The Wiener Distance Matrix for Acyclic Compounds and Polymers

Evgenii A. Smolenskii*

Laboratory of Mathematical Chemistry, N.D. Zelinsky Institute of Organic Chemistry, 47 Leninsky prosp., Moscow 117913 Russian Federation

Received July 31, 2003

The relationship between the Wiener indices and the topological structures of alkanes is analyzed. The expressions for the Wiener distances between elements of these structures are derived, and the distance matrix is constructed for them; this matrix is naturally called the Wiener distance matrix. The expressions for the Wiener indices of polymers with units of arbitrary structure are obtained.

I. INTRODUCTION

The Wiener index was proposed for describing the dependence between the physicochemical properties of alkanes and their structures.[1,2] It reaches the highest values for unbranched isomers and decreases as the degree of their branching increases. In terms of the distance matrix of a graph, according to Hosoya, this index is expressed as the halfsum of all its entries.[3] Numerous topological indices devised on the basis of the distance matrix[4−6] correlate with certain properties of alkanes and other chemical compounds. Several variants have been proposed for the generalization of the Wiener index; e.g., see refs 7−11.

Each entry $d_{i,j}$ in the distance matrix $\mathbf{D} = (d_{i,j})$ of a graph is an integer number equal to the topological distance between vertices i and j, that is, to the number of edges between them on the shortest path joining i and j. For alkanes, it is the number of C−C bonds between atoms i and j. By definition, we have $d_{ii} = 0$; therefore, the leading diagonal of the **D** matrix is zero. Let us consider the $\mathbf{D}_n$ matrix for a normal alkane C_nH_{2n+2} assuming that the atoms are sequentially numbered from 1 (one of the terminal atoms) to n (the other terminal atom):

$$\mathbf{D}_n = \begin{vmatrix} 0 & 1 & 2 & 3 & \ldots & \ldots & n-1 \\ & 0 & 1 & 2 & \ldots & \ldots & n-2 \\ & & 0 & 1 & \ldots & \ldots & n-3 \\ & & & \cdot & \cdot & \cdot & \cdot \\ & & & & 0 & 1 & 2 \\ & & & & & 0 & 1 \\ & & & & & & 0 \end{vmatrix} \qquad (1)$$

Since the matrix is symmetrical, we omitted the entries below the leading diagonal in order to simplify the notation. For entries above the leading diagonal, we have $d_{i,j} = j - i$. The Wiener index for the normal C_nH_{2n+2} alkane, that is, the sum of all the shown entries (the halfsum for the entire matrix), is expressed as

$$W(C_nH_{2n+2}) = \frac{(n+1)n(n-1)}{3!}$$

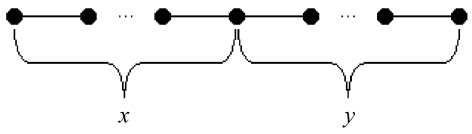
The succession of these values for $n = 1, 2, 3, 4, 5, 6, \ldots$ is 0, 1, 4, 10, 20, 35, 56, etc. This succession forms a series of the so-called tetrahedral numbers, which are found in the Pascal triangle. On the other hand, the same succession is formed by the numbers of combinations of $n + 1$ elements taken 3 at a time. Note that the generalization given in ref 7 uses the succession of the numbers of 5-element combinations rather than 3-element ones.

II. THE WIENER FUNCTION

Despite its seeming triviality, let us define the *Wiener function* for nonnegative variables:

$$W(x) = \frac{(x+2)(x+1)x}{3!} \qquad (2)$$

For a normal C_nH_{2n+2} alkane, which may be regarded as a chain of length $x = n - 1$, the function thus defined is the Wiener index. The term *function* is introduced here because eq 2 may be used for any real numbers, not only for integer x values. So, function 2 is the Wiener index for a chain of length x. Now let us see how this index behaves as the chain is extended. Let us connect two chains of lengths x and y via one common vertex:



The length of the chain thus formed is $x + y$ (the number of vertices is $x + y + 1$), and the Wiener index for it is expressed by formula 2 with the argument equal to $x + y$. Let us consider the distance matrix 1 for this new chain:

$$\mathbf{D}_{x+y} = \begin{vmatrix} 0 & 1 & 2 & \ldots & x & x+1 & \ldots & x+y \\ & 0 & 1 & \ldots & x-1 & x & \ldots & x+y-1 \\ & & & \ldots & \ldots & & \ldots & \\ & & & 0 & 1 & 2 & \ldots & y+1 \\ & & & & 0 & 1 & \ldots & y \\ & & & & & & \ldots & \\ & & & & & & 0 & 1 \\ & & & & & y & & 0 \end{vmatrix} \begin{matrix} \\ \\ \\ \\ x \\ \\ \\ \\ y \end{matrix} \qquad (3)$$

This matrix may be divided into three components: two triangles and one rectangle. The first triangle is the distance halfmatrix $\mathbf{D}_x$ for the chain of length x; the second triangle is the $\mathbf{D}_y$ halfmatrix for the chain of length y; and the rectangle of size xy consists of the lengths of the segments that start within the first subchain (x) and end within the

10.1021/ci030026k CCC: $27.50
Published on Web 02/26/2004

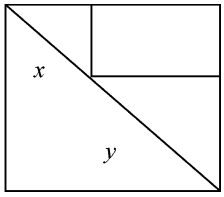
second one (y). The symmetrical values for the segments starting within the second subchain and ending within the first subchain are located in the rectangle (not shown here) of size yx in the bottom left corner of the $\mathbf{D}_{x+y}$ matrix. Let us consider the distance matrix for this new chain of length $x + y$ at $x = 3$ and $y = 2$:

