq i - 1. \tag{34}$$

(9) The distance polynomial of a complete graph K_N with N vertices is given by

$$\delta(K_N; x) = (x+1)^{N-1}(x - N + 1). \tag{35}$$

(10) If G is a star with N vertices, then its distance polynomial is given as

$$\delta(G; x) = (x+2)^{N-2}[(x^2 - 1 - (N-2)(2x+1)]. \tag{36}$$

2.3. THE DETERMINANT OF THE DISTANCE MATRIX

The determinant of the distance matrix and the determinant of the adjacency-plus-distance matrix have found use as topological indices [50,52]. Therefore, the properties of the determinant of the distance matrix are of interest and especially the question whether this determinant can be zero and, if so, for which structures.

The determinant of the distance matrix det$|\mathbf{D}|$ can be obtained by expansion of the matrix, but this is an unwieldy procedure. A much faster way is by computing first the distance polynomial and then using the constant term as det$|\mathbf{D}|$.

The determinant of the distance matrix for a tree with N vertices is given by a simple expression [18]:

$$\det|\mathbf{D}| = -(-2)^{N-2}(N-1). \tag{37}$$

Several conclusions follow from this formula:

(1) Isomeric trees necessarily possess the same value of det |**D**|.

(2) Det |**D**| alternates from negative values for trees with even N to positive values for trees with odd N.

(3) No tree has a zero value for det |**D**| but the simplest tree consisting of a single vertex.

In the case of monocycles, it has been found [111] that det |**D**| vanishes for $4n$ (n = integer) cycles.

2.4. THE WIENER NUMBER

The Wiener number $2\text{-}W = 2\text{-}W(G)$ of a graph G is defined by [28]

$$2\text{-}W = \frac{1}{2}\sum_i \sum_j (\mathbf{D})_{ij}. \tag{38}$$

For example, the Wiener numbers corresponding to a tree T from fig. 1 and an edge-weighted graph G_{EW} from fig. 2 are equal to

$$2\text{-}W(T) = 71, \quad 2\text{-}W(G_{EW}) = 85.$$

The Wiener numbers for n-alkanes with N carbon atoms can be obtained from a simple formula [28,39]:

$$2\text{-}W(n\text{-alkane}) = \frac{1}{6}N(N^2-1). \tag{39}$$

Similarly, the Wiener number for cycloalkanes can be computed by means of the following formula [80,112]:

$$2\text{-}W(n\text{-cycloalkane}) = \frac{1}{8}N\{N^2 - \frac{1}{2}[(1-(-1)^N]\}. \tag{40}$$

The Wiener number of a star can also be given in a closed form [39]:

$$2\text{-}W(\text{star}) = (N-1)^2. \tag{41}$$

The Wiener numbers for alkanes with up to 10 carbon atoms are given in table 4.

There are many algorithms for computing the distance matrix in the literature [35,72–80] and consequently the 2-W number for any structure. The Wiener number can also be computed by means of the adjacency matrix of a graph [113]:

Table 4

The Wiener numbers 2-W, 3-W_C and 3-W_{CH} for alkanes with up to two carbon atoms.

	Wiener number		
Alkane	2-W	3-W_C	3-W_{CH}
methane	0	0.000	15.36
ethane	1	1.532	56.88
propane	4	5.607	131.32
2-methylpropane	9	12.196	238.43
butane	10	13.603	251.72
2,2-dimethylpropane	16	21.264	377.74
2-methylbutane	18	23.372	395.20
pentane	20	26.694	425.97
2,2-dimethylbutane	28	35.640	571.17
2,3-dimethylbutane	29	36.945	583.43
2-methylpentane	32	41.036	620.14
3-methylpentane	31	39.726	608.00
hexane	35	46.216	669.74
2,2,3-trimethylbutane	42	52.490	797.28
2,2-dimenthylpentane	46	57.828	844.90
3,3-dimethylpentane	44	55.302	822.11
2,3-dimethylpentane	46	57.807	845.15
2,4-dimethylpentane	48	59.068	853.31
2-methylhexane	52	66.122	919.34
3-methylhexane	50	63.828	900.30
3-ethylpentane	48	57.762	838.30
heptane	56	73.384	989.89
2,2,3,3-tetramethylbutane	58	71.717	1054.03
2,2,3-trimethylpentane	63	77.860	1107.57
2,3,3-trimethylpentane	62	76.691	1097.57
2,2,4-trimethylpentane	66	79.921	1121.20
2,2-dimethylhexane	71	88.454	1200.12
3,3-dimethylhexane	67	80.679	1128.11
3-ethyl-3-methylpentane	64	77.052	1095.95
2,3,4-trimethylpentane	65	79.100	1116.04
2,3-dimethylhexane	70	87.304	1191.73
3-ethyl-2-methylpentane	67	80.317	1124.20
3,4-dimethylhexane	68	85.224	1175.40
2,4-dimethylhexane	71	85.219	1164.58
2,5-dimethylhexane	74	91.060	1219.64
2-methylheptane	79	100.205	1308.98
3-methylheptane	76	96.614	1277.61

. . . continued

Table 4 (continued)

Alkane	Wiener number		
	2-W	3-W_C	3-W_{CH}
4-methylheptane	75	95.497	1268.76
3-ethylhexane	72	86.846	1182.60
octane	84	109.510	1399.79
2,2,3,3-tetramethylpentane	82	100.374	1402.14
2,2,3,4-tetramethylpentane	86	103.723	1429.17
2,2,3-trimethylhexane	92	112.971	1511.19
2,2-dimethyl-3-ethylpentane	88	106.359	1450.39
3,3,4-trimethylhexane	88	108.463	1474.27
2,3,3,4-tetramethylpentane	84	101.733	1413.01
2,3,3-trimethylhexane	90	110.766	1493.25
2,3-dimethyl-3-ethylpentane	86	102.928	1420.38
2,2,4,4-tetramethylpentane	88	105.039	1435.09
2,2,4-trimethylhexane	94	114.216	1518.57
2,4,4-trimethylhexane	92	111.975	1500.13
2,2,5-trimethylhexane	98	118.774	1556.10
2,2-dimethylheptane	104	129.440	1658.81
3,3-dimethylheptane	98	122.411	1598.06
4,4-dimethylheptane	96	120.181	1580.25
3-ethyl-3-methylhexane	92	110.558	1487.49
3,3-diethylpentane	88	104.349	1430.61
2,3,4-trimethylhexane	92	112.167	1503.55
2,4-dimethyl-3-ethylpentane	90	107.346	1456.11
2,3,5-trimethylhexane	96	115.854	1530.93
2,3-dimethylheptane	102	126.890	1637.14
3-ethyl-2-methylhexane	96	114.766	1522.62
3,4-dimethylheptane	98	122.458	1599.99
3-ethyl-4-methylhexane	94	112.557	1504.60
2,4-dimethylheptane	102	125.983	1627.06
4-ethyl-2-methylhexane	98	115.115	1522.37
3,5-dimethylheptane	100	123.588	1605.47
2,5-dimethylheptane	104	128.838	1652.47
2,6-dimethylheptane	108	133.055	1687.31
2-methyloctane	114	144.393	1797.29
3-methyloctane	110	139.694	1757.22
4-methyloctane	108	137.212	1735.40
3-ethylheptane	104	125.580	1619.73
4-ethylheptane	102	123.555	1603.71

. . . continued

Table 4 (continued)

Alkane	Wiener number		
	2-W	3-W_C	3-W_{CH}
nonane	120	155.820	1909.68
2,2,3,3,4-pentamethylpentane	108	129.677	1764.19
2,2,3,3-tetramethylhexane	115	139.852	1852.33
3-ethyl-2,2,3-trimethylpentane	110	139.847	1770.99
3,3,4,4-tetramethylhexane	111	135.626	1818.76
2,2,3,4,4-pentamethylpentane	111	132.221	1782.18
2,2,3,4-tetramethylhexane	118	142.729	1877.27
3-ethyl-2,2,4-trimethylpentane	115	134.809	1801.01
2,3,4,4-tetramethylhexane	116	140.538	1858.26
2,2,3,5-tetramethylhexane	123	146.954	1906.31
2,2,3-trimethylheptane	130	159.492	2026.14
2,2-dimethyl-3-ethylhexane	122	146.342	1905.67
3,3,4-trimethylheptane	123	151.182	1954.57
3,3-dimethyl-4-ethylhexane	118	142.715	1877.96
2,3,3,4-tetramethylhexane	115	139.090	1846.20
3,4,4-trimethylheptane	122	150.067	1945.35
3,4-dimethyl-3-ethylhexane	117	139.514	1847.09
3-ethyl-2,3,4-trimethylpentane	112	132.675	1786.25
2,3,3,5-tetramethylhexane	120	144.147	1884.96
2,3,3-trimethylheptane	127	156.006	1996.23
2,3-dimethyl-3-ethylhexane	119	141.473	1864.18
3,3-diethyl-2-methylpentane	114	133.939	1794.84
2,2,4,4-tetramethylhexane	119	143.069	1987.23
2,2,4,5-tetramethylhexane	124	148.129	1916.57
2,2,4-trimethylheptane	131	159.916	2027.35
2,2-dimethyl-4-ethylhexane	126	146.346	1896.25
3,3,5-trimethylheptane	126	154.021	1976.33
2,4,4-trimethylheptane	127	155.349	1988.87
2,4-dimethyl-4-ethylhexane	122	146.131	1904.27
2,2,5,5-tetramethylhexane	127	151.408	1941.58
2,2,5-trimethylheptane	134	163.044	2052.19
2,5,5-trimethylheptane	131	159.644	2023.18
2,2,6-trimethylheptane	139	168.086	2092.28
2,2-dimethyloctane	146	181.606	2224.40
3,3-dimethyloctance	138	172.426	2146.83
4,4-dimethyloctane	134	167.584	2104.49
3-ethyl-3-methylheptane	129	154.783	1980.23
4-ethyl-4-methylheptane	126	151.621	1954.76

