can be described by the language.

So many interesting but difficult problems have to be tackled and solved to achieve the goal of a synthesis planning system.[19] Although KASP has been developed for a specific reaction database (RDB), the system's implemented facilities such as generic term manipulation and its matching mechanism should be useful for other similar systems.

ACKNOWLEDGMENT

This research is based on a reaction database that is a result of the working group for a reaction database system in Molecular Design Society in Japan, and I am grateful to Prof. Fujiwara of the University of Tsukuba, Dr. Matsuura of the Chuugai Corporation Company, and other members of the society.

REFERENCES AND NOTES

(1) Corey, E. J.; Wipke, W. T. Computer-assisted design of complex organic syntheses. *Science* **1969**, *166*, 178-192.
(2) Gasteiger, J. Computer-assisted synthesis design. *Chim. Ind. (Milan)* **1982**, *64*, 714-721.
(3) Hendrickson, J. B.; Braun-Keller, E.; Toczko, G. A. A Logic for Synthesis Design. *Tetrahedron* **1981**, *37* (Supplement 1), 359-370.
(4) Pensak, D. A.; Corey, E. J. LHASA-Logic and Heuristics Applied to Synthetic Analysis. *Computer-Assisted Organic Synthesis*; ACS Symposium Series 61; American Chemical Society: Washington, DC, 1978; pp 1-32.
(5) Wipke, W. T.; Ouchi, G. I.; Krishnan, S. Simulation and Evaluation of Chemical Synthesis-SECS. *Artif. Intell.* **1978**, *11*, 173-193.
(6) Lee, T. V. Expert Systems in Synthesis Planning: A User's View of the LHASA Program. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 259-272.
(7) Gund, P.; Grabowski, E. J. J.; Hoff, D. R.; Smith, G. M. Computer-Assisted Synthetic Analysis at Merck. *J. Chem. Inf. Comput. Sci.* **1980**, *20*, 88-93.
(8) Braun, H. W. Private communication.
(9) Yanaka, M.; Tomonaga, A. Organic Synthesis Design Program (SECS) and Its Knowledge Base Building. *The 8th Symposium on Chemical Information*, Kanazawa, Japan; Chemical Society of Japan: Tokyo, 1985; pp 138-141.
(10) Yanaka, M.; Kurumisawa, A.; Nakamura, K.; Wipke, W. T. Automatic Knowledge Base Building for the Organic Synthesis Design Program (SECS). *The 11th Symposium on Chemical Information*, Kyoto, Japan; Chemical Society of Japan: Tokyo, 1988; pp 72-75.
(11) Yanaka, M. An Application of SECS: Automatic Building of Knowledge Base. *Mol. Des.* **1989**, *11*, 2-57.
(12) Yanaka, M.; Nakamura, K.; Kurumisawa, A.; Wipke, W. T. Automatic Knowledge Base Building for the Organic Synthesis Design Program (SECS). *QSAR: Quantitative Structure-Activity Relationships in Drug Design*; Fauchere, J. L., Ed.; Alan R. Liss: New York, 1989; pp 147-150.
(13) The work of building RDB has been performed by the working group for a reaction database system that is an informal organization consisting of 16 company and university members of Molecular Design Society in Japan.
(14) Bersohn, M.; Mackay, K. Step toward the Automatic Compilation of Synthetic Organic Reactions. *J. Chem. Inf. Comput. Sci.* **1979**, *19*, 137-141.
(15) Matsuura, I. Computer-Assisted Reaction Design System Based on Reaction Database. *Mol. Des.* **1986**, *7*, 2-31.
(16) Matsuura, I. Development of PC-SYNTREX, A CAD System for Organic Synthetic Route Based on Reaction Database. *The 13th Symposium on Chemical Information*, Toyohashi, Japan; Chemical Society of Japan: Tokyo, 1990; pp 17-20.
(17) Structuring the transform base is not discussed in this paper.
(18) Corey, E. J. Computer-Assisted Analysis of Complex Synthetic Problems. *Chem. Soc. Rev.* **1971**, *25*, 455-482.
(19) Corey, E. J.; Long, A. K.; Rubinstein, S. D. Computer-Assisted Analysis in Organic Synthesis. *Science* **1985**, *228*, 408-418.