$\mathbf{D}_{3+2} =$

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

This example shows that the numbers in the subcolumn above the zero entry corresponding to the common vertex of chains x and y in matrix 3 belong to the $\mathbf{D}_x$ matrix, and those in the subrow on the right to the same entry belong to the $\mathbf{D}_y$ matrix. Further, we will display matrices of form 3 in a more schematic way:

$\mathbf{D}_{x+y} =$



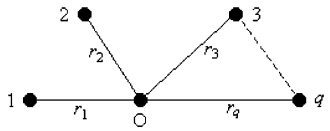
For the sums of the entries in the triangles, we have $W(x) \times W(y)$. Let us denote the sum of the entries in the rectangle by $W_0(x, y)$; hereafter, this value will be referred to as the *Wiener distance* between adjoining (having one common vertex) segments x and y. An evident relationship follows from the scheme for matrix 3

$$W(x + y) = W(x) + W(y) + W_0(x,y) \qquad (4)$$

and the resultant expression for the Wiener distance between x and y is

$$W_0(x,y) = W(x + y) - W(x) - W(y) \qquad (5)$$

Since formulas 4 and 5 may seem rather trivial, let us consider, for example, the expressions for the Wiener indices of star graphs with q rays of different lengths:



($\mathbf{G}_1$)

The r_i values for graph $\mathbf{G}_1$ are equal to the lengths of chains from the central vertex O to vertices i of degree 1. A generalization of formula 4 for $\mathbf{G}_1$ consists of the fact that its Wiener index is the sum of the Wiener indices of all the rays (subchains) plus the sum of all Wiener distances between the subchains, that is,

$$W(\mathbf{G}_1) = \sum_{i=1}^{q} W(r_i) + \sum_{(i,j)} W_0(r_i,r_j)$$

where the notation (i, j) implies that the summation is performed over vertex pairs, e.g., under the condition $i < j$. Substituting expression 5 for W_0 and considering that each r_i is characterized by $q - 1$ Wiener distances to other rays, we obtain

$$W(\mathbf{G}_1) = \sum_{i=1}^{q} W(r_i) + \sum_{(i,j)} W(r_i + r_j) - (q - 1)\sum_{i=1}^{q} W(r_i)$$

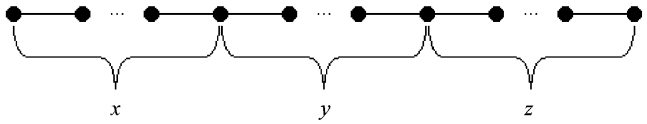
and thus the Wiener index may be represented in the form

$$W(\mathbf{G}_1) = \sum_{(i,j)} W(r_i + r_j) - (q - 2)\sum_{i=1}^{q} W(r_i) \qquad (6)$$

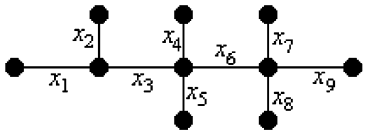
The explicit expression for the Wiener index of graph $\mathbf{G}_1$ was considered in ref 12 for the case of all the rays having equal lengths; however, our formulas for this case are simpler and more generalized.

Note that formula 6 at $q = 2$ is reduced to $W(r_1 + r_2)$. For $q = 3$ and $q = 4$, that is, for tri- and tetralkylmethanes,[13] it yields simple relationships for calculating the Wiener index without constructing the distance matrix.

Now let us find the Wiener index for three or more consecutively linked subchains:



For the $\mathbf{D}_{x+y+z}$ matrix, similarly to (3), we have



($\mathbf{G}_2$)

It is assumed that all vertices are numbered just as for formula 1. The meaning of the notations in the triangles and in the two rectangles adjacent to them has already been explained. As to the expression in the upper right rectangle, which is referred to as the Wiener distance between segments x and z separated by segment y, this is the set of the lengths of the segments that start within subchain x and end within subchain z. These segments are separated by segment y, and therefore we will denote this Wiener distance by $W_y(x, z)$. Thus the notations W_0 and W_y are compatible, because the terminal segments come into contact at $y = 0$, and $W_{y|y=0} = W_0$.

Analyzing the schematic structure of matrix 7, one can easily derive the relationship

$$W(x + y + z) = W(x + y) + W(y + z) + W_y(x,z) - W(y) \qquad (8)$$

the last term with the minus sign appears because the sum of $W(x + y)$ and $W(y + z)$ considers the $W(y)$ value (that is, the middle triangle) twice. Thus, we arrive at

$$W_y(x,z) = W(x + y + z) - W(x + y) - W(y + z) + W(y)$$

However, note that each entry in the triangle corresponding to $W_y(x, z)$ in scheme 7 contains distance y, and the number of such entries is xz. Therefore, it is evident that

$$W_y(x,z) = W_0(x,z) + xyz \quad (9)$$

Substituting (9) into (8) and making allowance for (5), we obtain

$$W(x+y+z) = W(x+y) + W(y+z) - W(y) + W(x+z) - W(x) - W(z) + xyz$$

whence

$$xyz = W(x+y+z) - W(x+y) - W(x+z) - W(y+z) + W(x) + W(y) + W(z) \quad (10)$$

The latter expression is true for any nonnegative integer numbers x, y, and z (however, some or all of these numbers may be zero). It is not worthwhile to consider expressions similar to (7) for four or more consecutively linked subchains, like a $\mathbf{D}_{x+y_1+y_2+z}$ matrix, because the expression $W_{y_1+y_2}(x,z) = W_0(x,z) + x(y_1 + y_2)z$ for this case coincides with (9) at $y = y_1 + y_2$.