. . . continued

Table 4 (continued)

Alkane	Wiener number 2-W	 3-W_C	 3-W_{CH}
3,3-diethylhexane	121	142.705	1872.36
2,3,4,5-tetramethylhexane	121	146.370	1907.95
2,3,4-trimethylheptane	128	156.365	1997.87
2,3-dimethyl-4-ethylhexane	123	144.134	1882.67
2,4-dimethyl-3-ethylhexane	122	144.730	1890.57
3,4,5-trimethylheptane	125	152.625	1965.00
2,4-dimethyl-3-isopropylpentane	117	138.256	1833.36
3-isopropyl-2-methylhexane	124	146.588	1905.28
2,3,5-trimethylheptane	131	158.829	2015.02
2,5-dimethyl-3-ethylhexane	127	147.984	1912.26
2,4,5-trimethylheptane	130	158.009	2009.28
2,3,6-trimethylheptane	136	164.668	2064.61
2,3-dimethyloctane	143	177.743	2191.63
3-ethyl-2-methylheptane	134	160.204	2026.81
3,4-dimethyloctane	137	171.126	2136.17
4-isopropylheptane	131	156.578	1997.73
4-ethyl-3-methylheptane	129	154.748	1981.17
4,5-dimethyloctane	135	168.681	2114.29
3-ethyl-4-methylheptane	130	155.372	1984.91
3,4-diethylhexane	125	145.646	1894.07
2,4,5-trimethylheptane	135	162.214	2039.83
2,4-dimethyloctane	142	175.753	2171.92
4-ethyl-2-methylheptane	134	158.719	2012.39
3,5-dimethyloctane	138	171.126	2132.56
3-ethyl-5-methylheptane	133	156.967	1994.81
2,5-dimethyloctane	143	177.291	2186.04
5-ethyl-2-methylheptane	138	163.675	2053.96
3,6-dimethyloctane	141	175.262	2169.96
2,6-dimethyloctane	146	180.599	2213.68
2,7-dimethyloctane	151	186.839	2268.53
2-methylnonane	158	200.100	2398.37
3-methylnonane	153	194.111	2346.10
4-methylnonane	150	190.381	2313.66
3-ethyloctane	145	175.850	2170.31
5-methylnonane	149	189.147	2302.73
4-ethyloctane	141	171.243	2130.45
4-propylheptane	138	162.527	2050.32
decane	165	213.635	2531.95

$$2\text{-}W = \frac{1}{2}\sum_{i=1}^{N}\sum_{l=1}^{l_{\max}} l(\mathbf{A}_l^2)_{ii}, \tag{42}$$

where l is the length of the shortest path between two vertices in G and $\mathbf{A}_l$ are the corresponding higher-order adjacency matrices.

The Wiener number is not a particularly discriminating index. Already in the heptane family appear two pairs of C_7 trees with the same Wiener numbers [39]. These trees are shown in fig. 7.

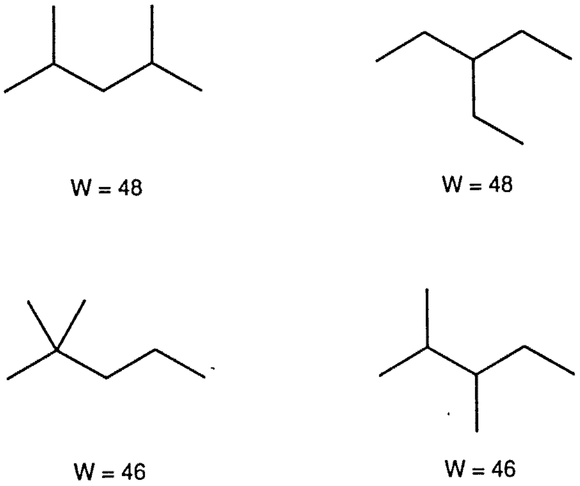


Fig. 7. Two pairs of heptane trees with identical Wiener numbers.

The Wiener number appears to be a good measure of the compactness of a molecule [39]: The smaller the Wiener number, the larger the compactness of a structure (in terms of structural features such as branching and cyclicity). Hence, it can be reliably used for correlations with those physical and chemical properties which depend on the ratio of the volume to the surface of the molecule. Chromatographic retention data for a homologous series of hydrocarbons are typical molecular properties for which the quantitative structure–property relationships (QSPR) with the Wiener number should yield trustworthy predictions [54, 114–117]. For example, the following relationship between the experimental gas-chromatographic retention indices I measured on squalane at 333 K [118] and at 373 K [119] and the 2-D Wiener numbers of the first 157 alkanes,

$$I = a(2\text{-}W)^b + c, \tag{43}$$

possesses very good statistical characteristics [54]: $a = 171.2$ (± 15.7), $b = 0.3347$ (± 0.0128), $c = 48.6$ (± 27.3), $r = 0.984$, $s = 33.0$ and $F = 2403$.

3. The geometric-distance matrix and the corresponding invariants

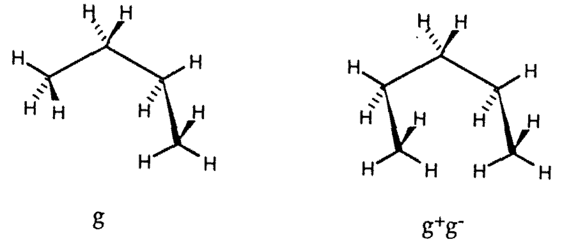
Most of the definitions given in the preceding section for the graph-theoretical distance matrix and the related topological invariants also apply in the modified form to the geometric-distance matrix and its invariants.

3.1. THE GEOMETRIC-DISTANCE MATRIX

While the graph-theoretical distance matrix 2-**D** captures the constitutional features of the molecule, its three-dimensional (3D) character is encoded in the geometric-distance matrix [5,6,11,13,44]. This matrix will be denoted here by 3-**D**.

The geometric-distance matrix $3\text{-}\mathbf{D} = 3\text{-}\mathbf{D}(G)$ of a molecular structure G is a real symmetric $N \times N$ matrix defined in the same way as the 2-**D** matrix (see (1)). However, the matrix elements $(\mathbf{D})_{ij}$ now represent the shortest Cartesian distances l_{ij} (in some arbitrary units of length) between sites i and j in G. The geometric-distance matrix (also called the topographic matrix [44]) can be constructed from the known geometry of a molecule. However, for many molecules their geometries are unknown. Therefore, one must approximate the geometry of a molecule in some way. This can be done, for example, by quantum mechanical computations [13,120] or by molecular mechanics computations [54,121,122] or by constructing idealized "frozen" 3D geometry of a molecule using standard bond lengths and bond angles [44,48]. It appears that at present the most convenient method to optimize the geometry of an arbitrary molecule is the molecular mechanics method [123]. All results reported here will be based on the molecular mechanics (MM) computations.

As the input data for the MM computations, the intuitively most stable conformation of a particular molecule was selected. This conformation was constructed from data for the standard bond lengths and bond angles. For example, in the case of alkanes it is possible to consider all conformations for smaller and simpler members of the series. In the case of larger and more complex alkanes, this cannot be done because the number of possible conformations increases enormously [124]. However, all conformations need not be considered. It is sufficient to take into account only those conformations which are present in the equilibrium mixtures in reasonable amounts, probably greater than a few percent. A simple estimate reveals that if one gauche-butane (g) interaction contributes to the enthalpy ≈3.4 kJ/mol, then it is not necessary to consider conformations with more than 3 or 4 gauche-butane interactions relative to the minimum energy conformation. The minimum energy conformation of an alkane is taken to be the most extended conformation with the minimum number of gauche-butane interactions, excluding the "forbidden" pentane (g^+g^-) interactions.