General Formulas for the Wiener Index

ISTVÁN LUKOVITS

Central Research Institute for Chemistry, Hungarian Academy of Sciences, P.O. Box 17, H-1525 Budapest, Hungary

Received February 8, 1991

The Wiener Index W of a graph is most often calculated by numerical methods. In this paper a method is proposed to deduce analytical formulas for W and for the connectedness index. It is shown that expressions for W can not include fourth or higher order terms in the number of vertices of the constituent strings. Formulas have been derived for graphs containing one, two, and three strings, and the general rules based on these formulas are discussed. Several numerical examples are given.

INTRODUCTION

The Wiener Index W is the sum of topological distances in the graph representing the structural formula of a molecule:[1]

$$W = \sum_{i<j} d_{ij} \tag{1}$$

where d_{ij} denotes the topological distance between atoms i and j ($i, j = 1, 2, ..., N$; N denotes the number of atoms, H atoms are neglected). The distances between all $N(N-1)/2$ pairs of atoms must be added. It has been found that in a series of alkane isomers W accounts for the compactness of the molecules,[1,2] the smaller the value of W the more compact the molecule.

W has been used to explain the variation in the physical and chemical properties of alkanes.[3-8] It has also been used to correlate the pharmacological properties of compounds with their structure.[9-11] Molecular branching[12] and molecular cyclicity[13,14] are topological characteristics that are properly accounted for by the Wiener Index. Congruence relationships for the Wiener Index of tree graphs (i.e., acyclic molecules) have also been investigated.[15] Algorithms have been proposed[16,17] supporting the reconstruction of the underlying graph from the actual value of the Wiener Index. Most applications of the Wiener Index are related to saturated hydrocarbons, but attempts have been undertaken to extend its definition for unsaturated hydrocarbons[18,19] and for heteroatoms.[20,21] For a series of molecules possessing a common parent structure, W could be decomposed into three portions. These are related to the parent structure, the substituents, and the interaction terms between the latter.[11]

Most often, numerical methods are used to calculate W. The first method suggested by Wiener,[1] can be used only for trees. The easiest way to obtain W is by eq 1, and the majority of the methods may be classified by the approach used to construct the matrix of distances. d_{ij} may be obtained by using the adjacency matrix approach,[22] the method of linked lists,[23] the algorithm by Müller et al.[24] or the Warshall algorithm.[25] Analytical expressions may be used to obtain W directly. The formulas were used to study the effects of structural variations and to find different structures with an identical W. Explicit formulas have been derived for chains,[1] simple cycles,[26] cyclic structures having acyclic branches,[27] spiro systems,[28] and polymers.[13,14] A recursion formula has been proposed for general trees.[21] The expected value of W for random benzenoid chains was also derived.[29]

The aim of the present work was to propose a numerical method, allowing deduction of an explicit formula for any structure. The next section introduces the necessary concepts of graph theory, and two theorems on which our method is based will be proved. The third section describes the numerical procedure used to obtain the analytical expressions. The expressions derived for several classes of structures are listed in the last section. In addition, formulas for the atomic connectedness indices[30] (see below) were also obtained, and the rules observed in these formulas will be discussed.

THEORY

The terms "molecular graph", "graph", and "structural formula" will be used interchangeably in this paper. The vertices v_i of the graph are the atoms, and the edges x_j are the bonds. The degree of a vertex (atom) v_i is the number of edges at v_i. By a string we shall denote a sequence of linked vertices. The degree of the first and of the last vertex of a string is either 1 or ≥ 3; the degree of the other vertices is equal to 2. The letters k, m, and n denote the number of vertices including the first and the last vertex. Note that branching points are taken into account more than once, so the total number of vertices N is

$$N = k + m + n - c \qquad (2)$$

where c is a constant which is equal to 1 in graphs 3 and 4; it is equal to 2 in graphs 5–7; and it is equal to 4 in graph 8 (Figure 1). The letters k, m, and n will also be used to label the actual strings; the same letter cannot be used to label different strings. Two strings beginning at the same vertex are said to be adjacent. There are four types of strings:

1. The simplest string is a chain in which the degree of the first vertex and of the last vertex is equal to 1 (graph 1 in Figure 1). A trivial string is a single vertex.
2. The degree of the first vertex is 1, and the degree of the last vertex is ≥ 3, or vice versa. In this case we have an acyclic branch, e.g., k in graph 3 (Figure 1).
3. The degree of the first vertex and of the last vertex is ≥ 3. The string may either belong to an acyclic subgraph (e.g., m in graph 7) or may be all or part of a cycle (e.g., k in graphs 7 and 8).
4. In a simple cycle (graph 2) the degree of each vertex is equal to 2.

The connectedness[30] s_i of a vertex v_i is

$$s_i = \sum_j d_{ij} \qquad (3)$$

The sum of the connectedness indices of a graph is equal to $2W$:

$$2W = \sum_i s_i \qquad (4)$$

In chains the connectedness of the endpoints is

$$s_1 = (k^2 - k)/2 \qquad (5)$$

In a cycle the connectedness of each vertex is

$$s_j = k^2/4 \qquad (j = 1, 2, ..., N, \text{ and } N \text{ is even}) \qquad (6)$$

and

$$s_j = (k^2 - 1)/4 \qquad (j = 1, 2, ..., N, \text{ and } N \text{ is odd}) \qquad (7)$$

In some cases s_j may be used to calculate W. It is easy to show that if subgraphs A and B share one common point, and

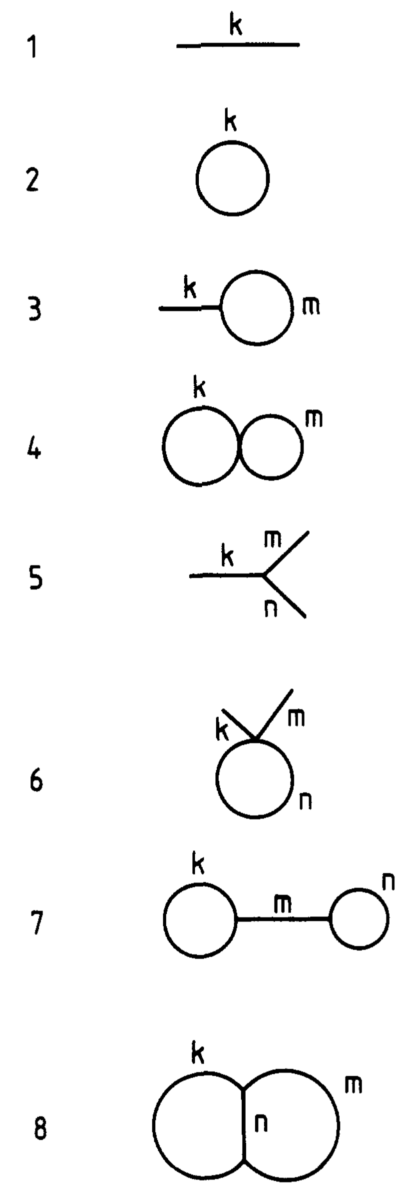

Figure 1. All graphs including not more than three strings. The degrees of the vertices are ≤ 4. The letters indicate the number of the vertices in a string.

W_A, W_B, s_A, and s_B are known[11] (s_A and s_B are related to this common point), then:

$$W = W_A + W_B + (N_A - 1)s_B + (N_B - 1)s_A \qquad (8)$$

where N_A and N_B denote the number of vertices in graphs A and B ($N = N_A + N_B - 1$), respectively.

The numerical procedure proposed in this paper is based on two theorems. These will be proved below. The proofs remain valid if there are more than three strings.

Theorem 1. The function $s(k,m,n)$ may include, at most, quadratic terms in k, m, and n.

Proof. 1. Let us first assume that the graph is acyclic. In this case there is only one path connecting vertex v_j with vertex v_i ($i \neq j$). Let us start the walk from an endpoint v_i (i.e., at a vertex whose degree is 1) toward vertex v_j. The contribution of this walk to s_j is equal to $(x^2 - x)/2$ by eq 5, where x denotes the number of vertices in the walk. If there is a branching point v_h between v_i and v_j, the contribution of the side branch to s_j is $(y^2 - y)/2 + (y - 1)d_{jh}$, provided that there are no more side branches. Here y denotes the number of vertices in the side branch. d_{jh} can only include linear terms in k, m, and n, by definition. If there are more side branches, the same procedure may be repeated.