Since any number p may be presented as the product of three numbers $p = p\cdot1\cdot1$, we obtain $p = p\cdot1\cdot1 = W(p+2) - 2W(p+1) + W(p) - W(2) + W(1)$. This is actually a more complicated modification of the simpler expression

$$p = W(p) - 2W(p-1) + W(p-2)$$

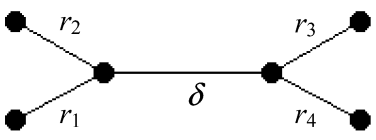
which mathematically corresponds to the second derivative of function 2 in the expressions for finite differences. So, any number may be represented in the form $p = p\cdot1\cdot1$, and expression 10 may be obtained for it. If p is a prime number, such a representation is unique; in the opposite case, p may be represented in the form $p = p_1\cdot p_2\cdot1$, where the cofactors p_1 and p_2 are either prime or composite. Formula 10 yields

$$p = W(p_1 + p_2 + 1) - W(p_1 + p_2) - W(p_1 + 1) - W(p_2 + 1) + W(p_1) + W(p_2) + W(1) \quad (11)$$

This, if p is not a prime number, eq 11 with respect to the p_1 and p_2 variables (the free term of this equation contains p) has one or several solutions in the form of pairs of integer numbers $1 < p_1, p_2 < p$. However, if p is a prime number, the third-order eq 11 has no integer solutions. This is where function 2 is used instead of the index.

III. THE WIENER DISTANCE MATRIX

Now let us analyze how the Wiener index is expressed for arbitrary trees using eqs 5 and 9. Evidently, one can present the topological structure of an alkane omitting vertices of degree 2. For example, let us consider the following graph:

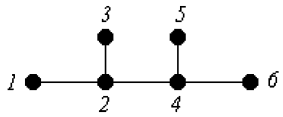


($\mathbf{G}_3$)

All x_i denote unbranched subchains whose length is equal to 1 plus the number of vertices of degree 2 inside the ith subchain. For these subchains, one can define the numbers $W_{ii} = W(x_i)$; that is, W_{ii} is the Wiener index for subchain x_i. For $i \neq j$, we have $W_{ij} = W_\delta(x_i,x_j)$ values, which are the Wiener distances between subchains x_i and x_j if these subchains are located at distance δ from each other. For example, for tree $\mathbf{G}_2$ we have $W_{12} = W_0(x_1,x_2)$, $W_{13} = W_0(x_1,x_3)$, $W_{34} = W_0(x_3,x_4)$, etc., as well as $W_{14} = W_{x_3}(x_1,x_4)$, $W_{16} = W_{x_3}(x_1,x_6)$, etc. and also $W_{17} = W_{x_3+x_6}(x_1,x_7)$, $W_{29} = W_{x_3+x_6}(x_2,x_9)$, etc. The values thus defined may be represented in the form of the symmetrical *Wiener distance matrix*:

$$\mathbf{D}_W = (W_{ij}) \quad (12)$$

The dimension of this matrix is usually much smaller than that of the known matrix 1. For the Wiener index, we have the expression $W = \sum W_{ii} + 1/2\sum W_{ij}$; that is, the sum of the diagonal entries plus the halfsum of all other entries of the Wiener distance matrix. (By the way, the same expression is formally true for the Wiener index calculated according to the ordinary distance matrix, but the sum of the diagonal entries is zero in the latter case.) As an example, let us consider a simpler graph: The Wiener distance matrix 12



for graph $\mathbf{G}_3$ will have the following form

$$\mathbf{D}_W(\mathbf{G}_3) = \begin{vmatrix} W(1) & W_0(1,2) & W_0(1,\delta) & W_\delta(1,3) & W_\delta(1,4) \\ & W(2) & W_0(2,\delta) & W_\delta(2,3) & W_\delta(2,4) \\ & & W(\delta) & W_0(\delta,3) & W_0(\delta,4) \\ & & & W(3) & W_0(3,4) \\ & & & & W(4) \end{vmatrix}$$

where all r_i are substituted by the corresponding symbols i for simplicity. The expression for the Wiener index is

$$W(\mathbf{G}_3) = \sum_{i=1}^{4} W(r_i) + W(\delta) + W_0(r_1,r_2) + W_0(r_3,r_4) + \sum_{i=1}^{4} W_0(r_i,\delta) + \sum_{\{(i,j):i=1,2;j=3,4\}} W_\delta(r_i,r_j)$$

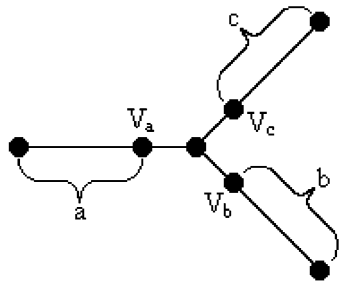
Let us assume that we have $r_1 = 3$, $r_2 = \delta = r_4 = 2$, $r_3 = 1$; that is, 4-ethyl-6-methyloctane. The Wiener index is equal to

$$W = W(3) + 3W(2) + W(1) + 3W(5) + 4W(4) + 3W(3) - 4W(3) - 12W(2) - 4W(1) + 3\cdot2\cdot1 + 3\cdot2\cdot2 + 2\cdot2\cdot2 + 2\cdot2\cdot1 = 3W(5) + 4W(4) - 9W(2) - 3W(1) + 30 = 105 + 80 - 36 - 3 + 30 = 176$$

We have written out this expression in such detail on purpose, to provide a better understanding of the summation process. In the first line, we have the W values for all the segments (r_i and δ); in the second line, the sum of all positive terms in formula 5; in the third line, the negative terms in eq 5; and the fourth line contains the sums of all the $r_i\cdot d\cdot r_j$ values ($i = 1, 2$; $j = 3, 4$) in formula 9. This example shows how the calculation of the *Wiener index* is simplified by the use of the formulas derived in this work.