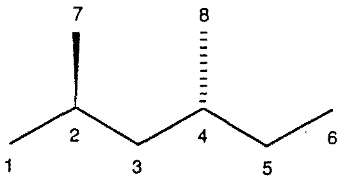


Scheme 1.

These input geometries were then refined via molecular mechanics computations, giving the chosen conformation optimum bond lengths, bond angles and interatomic distance. The alkane geometry obtained in this way was used to set up the corresponding geometric-distance matrix. The elements of 3-**D**, i.e. the distances between atoms in the alkane $(3\text{-}\mathbf{D})_{ij}$, were computed from the Cartesian coordinates, which are part of the standard output of the MM programs:

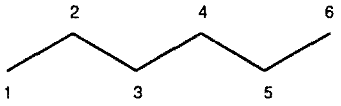
$$(3\text{-}\mathbf{D})_{ij} = [(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2]^{1/2}. \tag{44}$$

As an example, the geometric-distance matrix of the carbon skeleton of 2,4-dimethylhexane is given in fig. 8.



$$3\text{-}\mathbf{D} = \begin{bmatrix} 0.0000 & 1.5405 & 2.5256 & 3.9526 & 5.0748 & 5.4288 & 2.5165 & 4.5071 \\ 1.5405 & 0.0000 & 1.5429 & 2.6274 & 3.9754 & 4.5721 & 1.5377 & 3.0959 \\ 2.5256 & 1.5429 & 0.0000 & 1.5433 & 2.5563 & 3.1195 & 2.5599 & 2.5537 \\ 3.9526 & 2.6274 & 1.5433 & 0.0000 & 1.5454 & 2.5887 & 3.1335 & 1.5401 \\ 5.0748 & 3.9754 & 2.5563 & 1.5454 & 0.0000 & 1.5352 & 4.5670 & 2.5140 \\ 5.4288 & 4.5721 & 3.1195 & 2.5887 & 1.5352 & 0.0000 & 4.9832 & 3.9131 \\ 2.5165 & 1.5377 & 2.5599 & 3.1335 & 4.5670 & 4.9832 & 0.0000 & 3.6689 \\ 4.5071 & 3.0959 & 2.5537 & 1.5401 & 2.5140 & 3.9131 & 3.6689 & 0.0000 \end{bmatrix}$$

Fig. 8. The geometric-distance matrix corresponding to the carbon skeleton of 2,4-dimethylhexane.

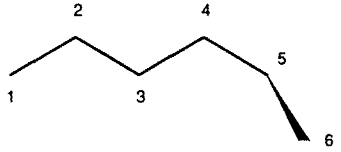


$$\text{3-D (aaa)} = \begin{bmatrix} 0.000 & 1.534 & 2.544 & 3.914 & 5.092 & 6.420 \\ 1.534 & 0.000 & 1.537 & 2.548 & 3.918 & 5.092 \\ 2.544 & 1.537 & 0.000 & 1.537 & 2.548 & 3.914 \\ 3.914 & 2.548 & 1.537 & 0.000 & 1.537 & 2.544 \\ 5.092 & 3.918 & 2.548 & 1.537 & 0.000 & 1.534 \\ 6.420 & 5.092 & 3.914 & 2.544 & 1.534 & 0.000 \end{bmatrix}$$

$$\delta(\text{aaa}; x) = x^6 - 176.80\,x^4 - 1065.35\,x^3 - 2338.49\,x^2 - 2209.57\,x - 761.64$$

spectrum: $\{ 15.9216, -1.0561, -1.1600, -1.4022, -2.9899, -9.3133 \}$

$$\text{3-W(aaa)} = 46.216$$



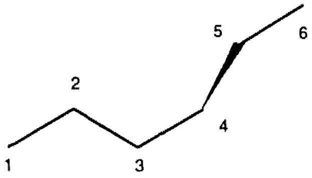
$$\text{3-D (aag)} = \begin{bmatrix} 0.000 & 1.535 & 2.544 & 3.914 & 5.117 & 5.555 \\ 1.535 & 0.000 & 1.538 & 2.548 & 3.934 & 4.586 \\ 2.544 & 1.538 & 0.000 & 1.537 & 2.574 & 3.161 \\ 3.914 & 2.548 & 1.537 & 0.000 & 1.538 & 2.571 \\ 5.117 & 3.934 & 2.574 & 1.538 & 0.000 & 1.534 \\ 5.555 & 4.586 & 3.161 & 2.571 & 1.534 & 0.000 \end{bmatrix}$$

$$\delta(\text{aag}; x) = x^6 - 156.86\,x^4 - 938.50\,x^3 - 2067.51\,x^2 - 1963.81\,x - 679.41$$

spectrum: $\{ 15.1184, -1.0254, -1.1493, -1.5317, -2.9382, -8.4739 \}$

$$\text{3-W(aag)} = 44.186$$

Fig. 9. Caption follows.

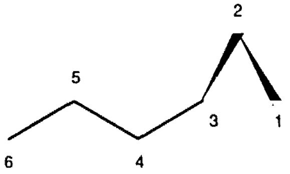


$$3\text{-D}\,(aga) = \begin{bmatrix} 0.000 & 1.535 & 2.543 & 3.930 & 4.594 & 6.070 \\ 1.535 & 0.000 & 1.537 & 2.576 & 3.168 & 4.595 \\ 2.543 & 1.537 & 0.000 & 1.539 & 2.576 & 3.930 \\ 3.930 & 2.576 & 1.539 & 0.000 & 1.537 & 2.543 \\ 4.594 & 3.168 & 2.576 & 1.537 & 0.000 & 1.535 \\ 6.070 & 4.595 & 3.930 & 2.543 & 1.535 & 0.000 \end{bmatrix}$$

$$\delta(aga; x) = x^6 - 158.01\,x^4 - 956.96\,x^3 - 2125.46\,x^2 - 2037.24\,x - 712.43$$

spectrum: { 15.1923, -1.0224, -1.3163, -1.3449, -3.0703, -8.4383 }

$$3\text{-W}(aga) = 44.209$$



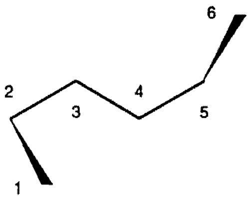
$$3\text{-D}\,(ag^+g^+) = \begin{bmatrix} 0.000 & 1.534 & 2.570 & 3.134 & 3.768 & 4.946 \\ 1.534 & 0.000 & 1.538 & 2.603 & 3.148 & 4.595 \\ 2.570 & 1.538 & 0.000 & 1.538 & 2.575 & 3.930 \\ 3.134 & 2.603 & 1.538 & 0.000 & 1.537 & 2.543 \\ 3.768 & 3.148 & 2.575 & 1.537 & 0.000 & 1.535 \\ 4.946 & 4.595 & 3.930 & 2.543 & 1.535 & 0.000 \end{bmatrix}$$

$$\delta(ag^+g^+; x) = x^6 - 133.24\,x^4 - 799.45\,x^3 - 1784.31\,x^2 - 1718.44\,x - 598.62$$

spectrum: { 14.1217, -0.9240, -1.3374, -1.5472, -3.0545, -7.2586 }

$$3\text{-W}(ag^+g^+) = 41.495$$

Fig. 9 (continued). Caption follows.