Table I. Analytical Expressions for the Wiener Index

graph[a]	remarks	W
1		$(k^3 - k)/6$
2a	k is even	$k^3/8$
2b	k is odd	$(k^3 - k)/8$
3a	m is even	$(4k^3 + 3m^3 + 12k^2m + 6km^2 - 12k^2 - 6m^2 - 12km + 8k)/24$
3b	m is odd	$(4k^3 + 3m^3 + 12k^2m + 6km^2 - 12k^2 - 6m^2 - 12km + 2k - 3m + 6)/24$
4a	k and m are even	$(k^3 + m^3 + 2k^2m + 2km^2 - 2k^2 - 2m^2)/8$
4b	k is even, m is odd	$(k^3 + m^3 + 2k^2m + 2km^2 - 2k^2 - 2m^2 - 2k - m + 2)/8$
4c	k and m are odd	$(k^3 + m^3 + 2k^2m + 2km^2 - 2k^2 - 2m^2 - 3k - 3m + 4)/8$
5		$(k^3 + m^3 + n^3 + 3k^2m + 3k^2n + 3km^2 + 3m^2n + 3kn^2 + 3mn^2 - 6k^2 - 6m^2 - 6n^2 - 6km - 6kn - 6mn + 5k + 5m + 5n)/6$
6a	n is even	$(4k^3 + 4m^3 + 3n^3 + 12k^2m + 12k^2n + 12km^2 + 12m^2n + 6kn^2 + 6mn^2 - 24k^2 - 24m^2 - 12n^2 - 24km - 12kn - 12mn + 20k + 20m)/24$
6b	n is odd	$(4k^3 + 4m^3 + 3n^3 + 12k^2m + 12k^2n + 12km^2 + 12m^2n + 6kn^2 + 6mn^2 - 24k^2 - 24m^2 - 12m^2 - 24km - 12kn - 12mn + 14k + 14m - 3n + 12)/24$
7a	k and n are even	$(3k^3 + 4m^3 + 3n^3 + 6k^2m + 6k^2n + 12m^2n + 12km^2 + 6kn^2 + 6mn^2 + 24kmn - 12k^2 - 24m^2 - 12n^2 - 36km - 24kn - 36mn + 24k + 44m + 24n - 24)/24$
7b	k is even, n is odd	$(3k^3 + 4m^3 + 3n^3 + 6k^2m + 6k^2n + 12m^2n + 12km^2 + 6kn^2 + 6mn^2 + 24kmn - 12k^2 - 24m^2 - 12n^2 - 36km - 24kn - 36mn + 18k + 38m + 21n - 12)/24$
7c	k and n are odd	$(3k^3 + 4m^3 + 3n^3 + 6k^2m + 6k^2n + 12m^2n + 12km^2 + 6kn^2 + 6mn^2 + 24kmn - 12k^2 - 24m^2 - 12n^2 - 36km - 24kn - 36mn + 15k + 32m + 15n)/24$
8a	k, m, and n are even; or k, m, and n are odd; $n \le k$, $n \le m$	$(k^3 + m^3 + 2n^3 + 2k^2m + 3k^2n + 3m^2n + 2km^2 + kn^2 + mn^2 + 4kmn - 10k^2 - 10m^2 - 8n^2 - 12km - 16kn - 16mn + 28k + 28m + 28n - 32)/8$
8b	k is odd, m and n are even; $n \le k$, $n \le m$	$(k^3 + m^3 + 2n^3 + 2k^2m + 3k^2n + 3m^2n + 2km^2 + kn^2 + mn^2 + 4kmn - 10k^2 - 10m^2 - 8n^2 - 12km - 16kn - 16mn + 27k + 26m + 27n - 26)/8$
8c	k and m are odd, n is even; $n \le k$, $n \le m$	$(k^3 + m^3 + 2n^3 + 2k^2m + 3k^2n + 3m^2n + 2km^2 + kn^2 + mn^2 + 4kmn - 10^{k2} - 10m^2 - 8n^2 - 12km - 16kn - 16mn + 26k + 27m + 27n - 26)/8$
8d	k and m are odd, n is even; or k and m a . even, n is odd; $n \le k$, $n \le m$	$(k^3 + m^3 + 2n^3 + 2k^2m + 3k^2n + 3m^2n + 2km^2 + kn^2 + mn^2 + 4kmn - 10k^2 - 10m^2 - 8n^2 - 12km - 16kn - 16mn + 25k + 25m + 30n - 24)/8$