The Wiener matrix 12 may be regarded as a certain generalization of the distance matrix for a tree. This becomes evident if the lengths of all subchains x_i (e.g., in graph $\mathbf{G}_2$) are assumed to be equal to unity. In this case, $\mathbf{D}_w$ is reduced

to the usual distance matrix with an accuracy up to the arrangement of its entries, that is, to the vertex numbering. In this case, we have $\lim \mathbf{D}_W = \mathbf{D}$ and $W_{ij} \rightarrow d_{ij}$ at $\max r_i \rightarrow 1$. Of course, the addition of a zero diagonal is assumed in this case, as is apparent from the following example. Let us consider the tree



$\mathbf{T}_{abc}$

which corresponds to 2,3-dimethylbutane. The Wiener distance matrix for it is

$$\mathbf{D}_W = \begin{vmatrix} 1 & 2 & 2 & 3 & 3 \\ & 1 & 2 & 3 & 3 \\ & & 1 & 2 & 2 \\ & & & 1 & 2 \\ & & & & 1 \end{vmatrix}$$

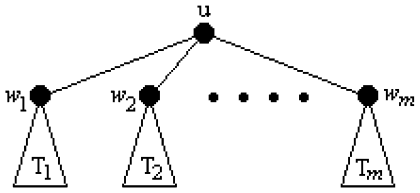
and the usual distance matrix is

$$\mathbf{D} = \begin{vmatrix} 0 & 1 & 2 & 2 & 3 & 3 \\ & 0 & 1 & 1 & 2 & 2 \\ & & 0 & 2 & 3 & 3 \\ & & & 0 & 1 & 1 \\ & & & & 0 & 2 \\ & & & & & 0 \end{vmatrix}$$

Thus, our assumption is confirmed, because all nonzero entries coincide with an accuracy up to the order of their arrangement.

IV. THE WIENER INDEX CALCULATIONS

Gutman proved the method of calculating the Wiener index described by Wiener himself[1] as a theorem in ref 14. In the same work, he considered a star graph $\mathbf{T}_{abc}$ with three rays of arbitrary lengths:



(T)

The picture of graph $\mathbf{T}_{abc}$ and all the symbols are taken from ref 14. After cumbersome calculations, having proved another interesting theorem important for certain applications, Gutman derived an expression for the Wiener index of this graph

$$W(\mathbf{T}_{abc}) = \binom{n+1}{3} - abc \qquad (13)$$

where a, b, and c are the numbers of vertices in the rays of the star and $n = a + b + c + 1$ is the total number of vertices in graph $\mathbf{T}_{abc}$. Allowing for the fact that $r_1 = a$, $r_2 = b$, and $r_3 = c$ for $\mathbf{T}_{abc}$, we obtain from formula 6 that

$$W(\mathbf{T}_{abc}) = W(a + b) + W(a + c) + W(b + c) - W(a) - W(b) - W(c)$$

Denoting $x = a$, $y = b$, and $z = c$ and using formula 10, we obtain

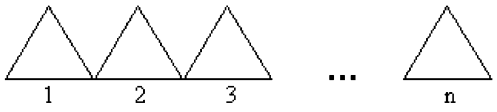
$$W(a + b) + W(a + c) + W(b + c) - W(a) - W(b) - W(c) = W(a + b + c) - abc$$

and this, with allowance for eq 13 (the left-hand parts coincide), yields

$$W(\mathbf{T}_{abc}) = W(a + b + c) - abc \qquad (14)$$

Since eq 2 implies that $W(a + b + c) = \binom{n+1}{3}$, formulas 13 and 14 coincide up to the notation, but it appears that we have derived this equation in a simpler way.

Let us consider graph **T** that was studied in ref 15 and considered as an example in ref 14:



(G_4)

Rouvray and coauthors[15] proposed a recurrent formula for calculating the Wiener index of this graph

$$W(\mathbf{T}) = \sum_{i=1}^{m} w(\mathbf{T}_i) + \sum_{i=1}^{m} [dw_i|(\mathbf{T}_i) + n_i] + \sum_{i<j} [d(w_i|\mathbf{T}_i) + d(w_j|\mathbf{T}_j) + 2n_i n_j] \qquad (15)$$

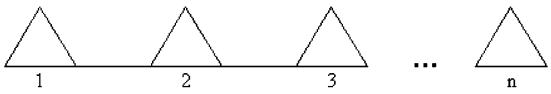
where the $d(w_i|\mathbf{T}_i)$ values are defined as

$$d(v_i|\mathbf{T}) = \sum_{j=1}^{n} d(v_i, v_j)$$

with $d(v_i, v_j)$ being the entries of the usual distance matrix $\mathbf{D}(\mathbf{T})$. The first sum in formula 15 is the sum of the Wiener indices for subtrees $\mathbf{T}_i$. The second sum is actually the sum of the Wiener distances between vertex u and subtree $\mathbf{T}_i$ (each of the distances from this vertex to all vertices of subtree $\mathbf{T}_i$ is one unit greater than the similar distance from vertex w_i, and the number of such distances is n_i). The third sum in (15) is the sum of the Wiener distances between subtrees $\mathbf{T}_i$. Each distance between a vertex of subtree $\mathbf{T}_i$ and a vertex of subtree $\mathbf{T}_j$ may be split into three components. The first component is the distance from this vertex of subtree $\mathbf{T}_i$ to vertex w_i; the second one is the distance from vertex w_i to w_j (equal to 2); and the third one is the distance from vertex w_j to the selected vertex of subtree $\mathbf{T}_j$. Since the number of such vertex pairs is $n_i n_j$, the term equal to $2n_i n_j$ in eq 15 follows from formula 9.