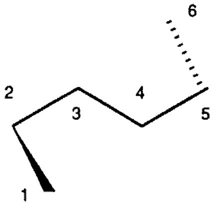


$$\text{3-D}\,(g^+ag^+) = \begin{bmatrix} 0.000 & 1.534 & 2.569 & 3.145 & 4.608 & 5.082 \\ 1.534 & 0.000 & 1.539 & 2.573 & 3.945 & 4.608 \\ 2.569 & 1.539 & 0.000 & 1.537 & 2.573 & 3.144 \\ 3.145 & 2.573 & 1.537 & 0.000 & 1.539 & 2.569 \\ 4.608 & 3.945 & 2.573 & 1.539 & 0.000 & 1.534 \\ 5.082 & 4.608 & 3.144 & 2.569 & 1.534 & 0.000 \end{bmatrix}$$

$$\delta(g^+ag^+; x) = x^6 - 141.89\,x^4 - 841.01\,x^3 - 1864.48\,x^2 - 1783.06\,x - 617.21$$

spectrum: $\{ 14.4740, -0.9439, -1.1845, -1.7866, -2.7250, -7.8340 \}$

$$\text{3-W}(g^+ag^+) = 42.500$$



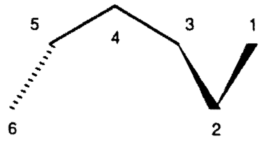
$$\text{3-D}\,(g^+ag^-) = \begin{bmatrix} 0.000 & 1.534 & 2.574 & 3.173 & 4.582 & 5.574 \\ 1.534 & 0.000 & 1.539 & 2.574 & 3.953 & 4.581 \\ 2.574 & 1.539 & 0.000 & 1.537 & 2.574 & 3.180 \\ 3.173 & 2.574 & 1.537 & 0.000 & 1.539 & 2.574 \\ 4.582 & 3.953 & 2.574 & 1.539 & 0.000 & 1.534 \\ 5.574 & 4.581 & 3.180 & 2.574 & 1.534 & 0.000 \end{bmatrix}$$

$$\delta(g^+ag^-; x) = x^6 - 147.18\,x^4 - 873.03\,x^3 - 1943.41\,x^2 - 1866.68\,x - 650.32$$

spectrum: $\{ 14.7021, -1.0140, -1.0830, -1.9451, -2.5549, -8.1051 \}$

$$\text{3-W}(g^+ag^-) = 43.024$$

Fig. 9 (continued). Caption follows.



$$3\text{-D}(g^-g^-g^-) = \begin{bmatrix} 0.000 & 1.534 & 2.569 & 3.138 & 3.743 & 4.763 \\ 1.534 & 0.000 & 1.538 & 2.601 & 3.149 & 3.743 \\ 2.569 & 1.538 & 0.000 & 1.538 & 2.601 & 3.138 \\ 3.138 & 2.601 & 1.538 & 0.000 & 1.538 & 2.569 \\ 3.743 & 3.149 & 2.601 & 1.538 & 0.000 & 1.534 \\ 4.763 & 3.743 & 3.138 & 2.569 & 1.534 & 0.000 \end{bmatrix}$$

$$\delta(g^-g^-g^-; x) = x^6 - 118.85\,x^4 - 706.57\,x^3 - 1590.45\,x^2 - 1541.83\,x - 537.09$$

spectrum: $\{ 13.4477, -0.9013, -1.2112, -2.0074, -2.7861, -6.5417 \}$

$$3\text{-W}(g^-g^-g^-) = 39.697$$

Fig. 9. The geometric-distance matrices and their invariants for the seven energy-lowest conformations of *n*-hexane.

It should be noted that the geometric-distance matrix determines uniquely a given structure to the extent that the corresponding conformation is unique to this structure. Therefore, the geometric-distance and consequently the related topographic invariants are different for different conformations of the molecule. To illustrate this statement, in fig. 9 are given the geometric-distance matrices and their invariants for the seven energy-lowest conformations out of 12 possible conformations of *n*-hexane.

3.2. THE GEOMETRIC-DISTANCE POLYNOMIAL AND THE GEOMETRIC-DISTANCE SPECTRUM

The geometric-distance (topographic) polynomial is defined in the same way as the graph-theoretical distance polynomial. Similarly, it is also computed using the modified form of the Le Verrier–Faddeev–Frame method [122]. As an example, a computing of the geometric-distance polynomial for butane is shown in table 5. The geometric-distance matrix is also diagonalized by the Householder-QL algorithm combination. The geometric-distance spectrum of butane is given by $\{6.9704, -1.1655, -1.5239, -4.2810\}$.

In table 6 and table 7, the geometric-distance polynomials and spectra, respectively, of the lowest alkanes are given.

The geometric-distance polynomials possess many interesting properties:

(1) They can differentiate isospectral molecules.

Table 5

The computation of the geometric-distance polynomial of butane by means of the modified Le Verrier–Faddeev–Frame method.

(i) Geometric-distance spectrum of butane

$\{6.97, -1.17, -1.52, -4.28\}$

(ii) $\sum_i (\mathbf{D})_{ii} = \sum_i (\mathbf{D}_1)_{ii} = 0$

$c_1 = 0$

(iii) $\{(\mathbf{B}_1)_{ii} = (\mathbf{D}_1)_{ii} - (c_1\mathbf{I})_{ii}\}_{i=1,\dots,4} = \{6.97, -1.17, -1.52, -4.28\}$

$\{(\mathbf{D}_2)_{ii} = (\mathbf{D})_{ii}(\mathbf{B}_1)_{ii}\}_{i=1,\dots,4} = \{48.59, 1.36, 2.32, 18.33\}$

$c_2 = \frac{1}{2}\sum_i (\mathbf{D}_2)_{ii} = 35.30$

(iv) $\{(\mathbf{B}_2)_{ii} = (\mathbf{D}_2)_{ii} - (c_2\mathbf{I})_{ii}\}_{i=1,\dots,4} = \{13.29, -33.94, -32.98, -16.97\}$

$\{(\mathbf{D}_3)_{ii} = (\mathbf{D})_{ii}(\mathbf{B}_2)_{ii}\}_{i=1,\dots,4} = \{92.63, 39.56, 50.26, 72.65\}$

$c_3 = \frac{1}{3}\sum_i (\mathbf{D}_3)_{ii} = 85.03$

(v) $\{(\mathbf{B}_3)_{ii} = (\mathbf{D}_3)_{ii} - (c_3\mathbf{I})_{ii}\}_{i=1,\dots,4} = \{7.60, -45.47, -34.77, -12.38\}$

$\{(\mathbf{D}_4)_{ii} = (\mathbf{D})_{ii}(\mathbf{B}_3)_{ii}\}_{i=1,\dots,4} = \{53.00, 53.00, 53.00, 53.00\}$

$c_4 = \frac{1}{4}\sum_i (\mathbf{D}_4)_{ii} = 53.00$

(vi) $\{(\mathbf{B}_4)_{ii} = (\mathbf{D}_4)_{ii} - (c_4\mathbf{I})_{ii}\}_{i=1,\dots,4} = \{0, 0, 0, 0\}$

(vii) The geometric-distance polynomial of butane

$\delta(\text{butane}; x) = x^4 - 35.30x^2 - 85.03x - 53.00$

For a pair of isospectral trees T_1 and T_2 from fig. 4 are given their geometric distance matrices, geometric-distance polynomials, geometric-distance spectra and three-dimensional Wiener numbers in table 8. All these quantities are different for each of these two trees.

(2) Two non-isomorphic molecules cannot possess identical geometric-distance polynomials unless they possess identical optimum conformations.

(3) Since different conformations of a molecule have different geometric-distance matrices, they consequently also possess different polynomials (and spectra). (see examples in fig. 8)

(4) The coefficient c_0 and c_1 at x^N and x^{N-1} are, respectively, equal to 1 and 0, as was the case for the graph-theoretical distance polynomial.

The coefficient c_1 is equal to zero because relationship (28) holds also for the elements of the geometric-distance spectrum.

(5) Similarly, the coefficient c_2 at x^{N-2} is equal to the half-sum of squares of elements of the geometric-distance matrix:

Table 6

The geometric-distance polynomials of the lowest alkanes.