[a] Figure 1.

2. Let us assume that our graph is not acyclic. For each v_j, a spanning subgraph can be constructed by removing those edges which are not contained in a path connecting v_j with the rest of the vertices. In this way we have obtained a tree, since there is only one path connecting v_j with v_i.

Theorem 2. The function $W(k,m,n)$ may include at most cubic terms in k, m, and n.

Proof. The statement is clearly valid for simple chains and cycles.[1,26] Let us assume that the statement is valid for graph A. Graph A may be extended to yield A_1 by attaching a single string k to it at vertex v_1. v_1 belongs to k and as well as to graph A. By eq 5 the Wiener Index W_1 of A_1 will not include fourth or higher order terms. In the same way graph A_2 may be obtained by attaching string k to graph A at v_2. The Wiener Index W_2 of A_2 may also not include higher order terms. Let us construct a third graph A_3 by connecting the endpoint of k in A_1 with v_2, the Wiener Index of A_3 is W_3. Now $W_3 < W_1$ and $W_3 < W_2$, since A_3 is more compact than A_1 or A_2, and W_3 depends on the same parameters k, m, and n as W_1 and W_2. If W_3 contained terms raised to a higher power, it would be possible to construct A_3 with $W_3 > W_1$ and $W_3 > W_2$, or $W_3 < 0$, since W_1 and W_2 can include at most cubic terms in k, m, and n. Thus it is not possible that W_3 include higher than cubic terms in k, m, and n. All graphs can be constructed by using this technique, except those graphs to which a cycle is added at vertices v_1 and v_2. However, the resulting graphs A_1 and A_2 cannot be altered to yield A_3 without introducing a new string.

CALCULATIONS

The derivation of W for graph 8 (Figure 1) was done by solving a system of linear equations. If the graph were composed of two strings, then W would be equal to

$$W = a_1k^2 + a_2m^2 + a_3km + a_4k + a_5m + a_6 \quad (11)$$

where a_i (i = 1, 2, 3, 4, 5, 6) represents the unknown constants that are to be determined. In general if there are p strings, $1 + p + p(p + 1)/2 + p(p + 1)(p + 2)/6$ equations will be needed to obtain the expression for W. This estimate is, however, the upper limit, because some terms might not appear. For the other graphs eq 8 was used.

The linear equations approach was used to obtain the expressions of connectedness s_j for selected points in graphs (Figure 2). The number of necessary equations is $1 + p + p(p + 1)/2$ in this case. Again, this expression is the upper limit, since the actual number of equations needed to obtain s_j may be lower. The actual equations used to derive the formulas were not given, since the results do not depend on the particular values of the parameters k, m, and n.

Expressions were derived for all graphs containing no more than three strings. The degrees of the branching points were equal to 3 or 4. However, the results are valid for more complex cases, too.

RESULTS AND DISCUSSION

Figure 1 shows all graphs which may be composed by combining one, two, or three strings. The degrees of the vertices are ≤4. Table I lists the respective formulas obtained for W. Formulas have been published,[1,21,26–28] sometimes in a different form, for graphs 1–6 and for special cases[14] of graph 8. Figure 2 shows some selected points in these graphs. The respective formulas of the connectedness index were listed in Table II. No formulas for s but the simplest cases have been published to date.

It seems that the concept of strings may simplify the classification of chemical structures. The graphs could be ordered by the number of the constituent strings, number of cycles, and number of branching points. The concept of strings may also be used to explain the rules discussed below.