Note that recurrent formula 15 from ref 15 still refers to a particular case, whereas the suggested Wiener distance matrix is more generalized. For example, one can consider not only unbranched subchains as elements of the topological structure but also more complex fragments, including those with cycles of an arbitrary length. The analytical expressions for the Wiener indices of some polyspirographs containing

three-, four-, five-, and six-membered rings were derived in ref 11. These expressions were calculated using the "layer matrix of cardinality".[16] However, using the Wiener distance matrix, one can easily derive such expressions. Let us consider some simple example, like polyspirocyclopropane:



(G_5)

The molecular formula of compound $\mathbf{G}_4$ is $C_{2n+1}H_{2n+4}$ (the number of carbon atoms is always odd). For cyclopropane alone, the Wiener index is $W^{\Delta} = 3$; the Wiener distance between the two cycles in dispirocyclopropane is $W_0^{\Delta,\Delta} = 8$, and the distance between two cycles separated by δ cycles is obtained according to formula 9: $W_\delta^{\Delta,\Delta} = 8 + 4\delta$ (here $x = z = 2$ and $y = \delta$). The Wiener distance matrix will look like this:

$$\mathbf{D}_W(\mathbf{G}_4) = \begin{vmatrix} 3 & 8 & 12 & 16 & \dots & 4n \\ & 3 & 8 & 12 & \dots & 4(n-1) \\ & & 3 & 8 & \dots & 4(n-2) \\ & & & \dots & \dots & \dots \\ & & & & 3 & 8 \\ & & & & & 3 \end{vmatrix}$$

that is, $W_{ii} = 3$ and $W_{ij} = 8 + 4\delta = 8 + 4(j - i - 1)$ (for the shown part of the matrix, we have $j > i$ and $\delta = j - i - 1$). To calculate $W(\mathbf{G}_4)$, one should summarize all the W_{ii} and W_{ij} entries shown here. Such a summation is conveniently performed along the diagonals. It is easily seen that such a procedure yields

$$W(\mathbf{G}_4) = n\cdot 3 + (n-1)8 + (n-2)12 + \dots + (n-k)(k+1)4 + \dots + 4n = 3n + 4\sum_{k=1}^{n-1}(n-k)(k+1) = 3n + 4\sum_{k=1}^{n-1}(nk - k^2 + n - k) = 3n + 4\sum_{k=1}^{n-1}(n-1)k + 4\sum n - 4\sum_{k=1}^{n-1}k^2 + 4\sum_{k=1}^{n-1}n - 4\sum_{k=1}^{n-1}k^2 = 3n + 4(n-1)\sum_{k=1}^{n-1}k + 4n(n-1) - 4\sum_{k=1}^{n-1}k^2$$

Hence, if we allow for the fact that $\sum_{k=1}^{n-1} k = n(n-1)/2$ and $\sum_{k=1}^{n-1} k^2 = 2n^3 - 3n^2 + n/6$, the summation yields $W(\mathbf{G}_4) = n/3\,(2n^2 + 6n + 1)$. This formula coincides with the one found in ref 11 using the layer matrix of cardinality.[16] However, our summation of the $\mathbf{D}_W(\mathbf{G}_4)$ entries may still seem cumbersome. One can simplify it and represent the matrix in the form of functions of matrix 1. Separating the leading diagonal and factoring 4 out for the rest of the matrix, we obtain that the sums of its entries amount to

$$\begin{vmatrix} 3 & 8 & 12 & 16 & \dots & 4n \\ & 3 & 8 & 12 & \dots & 4(n-1) \\ & & 3 & 8 & \dots & 4(n-2) \\ & & & \dots & \dots & \dots \\ & & & & 3 & 8 \\ & & & & & 3 \end{vmatrix} = 3n + 4 \begin{vmatrix} 2 & 3 & \dots & n \\ & 2 & \dots & n-1 \\ & & \dots & \\ & & & 2 \end{vmatrix}$$

Subtracting 1 from each entry of the latter matrix (the number of such entries will be $n(n-1)/2$) and considering the multiplication by 4, we arrive at

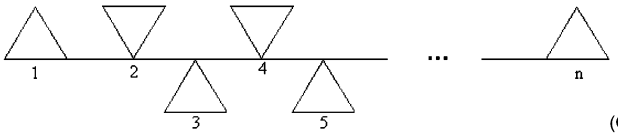
$$W(\mathbf{G}_4) = 3n + 4\frac{n(n-1)}{2} + 4 \begin{vmatrix} 1 & 2 & \dots & n-1 \\ & 1 & & n-2 \\ & & & \\ & & 1 & 2 \\ & & & 1 \end{vmatrix}$$

but the last term is the usual distance matrix for a chain of length $n - 1$ (without the zero diagonal) multiplied by 4; i.e., it may be expressed by formula 2. So, we arrive at

$$W(\mathbf{G}_4) = 3n + 4\frac{n(n-1)}{2} + \frac{4(n+1)n(n-1)}{6} = \frac{n(2n^2 + 6n + 1)}{3}$$

naturally, this expression coincides with the one derived above.