	Coefficients of the polynomial								
Alkane	c_0	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8
propane	1	0	-11.154	-11.95					
2-methylpropane	1	0	-26.262	-68.16	-45.3				
butane	1	0	-35.297	-85.03	-53.0				
2,2-dimethylpropane	1	0	-47.500	-199.23	-300.8	-151.4			
2-methylbutane	1	0	-60.462	-253.04	-360.0	-166.7			
pentane	1	0	-85.388	-346.52	-466.4	-205.9			
2,2-dimethylbutane	1	0	-91.882	-546.50	-1284.1	-1332.4	-499.3		
2,3-dimethylbutane	1	0	-100.311	-599.41	-1374.0	-1378.5	-502.6		
2-methylpentane	1	0	-131.363	-790.06	-1801.3	-1775.3	-633.8		
3-methylpentane	1	0	-121.364	-733.40	-1680.4	-1657.5	-590.0		
hexane	1	0	-176.801	-1065.35	-2338.5	-2209.6	-761.6		
2,2,3-trimethlybutane	1	0	-142.387	-1109.38	-3635.6	-5948.6	-4747.7	-1455.0	
2,2-dimethylpentane	1	0	-183.166	-1458.49	-4798.5	-7808.2	-6154.5	-1864.9	
3,3-dimethylpentane	1	0	-164.005	-1302.34	-4294.3	-6990.2	-5478.2	-1638.5	
2,3-dimethylpentane	1	0	-181.596	-1458.92	-4748.9	-7538.9	-5771.6	-1702.0	
2,4-dimethylpentane	1	0	-190.868	-1522.32	-4990.0	-8062.3	-6296.5	-1890.2	
2-methylhexane	1	0	-253.653	-2053.91	-6627.1	-10345.0	-7782.4	-2263.2	
3-methylhexane	1	0	-233.675	-1917.10	-6251.5	-9830.8	-7323.7	-2159.8	
3-ethylpentane	1	0	-179.080	-1477.89	-4909.3	-7878.4	-6037.5	-1764.7	
heptane	1	0	-326.644	-2716.99	-8597.4	-13007.2	-9487.8	-2687.0	
2,2,3,3-tetramethylbutane	1	0	-198.470	-1909.74	-8089.4	-18346.3	-23011.0	-14965.3	-3906.48
2,2,3-trimethylpentane	1	0	-243.866	-2438.63	-10522.4	-23998.1	-29950.4	-19215.6	-4930.08
2,3,3-trimethylpentane	1	0	-235.154	-2345.07	-10116.0	-23063.9	-28730.7	-18352.5	-4673.85
2,2,4-trimethylpentane	1	0	-259.393	-2570.98	-10981.3	-24846.2	-30853.9	-19747.4	-5070.54
2,2-dimethylhexane	1	0	-336.122	-3416.96	-14697.2	-33105.8	-40626.9	-25638.9	-6499.82
3,3-dimethylhexane	1	0	-264.032	-2729.81	-11977.7	-27395.5	-33955.0	-21533.7	-5458.13
3-ethyl-3-methylpentane	1	0	-235.538	-2401.15	-10513.0	-24071.8	-29834.8	-18865.1	-4750.49
2,3,4-trimethylpentane	1	0	-251.713	-2531.52	-10924.3	-24824.5	-30794.6	-19607.3	-4989.30
2,3-dimethylhexane	1	0	-323.059	-3356.55	-14497.6	-32286.7	-38870.8	-24016.5	-5968.48
3-ethyl-2-methylpentane	1	0	-259.556	-2680.47	-11666.0	-26384.2	-32318.1	-20258.5	-5076.28
3,4-dimethylhexane	1	0	-305.864	-3157.05	-13655.4	-30498.7	-36781.5	-22713.5	-5627.33
2,4-dimethylhexane	1	0	-300.331	-3110.47	-13549.6	-30664.3	-37567.3	-23560.3	-5916.72
2,5-dimethylhexane	1	0	-355.896	-3627.52	-15559.8	-34816.5	-42324.0	-26411.1	-6615.76
2-methylheptane	1	0	-451.316	-4747.84	-20283.0	-44519.9	-53797.9	-32188.8	-7914.79
3-methylheptane	1	0	-415.245	-4395.88	-18994.1	-41969.5	-49960.5	-30497.5	-7493.75
4-methylheptane	1	0	-403.220	-4341.69	-18904.1	-41948.2	-50085.3	-30644.7	-7542.73
3-ethylhexane	1	0	-316.259	-3373.85	-14929.1	-33810.3	-41075.0	-25419.2	-6281.49
octane	1	0	-556.748	-6090.57	-25969.8	-55848.4	-64624.8	-38478.3	-9267.11

Table 7

The geometric-distance spectra corresponding to the alkane geometric-distance polynomials shown in table 6.

	Geometric-distance spectrum							
Alkane	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
propane	3.7832	−1.2441	−2.5391					
2-methylpropane	6.1996	−1.1434	−2.5281	−2.5281				
butane	6.9704	−1.1655	−1.5239	−4.2810				
2,2-dimethlypropane	8.6483	−1.0985	−2.5164	−2.5165	−2.5168			
2-methylbutane	9.5454	−1.0101	−1-5251	−2.5326	−4.4776			
pentane	10.9961	−1.1209	−1-1711	−2.1892	−6.5148			
2,2-dimethylbutane	12.1062	−0.9206	−1.5272	−2.5154	−2.5263	−4.6168		
2,3-dimethylbutane	12.5504	−0.9533	−1.5171	−1.6063	−3.3972	−5.0775		
2-methylpentane	14.0492	−1.0074	−1.1309	−2.1449	−2.5632	−7.2028		
3-methylpentane	13.6028	−0.9914	−1.1159	−2.1402	−2.7905	−6.5648		
hexane	15.9216	−1.0561	−1.1600	−1.4022	−2.9899	−9.3133		
2,2,3-trimethylbutane	15.2492	−0.8559	−1.5190	−1.6068	−2.5167	−3.3772	−5.3735	
2,2-dimethylpentane	16.9571	−0.9182	−1.1318	−2.1319	−2.5155	−2.5645	−7.6952	
3,3-dimethlypentane	16.2011	−0.8724	−1.1098	−2.1355	−2.5124	−2.9326	−6.6383	
2,3-dimethylpentane	16.9139	−0.9451	−1.0955	−1.5911	−2.2513	−3.7018	−7.3290	
2,4-dimethylpentane	17.2492	−0.9361	−1.0752	−2.1375	−2.2786	−2.7794	−8.0424	
2-methylhexane	19.4485	−1.0015	−1.1002	−1.3073	−2.5614	−3.0141	−10.4640	
3-methylhexane	18.8318	−0.9816	−1.0664	−1.3837	−2.3873	−3.4787	−9.5340	
3-ethylpentane	16.8955	−0.8727	−1.1150	−1.5906	−2.3230	−4.4172	−6.5771	
heptane	21.7105	−1.0250	−1.1104	−1.3128	−1.6389	−4.0059	−12.6176	
2,2,3,3-tetramethylbutane	18.1965	−0.8256	−1.5150	−1.5966	−1.5966	−3.3741	−3.3741	−5.9146
2,2,3-trimethylpentane	19.9044	−0.8532	−1.1040	−1.5836	−2.2041	−2.5586	−3.7509	−7.8499
2,3,3-trimethylpentane	19.6049	−0.8300	−1.0912	−1.5913	−2.1884	−2.6799	−3.7977	−7.4263
2,2,4-trimethylpentane	20.3752	−0.9106	−1.0100	−1.7455	−2.1413	−2.5521	−3.2280	−8.7877
2,2-dimethylhexane	22.7522	−0.9179	−1.0919	−1.2870	−2.5155	−2.5642	−3.0251	−11.3506
3,3-dimethylhexane	20.6735	−0.8690	−1.1283	−1.3092	−2.1655	−2.5697	−4.6052	−8.0267
3-ethyl-3-methylpentane	19.6954	−0.8415	−1.1102	−1.3111	−2.1412	−2.9372	−4.7191	−6.6352
2,3,4-trimethylpentane	20.1793	−0.8554	−1.0673	−1.5910	−2.1741	−2.4508	−3.9470	−8.0937
2,3-dimethylhexane	22.4563	−0.9449	−1.0648	−1.2927	−1.6368	−2.8256	−4.2192	−10.4723
3-ethy-2-methylpentane	20.5212	−0.8585	−1.0895	−1.3866	−2.0525	−2.3550	−5.2202	−7.5588
3,4-dimethylhexane	21.9365	−0.9377	−1.0064	−1.3784	−1.5931	−3.3342	−3.7285	−9.9583
2,4-dimethylhexane	21.7890	−0.9009	−1.0676	−1.2834	−2.1750	−2.5083	−4.1596	−9.6943
2,5-dimethylhexane	23.3045	−0.9337	−1.0294	−1.2751	−2.2911	−2.8035	−3.0167	−11.9551
2-methylheptane	25.8407	−0.9737	−1.0988	−1.1815	−1.6219	−2.5700	−4.0535	−14.3413
3-methylheptane	24.9881	−0.9728	−1.0629	−1.1946	−1.6111	−2.6949	−4.2292	−13.2227
4-methylheptane	24.7474	−0.9789	−1.0206	−1.2598	−1.6398	−2.4291	−4.8291	−12.5901
3-ethylhexane	22.3642	−0.8727	−1.0579	−1.3716	−1.5902	−3.0843	−4.6388	−9.7487
octane	28.3908	−1.0206	−1.0425	−1.3052	−1.3595	−2.0304	−5.1739	−16.4587

Table 8

The geometric-distance matrices and the corresponding topographic invariants for a pair of isospectral trees T_1 and T_2 from fig. 4.