Let us consider graph 5 (Figure 1) and try to derive an expression for W. W must contain $(k^3 - k)/6$, since the sum of distances between the vertices of this string is given by this formula.[1] Similarly, the contributions of strings m and n are $(m^3 - m)/6$ and $(n^3 - n)/6$, respectively. Next we have to add all distances between the vertices of string k and of string m. This contribution is equal to $[(k + m - 1)^3 - (k + m - 1)]/6 - (k^3 - k)/6 - (m^3 - m)/6$. The second and third terms must be subtracted, because the first term already includes all intrastring distances. Thus, the interaction terms will not include

Table II. Analytical Expressions for the Connectedness Index

graph[a]	remarks	s
1		$(k^2 - k)/2$
2a	k is even	$k^2/4$
2b	k is odd	$(k^2 - 1)/4$
3a	m is even	$(2k^2 + m^2 + 4km - 6k - 4m + 4)/4$
3b	m is odd	$(2k^2 + m^2 + 4km - 6k - 4m + 3)/4$
4a	$m + n$ is even, $m \geq n$	$(2k^2 + m^2 + n^2 + 4kn + 2mn - 6k - 4m - 8n + 8)/4$
4b	$m + n$ is odd, $m \geq n$	$(2k^2 + m^2 + n^2 + 4kn + 2mn - 6k - 4m - 8n + 7)/4$
5		$(k^2 + m^2 + n^2 + 2km + 2kn - 5k - 3m - 3n + 4)/2$
6a	$m + n$ is even, $m \geq n$, k is even	$(k^2 + m^2 + n^2 + 4kn + 2mn - 4k - 4m - 8n + 8)/4$
6b	$m + n$ is even, $m \geq n$; k is odd	$(k^2 + m^2 + n^2 + 4kn + 2mn - 4k - 4m - 8n + 7)/4$
6c	$m + n$ is odd, $m \geq n$; k is even	$(k^2 + m^2 + n^2 + 4kn + 2mn - 4k - 4m - 8n + 7)/4$
6d	$m + n$ is odd, $m \geq n$; k is odd	$(k^2 + m^2 + n^2 + 4kn + 2mn - 4k - 4m - 8n + 6)/4$
7a	n is even	$(2k^2 + 2m^2 + n^2 + 4km + 4kn -10k - 6m - 4n + 8)/4$
7b	n is odd	$(2k^2 + 2m^2 + n^2 + 4km + 4kn - 10k - 6m - 4n + 7)/4$

[a] Figure 2.

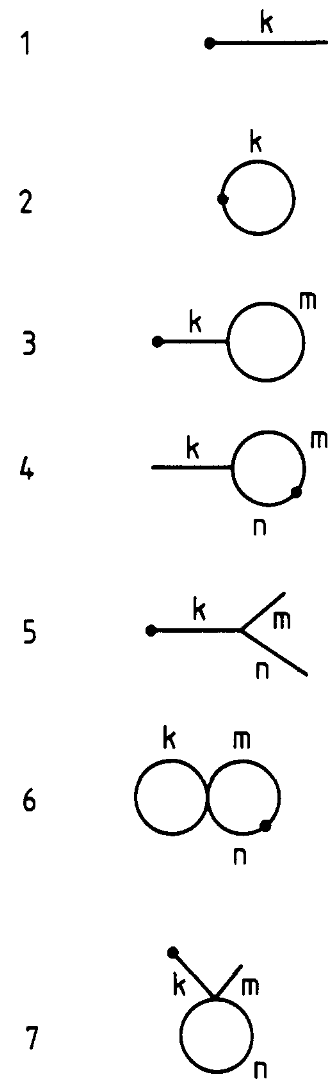

Figure 2. Some selected points, the connectedness of which is given in Table II.

pure cubic terms (e.g., k^3, m^3, etc.). Similar contributions are due to strings k and n and due to strings m and n. The sum of these contributions is equivalent with formula 5 (Table I). The cubic term kmn does not appear in formula 5, because there is no path encompassing all three strings. The same logic is valid for nonadjacent strings, too. If k and m are nonadjacent, there will be a term

$$(k + ... + m - c)^3 = (k + ... + m)^3 - 3(k + ... + m)^2c + 3(k + ... + m)c^2 - c^3 \quad (12)$$

since in connected graphs there is always a path leading from string k to string m. A kmn term cannot occur if there is no path including strings k, m, and n.