Let us consider two more complicated cases of triangulanes that were not analyzed in ref 11:



(G_6)

Just as before, denoting the Wiener distances between the cyclopropanes by $W_\delta^{\Delta,\Delta}$ and W_δ^{Δ}, we have $W_\delta^{\Delta} = 3$ (as in the previous case), but the distances δ between cycles i and j at $i < j$ amount to $\delta(i, j) = 2(j - i) - 1$; that is, δ varies from 1 for adjacent cycles ($j = i + 1$) to $2n - 3$ for $i = 1$ and $j = n$. One can easily calculate that $W_0^{\Delta,\Delta} = 21$. The $\delta(i, j)$ values are always odd, and, as $\delta(i, j)$ increases in passing from cycle j to $j + 1$, we have $\Delta W = 2\cdot 3\cdot 3 = 18$. Thus, one can easily construct the Wiener distance matrix for $\mathbf{G}_5$:

$$\mathbf{D}_W(\mathbf{G}_5) = \begin{vmatrix} 3 & 21 & 39 & 57 & \dots & [3 + 18(n-1)] \\ & 3 & 21 & 39 & \dots & [3 + 18(n-2)] \\ & & & 3 & \dots & [3 + 18(n-k)] \\ & & & & \dots & \dots \\ & & & & \dots & 3[3 + 18] \\ & & & & & 3 \end{vmatrix}$$

If we subtract 3 from each entry of this matrix (there will be $n(n + 1)/2$ such numbers) and then divide each of the remaining terms by 18, we obtain a simpler and more familiar matrix:

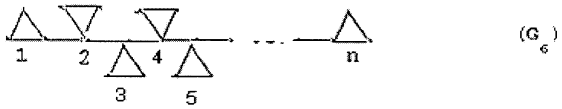
$$\mathbf{D}_W(\mathbf{G}_5) = \begin{vmatrix} 3 & 3 & \dots & 3 & & & \\ & 3 & \dots & 3 & & & \\ & & \dots & & & & \\ & & & 3 & & & \\ & & & & & & \\ & & & & & & \end{vmatrix} + 18 \begin{vmatrix} 0 & 1 & 2 & 3 & \dots & n-1 \\ & 0 & 1 & 2 & \dots & n-2 \\ & & 0 & 1 & \dots & n-3 \\ & & & & & \\ & & & & 0 & 1 \\ & & & & & 0 \end{vmatrix}$$

Since the second matrix (with the zero diagonal) coincides with matrix 1, the sum of its entries is expressed by formula 2, that is,

$$W(\mathbf{G}_5) = 3\frac{n(n+1)}{2} + 18\frac{(n+1)n(n-1)}{6} = \frac{3n(2n^2+n-1)}{2}$$

Finally, the Wiener index for graph $\mathbf{G}_5$ with n cyclopropanes is $W(n) = \{3n(2n^2 + n - 1)\}/2$.

Similarly calculating the Wiener index using the Wiener distance matrix for graph



($\mathbf{G}_6$)

we obtain $W_n(\mathbf{G}_6) = \{3n(n^2 + 4n - 3)\}/2$. Since compounds $\mathbf{G}_5$ and $\mathbf{G}_6$ with equal n are isomeric, let us compare their Wiener indices using the formulas derived above. Naturally, the structures coincide for $n = 1$ and $n = 2$. However, the results for $n = 3$ are already different: $W(\mathbf{G}_5) = 90$ and $W(\mathbf{G}_6) = 81$. The results for $n = 4$ are 210 and 174, respectively. If we consider the Wiener index as the measure of branching, these numbers show that isomers $\mathbf{G}_6$ is "more branched" than $\mathbf{G}_5$; actually, this is just a quantitative confirmation of the intuitive feeling.

In conclusion of this section, let us derive a general formula for calculating the Wiener indices of polymeric compounds with n arbitrary constituent units **R**:

$$\underset{1\ \ 2\qquad n}{\mathbf{R}\text{–}\mathbf{R}\text{–}\dots\text{–}\mathbf{R}} \qquad (\mathbf{G}_7)$$

Any structure, including cycles or their combinations, may be regarded as **R**. In ref 11, these units were represented by spiroconnected cyclopropanes, cyclobutanes, cyclopentanes, and cyclohexanes. In ref 17, the general formulas were also derived for W of polymeric compounds with units represented by chains. Let us assume that we have calculated the Wiener index W_0^R for a separate **R**. For the **R**−**R** pair, we have a Wiener index equal to $W_0^{R,R}$, and the distance between the beginnings of the chains is denoted by δ. Let the number of atoms in **R** be equal to m. Then the Wiener distance matrix 12 assumes the form

$$\mathbf{D}_W(\mathbf{G}_7) = \begin{vmatrix} W_0^R & W_0^{R,R} & W_0^{R,R}+\delta m^2 & W_0^{R,R}+2\delta m^2 & \dots & W_0^R+(n-2)\delta m^2 \\ & W_0^R & W_0^{R,R} & W_0^{R,R}+\delta m^2 & \dots & W_0^R+(n-3)\delta m^2 \\ & & & & & \\ & & & & W_0^R & W_0^{R,R} \\ & & & & & W_0^R \end{vmatrix} \quad (16)$$