(i) T_1

$$\mathbf{D}(T_1) = \begin{bmatrix}
0.000 & 1.542 & 2.577 & 3.948 & 5.184 & 6.474 & 7.727 & 9.008 & 2.495 & 3.027 \\
1.542 & 0.000 & 1.554 & 2.567 & 3.971 & 5.112 & 6.466 & 7.654 & 1.542 & 2.558 \\
2.577 & 1.554 & 0.000 & 1.547 & 2.608 & 3.956 & 5.154 & 6.468 & 2.578 & 1.541 \\
3.948 & 2.567 & 1.547 & 0.000 & 1.539 & 2.551 & 3.921 & 5.095 & 3.045 & 2.540 \\
5.184 & 3.971 & 2.608 & 1.539 & 0.000 & 1.539 & 2.550 & 3.916 & 4.490 & 3.043 \\
6.474 & 5.112 & 3.956 & 2.551 & 1.539 & 0.000 & 1.537 & 2.544 & 5.323 & 4.519 \\
7.727 & 6.466 & 5.154 & 3.921 & 2.550 & 1.537 & 0.000 & 1.535 & 6.766 & 5.447 \\
9.008 & 7.654 & 6.468 & 5.095 & 3.916 & 2.544 & 1.535 & 0.000 & 7.769 & 6.876 \\
2.495 & 1.542 & 2.578 & 3.045 & 4.490 & 5.323 & 6.766 & 7.769 & 0.000 & 3.940 \\
3.027 & 2.558 & 1.541 & 2.540 & 3.043 & 4.519 & 5.447 & 6.876 & 3.940 & 0.000
\end{bmatrix}$$

$$\delta(T_1;x) = x^{10} - 890.23710x^8 - 14050.002x^7 - 94413.665x^6 - 346165.72x^5 - 757079.31x^4 - 1015388.6x^3 - 819381.17x^2 - 365316.45x - 69157.02$$

spectrum: $\{36.742, -0.9354, -1.0322, -1.1136, -1.3271, -1.5399, -2.1225, -3.6744, -5.6886, -19.308\}$

$3\text{-}W(T_1) = 177.74$

(ii) T_2

$$\mathbf{D}(T_2) = \begin{bmatrix}
0.000 & 1.536 & 2.588 & 3.921 & 5.214 & 6.440 & 7.781 & 8.973 & 3.129 & 5.509 \\
1.536 & 0.000 & 1.546 & 2.526 & 3.962 & 5.039 & 6.447 & 7.550 & 2.547 & 4.524 \\
2.588 & 1.546 & 0.000 & 1.546 & 2.641 & 3.961 & 5.221 & 6.481 & 1.538 & 3.126 \\
3.921 & 2.526 & 1.546 & 0.000 & 1.546 & 2.525 & 3.926 & 5.053 & 2.552 & 2.552 \\
5.214 & 3.962 & 2.641 & 1.546 & 0.000 & 1.546 & 2.594 & 3.944 & 3.130 & 1.538 \\
6.440 & 5.039 & 3.961 & 2.525 & 1.546 & 0.000 & 1.538 & 2.544 & 4.528 & 2.550 \\
7.781 & 6.447 & 5.221 & 3.926 & 2.594 & 1.538 & 0.000 & 1.535 & 5.528 & 3.130 \\
8.973 & 7.550 & 6.481 & 5.053 & 3.944 & 2.544 & 1.535 & 0.000 & 6.855 & 4.568 \\
3.129 & 2.547 & 1.538 & 2.552 & 3.130 & 4.528 & 5.528 & 6.855 & 0.000 & 3.700 \\
5.509 & 4.524 & 3.126 & 2.552 & 1.538 & 2.550 & 3.130 & 4.568 & 3.700 & 0.000
\end{bmatrix}$$

$$\delta(T_2;x) = x^{10} - 815.77130x^8 - 12909.212x^7 - 87555.643x^6 - 325954.92x^5 - 725877.12x^4 - 991173.44x^3 - 812618.11x^2 - 366912.23x - 70084.37$$

spectrum: $\{35.458, -0.9061, -0.9912, -1.0954, -1.3436, -1.8074, -2.6972, -2.8857, -5.9918, -17.740\}$

$3\text{-}W(T_2) = 171.13$

$$c_2 = \frac{1}{2}\sum_i \sum_j (3-\mathbf{D}^2)_{ij}. \tag{45}$$

(6) The last coefficient c_N of a geometric-distance polynomial for a tree T is also equal to the determinant of the geometric-distance matrix:

$$c_N = (-1)^N \det |3-\mathbf{D})|. \tag{46}$$

The geometric-distance spectra exhibit many of the properties summarized in the preceding section for the distance spectra of graphs, so they will not be repeated here.

This striking similarity between the graph-theoretical distance matrix and the geometric-distance matrix and their invariants may perhaps be explained by exploring the idea that the 3D structure of a molecule may be represented by the edge-weighted complete graph [122] (the 2D model of a molecule). The edge-weights in such a graph may correspond to geometric distances between the atoms in a molecule. For example, the anti,anti-conformation (aa) of n-pentane may be depicted by the edge-weighted graph G_{EW} shown in fig. 10. Thus, the graph-theoretical distance matrix G_{EW} is identical to the geometric-distance matrix of the carbon skeleton of the n-pentane (aa) conformation.

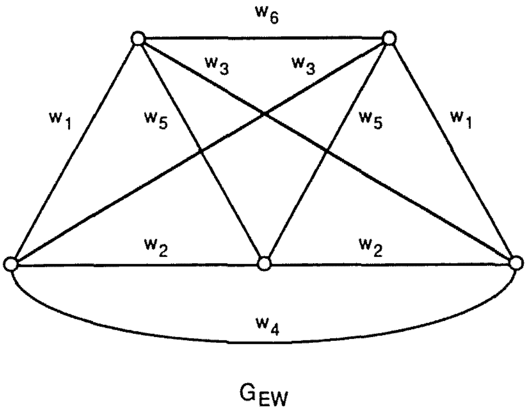


Fig. 10. The edge-weighted graph G_{EW} corresponding to the anti,anti (aa) conformation of n-pentane. w_i's stand for the edge-weights which are equal to the geometric distances between the carbon atoms in the (aa) conformation of n-pentane. Because of the symmetry, there are only six different weights (different geometric distances): w_1 (1.534 Å), w_2 (2.544 Å), w_3 (3.914 Å), w_4 (5.087 Å), w_5 (1.537 Å) and w_6 (2.548 Å).

Since most of the above graph-theoretical results apply also to weighted graphs, they likewise apply to the geometric-distance matrix and its invariants.

3.3. THE DETERMINANT OF THE GEOMETRIC-DISTANCE MATRIX

No work has so far been reported in which any use of the determinant of the geometric-distance matrix det |3-**D**| is mentioned. Therefore, the only use that one can see for det |3-**D**| is for checking the constant term in the geometric-distance polynomial of a tree (see (46)). In addition, det |3-**D**| should also be different for different structures, be they isospectral molecules or conformers of a given molecule.

3.4. THE 3D WIENER NUMBER

The 3D Wiener number, denoted as 3-W, is defined as the half-sum of the elements of the geometric-distance matrix [48]. The half-sum of the elements of the geometric-distance matrix is named the 3D Wiener number because of its similarity with the definition of the 2D Wiener number (see (38)).

The 3D Wiener numbers for the carbon skeletons (3-W_C) and the whole CH skeletons (3-W_{CH}) of alkanes with up to 10 carbon atoms are also given in table 4.

The 3D Wiener number appears to be a very discriminating index. So far, a pair of non-isomorphic structures with the same 3-W number has not been detected. However, such a case may emerge because there is a considerable loss of information on going from the geometric-distance matrix to the Wiener number. For example, a given structure may uniquely be reconstructed from its geometric-distance matrix, but not on the basis of its Wiener number.

The 3D Wiener number also differentiates between the conformations of a molecule (see fig. 9). It should be noted that the 3D Wiener number changes with the change in the structure of each conformation, being the largest for the most extended conformation and the smallest for the most compact conformation. The similar decrease in the Wiener number can be observed on going from a linear isomer to a branched isomer. Therefore, the 3-W number decreases with increasing spheroidicity of a molecule. This parallels the observation for the 2D Wiener number, which decreases with increasing branching (and/or cyclicity) of a structure [39]. Hence, it appears that the 3D Wiener number may be a convenient index to model the shape of a molecule. Thus, it is expected to be the most useful in the correlations with shape-dependent physical, chemical or biological properties.

For illustrative purposes, here will be given the quantitative structure–chromatographic retention relationships (QSCRR) with 3-W_C and 3-W_{CH} indices, respectively. The same algebraic formula (43) as before and the same set of experimental gas-chromatographic data for the first 157 alkanes [118,119] will be used. The following QSCRR models, after the appropriate statistical work, were obtained [54]:

$$I_{\mathrm{calc}} = 170.6\ (\pm 17.0)\ (3\text{-}W_C)^{0.3245\ (\pm 0.0133)} - 31.8\ (\pm 30.2) \qquad (45)$$

with fair statistical characteristics: $r = 0.982$, $s = 35.6$ and $F = 2048$, and:

$$I_{calc} = 119.5\ (\pm 21.8)\ (3\text{-}W_{CH})^{0.2947\ (\pm 0.0174)} - 202.9\ (\pm 57.0) \tag{46}$$

with better statistical parameters: $r = 0.986$, $s = 31.0$ and $F = 2728$.