Using these simple arguments, the rules below may be explained easily (Table I). Rules which have not been proved are denoted with an asterisk.

Rule 1. In acyclic graphs the denominator is equal to 6.

Rule 2. The denominator is equal to 24 if the graphs contain both cyclic and acyclic subgraphs.

*Rule 3.** The denominator is equal to 8 if the underlying graphs include cycles only.

Rule 4. The coefficients of all cubic terms are positive.

Rule 5. The coefficients of quadratic terms are negative.

Rule 6. All possible cubic terms involving one or two letters contribute to W (e.g., k^3, m^3, k^2m, etc.).

Rule 7. Terms involving three letters (e.g., kmn) appear if there is a path which involves the respective strings.

Rule 8. The numerical value of the coefficients of terms k^3, m^3, and n^3 is 1/6, if k, m, and n do not belong to a cycle.

*Rule 9.** The numerical value of terms k^3, m^3, and n^3 is $r/8$ if k, m, and n do belong to a cycle, where r denotes the number of cycles to which k, m, or n belong. It is easy to show that the expressions given for simple chains and cycles remain valid if $k = 1$. The Wiener index is equal to 0 for a trivial graph. Thus a trivial graph is the simplest chain or the simplest cycle. Similarly, all graphs in Figure 1 can be shrunk to yield a trivial graph. If k, m, n, etc. are equal to 1, W will be 0.

Rule 10. The sum of the coefficients (including the constant) is 0 in formulas 1, 2b, 3b, 4c, 5, 6b, 7c, and 8a (Table I).

Similar rules could also be derived for the connectedness index s (Table II).

Rule 11. The denominator for acyclic graphs is equal to 2.

Rule 12. The denominator for graphs containing cyclic subgraphs is equal to 4.

Rule 13. The coefficients of the quadratic terms are positive.

Rule 14. The coefficients of the linear terms are negative.

Rule 15. The coefficient of a pure quadratic term (e.g., k^2, m^2, etc.) is 1/2, if the string does not belong to a cycle.

Rule 16. The coefficient of a pure quadratic term is 1/4, if the string does belong to a cycle.

Table III. Pairs of Structures with an Identical Wiener Index

	first structure				second structure				
no.	code[a]	k	m	n	code[a]	k	m	n	W
1	1	10			2	11			165
2	1		none		3		none		
3	1	12			4	8	7		286
4	1	26			5	10	10	10	2925
5	1	11			6	5	5	4	220
6	1	11			7	5	4	5	220
7	1	11			8	6	6	5	220
8	2	7			3	2	6		42
9	2		none		4		none		
10	2	15			5	8	6	2	420
11	2	18			6	9	5	6	729
12	2	10			7	3	3	6	125
13	2	14			8	8	7	4	343
14	3	7	14		4	15	7		966
15	3	4	3		5	3	3	2	31
16	3	2	5		6	2	2	4	26
17	3	5	14		7	6	8	5	689
18	3	6	3		8	5	5	3	78
19	4	6	5		5	5	4	2	108
20	4	11	7		6	8	4	6	507
21	4	14	4		7	8	7	3	550
22	4	10	7		8	10	7	3	425
23	5	6	4	2	6	3	2	8	150
24	5	7	4	3	7	4	6	4	250
25	5	7	7	5	8	8	8	7	672
26	6	4	2	9	7	7	2	6	246
27	6	6	3	4	8	6	5	5	174
28	7	5	8	4	8	9	6	6	491

[a] See Figure 1.

Rule 17. The sum of the coefficients (including the constant) is 0 in formulas 1, 2b, 3b, 4a, 5, 6b, and 7a (Table II).

Rule 18. The connectedness of cutpoints is equal to the sum of the connectedness indices in the respective subgraphs, e.g., for graph 5 (Figure 1):

$$s = (k^2 - k)/2 + (m^2 - m)/2 + (n^2 - n)/2 \qquad (13)$$

For this reason the connectedness indices of other cutpoints of the graphs in Figure 1 were not listed in Table II.