One can easily find that this matrix contains n entries equal to W_0^R (the leading diagonal) and $n(n-1)/2$ entries equal to $W_0^{R,R}$ (in each matrix entry except for the leading diagonal). All the remaining entries of matrix 16 are represented by the δm^2 value multiplied by a number from 1 to $n - 2$. Factoring this value out, we obtain the third term:

$$\delta m^2 \begin{vmatrix} 0 & 1 & 2 & \dots & n-2 \\ & 0 & 1 & \dots & n-1 \\ & & & & \\ & & 0 & 1 & \\ & & & 0 & \end{vmatrix}$$

Thus, summing the terms described above with entries of the latter matrix according to formula 2, we arrive at

$$W(\mathbf{G}_7) = nW_0^R + \frac{n(n-1)}{2}W_0^{R,R} + \frac{n(n-1)(n-2)}{6}\delta m^2 \quad (17)$$

This expression may be represented by a polynomial

$$W(\mathbf{G}_7) = an^3 + bn^2 + cn \quad (18)$$

where the a, b, and c coefficients may be expressed using three quantities: W_0^R, $W_0^{R,R}$, and δm^2:

$$a = \delta m^2,\ b = \frac{W_0^{R,R} - \delta m^2}{2},\ c = \frac{6W_0^R - 3W_0^{R,R} + 2\delta m^2}{6}$$

However, the coefficients of formula 18 may be determined for any actual case according to the method of undetermined coefficients, without using the layer matrix of cardinality[16] of matrix 12. For this purpose, one should calculate the Wiener indices for any three terms of the series with the general formula $W(\mathbf{G}_7)$ (e.g., for $n = 1, 2, 3$) and solve a set of three linear equations. Nevertheless, before this one has to know that the formula for $W(\mathbf{G}_7)$ is of the general form 18. By the way, if we take carbon atoms as **R** (that is, take a normal alkane C_nH_{2n+2} instead of $\mathbf{R}_n$), we have $W_0^R = 0$, $W_0^{R,R} = 1$, and $m^2 = 1$. Formula 17 is transformed into formula 2, just as should have been expected.

V. CONCLUSION

One can use matrix 12 not only for calculating the Wiener index. Above all, note that the aforementioned dependence of the Wiener index on the branching consists of the fact that the index W is the sum of C_n^2 distances between all pairs of C atoms in the alkane, or, which is the same, between all vertex pairs in the tree. Divided by C_n^2, this index yields the average topological distance between the atoms:

$$\bar{\rho} = \frac{1}{N}\sum_{i=1}^{N}\rho_i = \frac{W}{C_n^2} \quad (19)$$

For normal alkanes, the resultant value is

$$\bar{\rho} = \frac{(n+1)n(n-1)}{3!} : \frac{n(n-1)}{2!} = \frac{n+1}{3}$$

For ethane ($n = 2$), the average (and only) interatomic distance is $2+1/3 = 1$. For propane, butane, and pentane, we have 4/3, 5/3, and 2, respectively, etc. That is, these numbers for unbranched alkanes form an arithmetical progression with a difference of $d = 1/3$. Since the numbers

of the atom pairs are equal among isomers, the decrease in the Wiener index for alkanes with increasing branching just reflects the decrease in the average topological distance between the atoms.

Randić[18] proposed new topological indices P' and P'/P, where $P = C_n^2$, that is, the number of paths in tree **G**. He defined P' as a sum over bonds, where the value put in correspondence with each bond is the sum of the numbers of paths in subgraphs $\mathbf{G}'_k$ and $\mathbf{G}''_k$ formed from graph **G** by deleting this bond. Although P' is a new index, Randić himself indicated that it is linearly related to the Wiener index,[18] although he did not express this dependence for the general case. One can show that these indices are related to each other via the distance $\bar{\rho}$. Denoting the numbers of vertices in subgraphs $\mathbf{G}'_k$ and $\mathbf{G}''_k$ by n'_k and n''_k (naturally, $n'_k + n''_k = n$), we obtain

$$P' = \sum_{k=1}^{n-1} [C_{n'_k}^2 + C_{n''_k}^2]$$

with the summation over all $n - 1$ bonds.

Let us show the relationship between P, W, and $\bar{\rho}$ by using the equalities $P = C_n^2 = n(n-1)/2$, $n = n'_k + n''_k$, and $n^2 = (n'_k + n''_k)^2 = n'^2_k + n''^2_k + 2n'_k n''_k$:

$$P' = \sum_{k=1}^{n-1}\left[\frac{n'_k(n'_k - 1)}{2} + \frac{n''_k(n''_k - 1)}{2}\right] = \frac{1}{2}\sum_{k=1}^{n-1}[n'^2_k + n''^2_k - (n'_k + n''_k)]$$

$$= \frac{1}{2}\sum_{k=1}^{n-1}[n^2 - 2n'_k n''_k - n] = \frac{n(n-1)}{2}\cdot(n-1) - \sum_{k=1}^{n-1} n'_k n''_k$$

Considering that the last sum is the Wiener index,[1,14] we obtain

$$P' = P\cdot(n-1) - W \qquad (20)$$

Dividing this equality by P and allowing for (19), we obtain

$$\frac{P'}{P} = \rho_{\max} - \bar{\rho} \qquad (21)$$

Equation 20 shows that Randić actually proposed a new (but hardly the simplest) way of calculating the W index.[18] Considering that $n - 1$ is the longest distance $\rho_{\max}$ between vertices in a graph with n vertices (and can be reached only in an unbranched alkane), eq 21 shows that the P'/P index is the difference between the maximum and average topological distances in graph **G**.