It is interesting to note that the QSCRR model with 2-W index possesses slightly better statistical characteristics than the QSCRR model with 3-W_C index. The QSCRR model with 3-W_{CH} index produced the best statistical parameters of the three models considered. The 3-W_{CH} reflects well the structure of a molecule via the geometric-distance matrix. Thus, it appears that the structural characteristics of molecules are one of the most important factors responsible for molecular migration in the chromatographic process.

4. Concluding remarks

In this report, the graph-theoretical (topological) distance matrix and the geometric-distance (topographic) matrix and their invariants (polynomials, spectra, determinants and the Wiener numbers) are discussed. The methods for computing all these quantities are presented. It is worth noting that in most cases the results obtained by means of the topological distance matrix and its invariants are comparable to those obtained by means of the topographic distance matrix and related invariants. The topological and topographic distance matrices are related via the complete weighted graphs, subgraphs of which correspond to graphs of molecules, while the weights correspond to geometric distances between atoms in the molecules in question.

The discriminating power of topological invariants is less than that of topographic invariants. At this point, it appears that the geometric-distance matrix and its invariants are unique molecular descriptors. These quantities can be used to differentiate the isospectral molecules and conformational isomers. In this way, they represent a novel basis for modeling 3D properties of molecules via appropriate QSPR approaches [14,44,45,48,54,57,121,122,125].

Acknowledgements

This work was supported in part by the Ministry of Science, Technology and Informatics of the Republic of Croatia. We thank Professor Haruo Hosoya (Tokyo), Dr. Teodor-Silviu Balaban (Bucharest), Dr. Milan Kunz (Brno), Professor Ante Graovac (Zagreb) and Professor Tomislav Živković (Zagreb) for useful suggestions and correspondence. We also thank the referees for helpful comments.

References

[1] V. Prelog, in: *Mathematics and Computational Concepts in Chemistry*, ed. N. Trinajstić (Horwood, Chichester, 1986), p. 1.

[2] F. Harary, *Graph Theory*, 2nd Ed. (Addison-Wesley, Reading, MA, 1971), p. 203.

[3] L.W. Beineke and R.J. Wilson, in: *Selected Topics in Graph Theory*, ed. L.W. Beineke and R.J. Wilson (Academic Press, London, 1978), p. 13.

[4] R.L. Graham, in: *Selected Topics in Graph Theory 3*, ed. L.W. Beineke and R.J. Wilson (Academic Press, London, 1983), p. 133.
[5] G.M. Crippen, J. Comput. Phys. 24(1977)96; J. Math. Chem. 6(1991)307.
[6] T.F. Havel, I.D. Kuntz and G.M. Crippen, Bull. Math. Biol. 45(1983)665.
[7] R.L. Graham and P.M. Winkler, Trans. Amer. Math. Soc. 288(1985)527.
[8] D.H. Rouvray, in: *Chemical Applications of Graph Theory*, ed. A.T. Balaban (Academic Press, London, 1976), p. 175.
[9] N. Trinajstić, *Chemical Graph Theory*, Vol. 2 (CRC Press, Boca Raton, 1983), Ch. 4.
[10] D.H. Rouvray, in: *Mathematics and Computational Concepts in Chemistry* (Horwood, Chichester, 1986), p. 295.
[11] G.M. Crippen and T.F. Havel, *Distance Geometry and Molecular Conformation* (Wiley, New York, 1988).
[12] P.J. Hansen and P.C. Jurs, J. Chem. Educ. 65(1988)574.
[13] K. Balasubramanian, Chem. Phys. Lett. 169(1990)224.
[14] N. Trinajstić, *Chemical Graph Theory*, 2nd rev. Ed. (CRC Press, Boca Raton, 1992), Ch. 4.
[15] R.C. Prim, Bell Syst. Tech. J. 36(1957)1389.
[16] S.L. Hakimi and S.S. Yau, Quart. Appl. Math. 22(1964)305.
[17] F.T. Boesch, Quart. Appl. Math. 26(1968)607.
[18] R.L. Graham and H.O. Pollak, Bell Syst. Tech. J. 50(1971)2495.
[19] M. Edelberg, M.R. Garey and R.L. Graham, Discr. Math. 14(1976)23.
[20] R.L. Graham, A.J. Hoffman and H. Hosoya, J. Graph Theory 1(1977)85.
[21] J.M.S. Simões-Pereira and C.M. Zamfirescu, Linear Algebra and Its Applications 44(1982)1.
[22] R.E. Prather, *Elements of Discrete Mathematics* (Houghton Miffin, Boston, 1986), p. 413.
[23] M.R. Hoare, Adv. Chem. Phys. 40(1979)49.
[24] A. Cayley, Cambridge Math. J. 2(1841)267.
[25] M.G. Brunel, Mémoires Soc. Sci. Bordeaux 5(1895)165.
[26] M.J. Clark and S.F.A. Kettle, Inorg. Chim. Acta 14(1975)201.
[27] I. Tinoco, Jr., O.C. Uhlenbeck and M.D. Levine, Nature 230(1971)362.
[28] H. Hosoya, Bull. Chem. Soc. Japan 44(1971)2332.
[29] D.H. Rouvray, Sci. Amer. 254(1986)40.
[30] A.T. Balaban (ed.), *Chemical Applications of Graph Theory* (Academic Press, London, 1976).
[31] R.B. King (ed.), *Chemical Applications of Topology and Graph Theory* (Elsevier, Amsterdam, 1983).
[32] N. Trinajstić (ed.), *Mathematics and Computational Concepts in Chemistry* (Horwood, Chichester, 1986).
[33] H. Hosoya, M. Murakami and M. Gotoh, Nat. Sci. Rep. Ochanomizu Univ. (Tokyo) 24(1973)27.
[34] P. Křivka and N. Trinajstić, Appl. Mat. (Prague) 28(1983)357; A. Graovac, G. Jashari and M. Strunje, Appl. Mat. 30(1985)286.
[35] K. Balasubramanian, J. Comput. Chem. 11(1990)829.
[36] H. Wiener, J. Amer. Chem. Soc. 69(1947)17.
[37] D. Bonchev, *Information Theoretic Indices for Characterization of Chemical Structures* (Wiley, Chichester, 1983), p. 71.
[38] I. Gutman and O.E. Polansky, *Mathematical Concepts in Organic Chemistry* (Springer, Berlin, 1986), p. 124.
[39] D. Bonchev and N. Trinajstić, J. Chem. Phys. 67(1977)4517.
[40] D. Bonchev and N. Trinajstić, Int. J. Quant. Chem. Quant. Chem. Symp. 12(1978)293.
[41] A.T. Balaban, Chem. Phys. Lett. 89(1982)399.
[42] P.G. Seybold, M. May and U.A. Bagal, J. Chem. Educ. 64(1987)575.
[43] D.E. Needham, I.-C. Wei and P.G. Seybold, J. Amer. Chem. Soc. 110(1988)4186.

[44] M. Randić, in: *MATH/CHEM/COMP 1987*, ed. R.C. Lacher (Elsevier, Amsterdam, 1988), p. 101.

[45] M. Randić, Int. J. Quant. Chem. Quant. Biol. Symp. 15(1988)201.

[46] I. Lukovits, J. Chem. Soc. Perkin Trans. II(1988)1667; Quant. Struct.-Act. Relat. 9(1990)227; Rep. Mol. Theory 1(1990)127.

[47] H.P. Schultz, J. Chem. Inf. Comput. Sci. 29(1989)227.

[48] B. Bogdanov, S. Nikolić and N. Trinajstić, J. Math. Chem. 3(1989)291; 5(1990)305.

[49] S.S. Tratch, M.I. Stankevich and N.S. Zefirov, J. Comput. Chem. 11(1990)899.

[50] H.P. Schultz, E.B. Schultz and T.P. Schultz, J. Chem. Inf. Comput. Sci. 30(1990)27.

[51] W.R. Müller, K. Szymanski, J.V. Knop and N. Trinajstić, J. Chem. Inf. Comput. Sci. 30(1990)160.

[52] J.V. Knop, W.R. Müller, K. Szymanski and N. Trinajstić, J. Chem. Inf. Comput. Sci. 31(1991)83.

[53] H.P. Schultz and T.P. Schultz, J. Chem. Inf. Comput. Sci. 31(1991)145.

[54] N. Bošnjak, Z. Mihalić and N. Trinajstić, J. Chromatogr. 540(1991)430.

[55] A.T. Balaban, Pure. Appl. Chem. 55(1983)199.

[56] M.I. Stankevich, I.V. Stankevich and N.S. Zefirov, Russian Chem. Rev. 57(1988)191.

[57] Z. Mihalić, S. Nikolić and N. Trinajstić, J. Chem. Inf. Comput. Sci. 32(1992)28.

[58] F. Harary, in: *Applied Combinatorial Mathematics*, ed. E.F. Beckenbach (Wiley, New York, 1964), p. 185.

[59] F. Harary, in: *A Seminar on Graph Theory*, ed. F. Harary and L.W. Beineke (Holt, Reinhart and Winston, New York, 1967), p. 9.

[60] A.N. Patrinos and S.L. Hakimi, Quart. Appl. Math. 30(1973)255.

[61] J.K. Kruskal, Jr., Proc. Amer. Math. Soc. 7(1956)48.

[62] W. Mayeda, *Graph Theory* (Wiley, New York, 1972).

[63] D.E. Johnson and J.R. Johnson, *Graph Theory with Engineering Applications* (Ronald Press, New York, 1972).

[64] N. Trinajstić, *Chemical Graph Theory*, 2nd rev. Ed. (CRC Press, Boca Raton, 1991), Ch. 3.

[65] A. Graovac, O.E. Polansky, N. Trinajstić and N. Tyutyulkov, Z. Naturforsch. 30a(1975)1696.

[66] M.J. Rigby, R.B. Mallion and A.C. Day, Chem. Phys. Lett. 51(1977)178; erratum 53(1978)418.

[67] M. Barysz, G. Jashari, R.S. Lall, V.K. Srivastava and N. Trinajstić, in: *Chemical Applications of Topology and Graph Theory*, ed. R.B. King (Elsevier, Amsterdam, 1983), p. 222.

[68] N. Trinajstić, Kem. Ind (Zagreb) 33(1984)311.

[69] A.T. Balaban, Math. Chem. (Mülheim/Ruhr) 21(1986)115.

[70] M. Randić, A. Sabljić, S. Nikolić and N. Trinajstić, Int. J. Quant. Chem. Quant. Chem. Symp. 15(1988)267.

[71] A.T. Balaban and O. Ivanciuc, in: *MATH/CHEM/COMP 1988*, ed. A. Graovac (Elsevier, Amsterdam, 1989), p. 193.

[72] C. Berge, *Théorie des Graphes et ses Applications* (Dunod, Paris, 1958), p. 136.

[73] F.S. Roberts, *Discrete Mathematical Model* (Prentice-Hall, Englewood Cliffs, 1976), p. 58; see also H. Hosoya, in: *Proceedings of the Conference on Molecular Structure* (Tokyo, 1970), p. 291.

[74] R.W. Floyd, Comm. ACM 5(1962)345.

[75] M. Bershon, J. Comput. Chem. 4(1983)110.

[76] I.V. Peredunova, V.E. Kuzmin and Y.P. Konovortskii, Russian J. Struct. Chem. 24(1983)645.

[77] N. Deo and C. Pang, Networks 14(1984)275.

[78] W.R. Müller, K. Szymanski, J.V. Knop and N. Trinajstić, J. Comput. Chem. 8(1987)170.

[79] B. Mohar and T. Pisanski, J. Math. Chem. 2(1988)267.

[80] P. Senn, Comput. Chem. 12(1988)219.

[81] C. Jochum and J. Gasteiger, J. Chem. Inf. Comput. Sci. 17(1977)113.

[82] W. Schubert and I. Ugi, J. Amer. Chem. Soc. 100(1978)37.

[83] M. Randić and C.L. Wilkins, J. Chem. Inf. Comput. Sci. 19(1979)31.

[84] M. Uchino, J. Chem. Inf. Comput. Sci. 20(1980)116.
[85] D. Bonchev, A.T. Balaban and O. Mekenyan, J. Chem. Inf. Comput. Sci. 20(1980)106.
[86] D. Bonchev, A.T. Balaban and M. Randić, Int. J. Quant. Chem. 19(1981)61; 22(1982)441.
[87] D. Bonchev, O. Mekenyan and A.T. Balaban, J. Chem. Inf. Comput. Sci. 29(1989)91.
[88] U.J.J. Le Verrier, J. Math. 5(1840)95, 220.
[89] V.N. Faddeeva, *Computational Methods of Linear Algebra* (Dover, New York, 1959).
[90] D.K. Faddeev and I.S. Sominskii, *Problems in Higher Algebra* (Freeman, San Francisco, 1965).
[91] P.S. Dwyer, *Linear Computations* (Wiley, New York, 1951), p. 225.
[92] K. Balasubramanian, Theor. Chim. Acta 65(1984)49; J. Comput. Chem. 5(1984)387; 6(1985)659; 9(1988)204; 12(1991)248.
[93] R. Ramaraj and K. Balasubramanian, J. Comput. Chem. 6(1985)122.
[94] P. Křivka, Ž. Jeričević and N. Trinajstić, Int. J. Quant. Chem. Quant. Chem. Symp. 19(1986)129.
[95] T.P. Živković, J. Comput. Chem. 11(1990)217.
[96] C.-E. Fröberg, *Introduction to Numerical Analysis*, 2nd Ed. (Addison-Wesley, Reading, MA, 1970), Ch. 6.
[97] I. Gutman and N. Trinajstić, Topics Curr. Chem. 42(1973)49.
[98] W.C. Herndon, Tetrahedron Lett. (1974)671.
[99] T. Živković, N. Trinajstić and M. Randić, Mol. Phys. 30(1975)517.
[100] W.C. Herndon and M.L. Ellzey, Jr., Tetrahedron 31(1975)99.
[101] M. Randić, N. Trinajstić and T. Živković, J. Chem. Soc. Faraday Trans. II (1976)244.
[102] S.S. D'Amato, B.M. Gimarc and N. Trinajstić, Croat. Chem. Acta 54(1981)1.
[103] N. Trinajstić, *Chemical Graph Theory*, Vol. 1 (CRC Press, Boca Raton, 1983), Ch. 7.
[104] M. Randić, M. Barysz, J. Nowakowski, S. Nikolić and N. Trinajstić, J. Mol. Struct. (THEOCHEM) 185(1989)95.
[105] J.R. Dias, *Molecular Orbital Calculations using Chemical Graph Theory* (Springer, Berlin, 1992).
[106] A.J. Schwenk, in: *New Directions in the Theory of Graphs*, ed. F. Harary (Academic Press, New York, 1973), p. 275.
[107] T. Živković, Report at the Quantum Chemistry School, Repino near Leningrad, December 1973.
[108] B.B. McKay, Ars Combinatoria 3(1977)219.
[109] R.L. Graham and L. Lovász, Adv. Math. 29(1978)60;
see also K.L. Collins, Discr. Appl. Math. 25(1989)27.
[110] R.L. Graham and L. Lovász, Lecture Notes Math. 642(1978)186.
[111] D. Plavšić and N. Trinajstić, unpublished.
[112] D. Bonchev, O. Mekenyan and N. Trinajstić, Int. J. Quant. Chem. 17(1980)845.
[113] M. Barysz, D. Plavšić and N. Trinajstić, Math. Chem. (Mülheim/Ruhr) 89(1980)149.
[114] D. Bonchev, O. Mekenyan, G. Protić and N. Trinajstić, J. Chromatogr. 176(1979)149.
[115] E.R. Canfield, R.W. Robinson and D.H. Rouvray, J. Comput. Chem. 6(1985)598.
[116] N. Adler, D. Babić and N. Trinajstić, Fresenius Z. Anal. Chem. 322(1985)246.
[117] N. Bošnjak, N. Adler, M. Perić and N. Trinajstić, in: *Modelling of Structures and Properties of Molecules*, ed. Z.B. Maksić (Horwood, Chichester, 1987), p. 103.
[118] J.R. Chrétien and J.-E. Dubois, J. Chromatogr. 126(1976)171.
[119] G. Schomburg and G. Dielman, J Chromatogr. Sci. 11(1973)151.
[120] J.V. King and R.J. Kassel, Int. J. Quant. Chem. Quant. Biol. Symp. 232(1991)65.
[121] Z. Mihalić and N. Trinajstić, J. Mol. Struct. (THEOCHEM), in press.
[122] S. Nikolić, N. Trinajstić, Z. Mihalić and S. Carter, Chem. Phys. Lett. 179(1991)21.
[123] K. Burkert and N.L. Allinger, Molecular Mechanics (ACS, Washington, 1982).
[124] A.T. Balaban, Rev. Roum. Chim. 21(1970)1049.
[125] M. Randić, B. Jerman-Blažič and N. Trinajstić, Comput. Chem. 14(1990)237.