As to the Wiener distance matrix 12, its diagonal entries W_{ii} depend on the average distance between vertices of each segment $\mathbf{X}_i$ in a similar manner. The nondiagonal entries W_{ij}, which we termed the Wiener distances between segments $\mathbf{X}_i$ and $\mathbf{X}_j$, are also obtained by multiplying the number of the vertex pairs with one vertex from $\mathbf{X}_i$ and the other from $\mathbf{X}_j$ by the average distance between such vertices. But, since matrix 12 usually contains a much smaller number of entries than the usual distance matrix, we can now explore the dependences of the chemical compound properties not only on the sum of all the matrix entries (the Wiener index) but directly on separate components of the Wiener distance matrix.

In addition, by analogy with the construction of topological indices based on the usual distance matrix,[4-6] one can construct the corresponding topological indices based on the Wiener distance matrix. This procedure was performed for entries of matrix **D** for the distances $d_{ij} = 3$[1,2] and $d_{ij} = 2$[19] and considered in a more generalized form in refs 20 and 21.

ACKNOWLEDGMENT

The author is sincerely grateful to Prof. S. S. Tratch, Prof. N. S. Zefirov, and Dr. M. S. Molchanova for their critical reviews of this manuscript and helpful comments.

REFERENCES AND NOTES

(1) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, *69*, 17−20.
(2) Wiener, H. Correlation of Heats of Isomerization, and Differences in Heats of Vaporisation of Isomers, among the Paraffin Hydrocarbons. *J. Am. Chem. Soc.* **1947**, *69*, 2636−2638.
(3) Hosoya, H. Topological Index. A Newly Proposed Quantity Characterising the Topological Nature of Structural Isomers of Saturated Hydrocarbons. *Bull. Chem. Soc. Jpn*. **1971**, *4*, 2332−2339.
(4) Balaban, A. T. Topological indices based on topological distances in Molecular Graphs. *Pure Appl. Chem.* **1983**, *55*, 199−206.
(5) Randić, M.; Balaban, A. T.; Basalc, S. C. On structural Interpretation of several Distance Related Topological Indices. *J. Chem. Inf. Comput. Sci*. **2001**, *41*, 593−601.
(6) Stankevich, M. I.; Stankevich, I. V.; Zefirov, N. S. Topological Indices in Organic Chemistry. *Russ. Chem. Rev.* **1988**, *57*, 191−208.
(7) Tratch, S. S.; Stankevich, M. I.; Zefirov, N. S. Combinatorial models and algorithms in chemistry. The expanded Wiener number − a novel topological index. *J. Comput. Chem.* **1990**, *11*, 899−908.
(8) Randić, M.; Guo, X.; Oxley, X.; Krishnapriyan, H. Wiener Matrix: Source of Novel Graph Invariants. *J. Chem. Inf. Comput. Sci*. **1993**, *33*, 709−716.
(9) Diudea, M. V.; Pop, C. M. Molecular topology 27: A Shultz-type index based on the Wiener matrix. *Ind. J. Chem.* **1996**, *35A*, 257−261.
(10) Diudea, M. V. Walk Numbers $^{e}W_M$: Wiener-Type Numbers of Higher Rank. *J. Chem. Inf. Comput. Sci*, **1996**, *36*, 535−540.
(11) Diudea, M. V.; Katona G.; Minailiuc, O. M.; Parv, B. Molecular topology 24: Wiener and hyper Wiener indices in spiro-graphs. *Izvestia Akad. Nauk, Ser. Khim.* **1995**, 1674−1679.
(12) Bonchev, D.; Trinajstić, N. Information theory, distance matrix and molecular branching. *J. Chem. Phys.* **1977**, *67*, 4517−4533.
(13) Smolenskii, E. A.; Zefirov, N. S. Cutting trees and Classification of Alkanes. *Doklady Chem.* **2001**, *379*, 774−780.
(14) Gutman, I. A new Method for the Calculation of the Wiener Number of Acyclic Molecules. *J. Mol. Struct. (THEOCHEM)* **1993**, *285*, 137−142.
(15) Canfield, E. R.; Robinson, R. W.; Rouvray, D. H. Determination of the Wiener Molecular branching Index to the General Trees. *J. Comput. Chem*. **1985**, *6*, 598−609.
(16) Diudea, M. V. Layer Matrices in Molecular Graphs. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1064−1078.
(17) Bonchev, D.; Markel, E.; Dekmezian, A. Topological Analysis of Long-Chain Branching Patterns in Polyolefins. *J. Chem. Inf. Comput. Sci*. **2001**, *41*, 1274−1285.
(18) Randić, M. Generalized Molecular Descriptors. Generalized Molecular Descriptors. *J. Math. Chem.* **1991**, *7*, 155−168.
(19) Platt, J. R. Prediction of Isomeric Differences in Paraffin Properties. *J. Phys. Chem*. **1952**, *56*, 328−336
(20) Randić, M.; Wilkins, C. L. Graph Theoretical Ordering of Structures as a Basis for Systematic Searches for Regularities in Molecular Data. *J. Phys. Chem*. **1979**, *83*, 1525−1540.
(21) Needham, D. E.; Wei, I. C.; Seybold, P. C. Molecular Modeling of the Physical Properties of the Alkanes. *J. Am. Chem. Soc*. **1988**, *110*, 4186−4194.

CI030026K