The method presented in this paper, may be applied in a systematic way to obtain formulas for more complex structures. However, with an increasing number of strings the number of equations grows rapidly. The rules listed above are clearly not sufficient to derive a formula for any structure. It remains to be examined whether there is a general scheme that could produce the coefficients without using the proposed numerical approach. As analytical formulas seem to be most efficient if someone wants to discover different structures possessing the same topological index, it may be expected that analytical formulas will be derived for other topological invariants, too.[31]

NUMERICAL EXAMPLES

The formulas (Table I) are most convenient if someone wants to find different structures with an identical Wiener index. Table III lists several examples found for pairs of structures with $N \leq 30$. In two cases (nos. 2 and 9) the search was unsuccessful. Of course the search could have been repeated in order to find equivalent W within the same class of structures, too. According to Table III n-decane and cycloundecane (no. 1) are equally compact (this is the only equivalence found between n-alkanes and cycloalkanes). Similarly 3-methylpentane and propylcyclopropane (no. 15), 1,1-dimethylcyclobutane and methylcyclopentane (no. 16), and also 4-methyloctane and spiro[cyclohexane–cyclopentane] (no. 19) are also equally compact.

In a series of alkanes highly significant correlation was detected[31] between W and the logarithm of the 1-octanol/water partition coefficients (log P). For molecules being equally compact, it may be expected that the respective properties of the molecules are also similar, provided that the correlation is good enough. Example 8 (Table III) shows that the Wiener Index of cycloheptane and of methylcyclohexane is 42. The respective values of log P should be rather similar. In fact[8] log P of cycloheptane is 2.87 and log P of methylcyclohexane is 2.76.

REFERENCES AND NOTES

(1) Wiener, H. *J. Am. Chem. Soc.* **1947**, *69*, 17–20.
(2) Wiener, H. *J. Am. Chem. Soc.* **1947**, *69*, 2636–2638.
(3) Rouvray, D. H. *J. Comput. Chem.* **1987**, *8*, 470–480.
(4) Platt, J. R. *J. Phys. Chem.* **1952**, *56*, 328–336.
(5) Bošnjak, N.; Alder, N.; Perić, M.; Trinajstić, N. In *Modelling of Structure and Properties of Molecules*; Maksic, Z. B., Ed.; Ellis Horwood: Chichester, 1987; p 103.
(6) Seybold, P. G.; May, M. A.; Gargas, M. L. *Acta Pharm. Jugosl.* **1986**, *36*, 253–265.
(7) Mekenyan, O.; Bonchev, D.; Trinajstić, N. *Int. J. Quantum Chem.* **1986**, *18*, 369–380.
(8) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976; p 38.
(9) Basak, S. C.; Monsrud, L. J.; Rosen, M. E.; Frane, C. M.; Magnuson, V. R. *Acta Pharm. Jugosl.* **1986**, *36*, 81–95.
(10) Seybold, P. G. *Int. J. Quantum Chem., Quantum Piol. Symp.* **1983**, *No. 10*, 103–108.
(11) Lukovits, I. *J. Chem. Soc., Perkin Trans. 2* **1988**, 1667–1671.
(12) Bonchev, D.; Trinajstić, N. *J. Chem. Phys.* **1977**, *67*, 4517–4533.
(13) Bonchev, D.; Mekenyan, D.; Trinajstić, N. *Int. J. Quantum Chem.* **1980**, *17*, 845–893.
(14) Mekenyan, O.; Bonchev, D.; Trinajstić, N. *Int. J. Quantum Chem.* **1981**, *19*, 929–955.
(15) Gutman, I.; Rouvray, D. H. *Comput. Chem.* **1990**, *14*, 29–32.
(16) Baskin, I. I.; Gordeeva, E. V.; Devdariani, R. O.; Zefirov, N. S.; Paliulin, V. A.; Stankevich, M. I. *Dokl. Akad. Nauk SSSR* **1989**, *307*, 613–617.
(17) Kvasnička, V.; Pospichal, J. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 99–105.
(18) Lukovits, I. *Quant. Struct.-Act. Relat.* **1990**, *9*, 227–231.
(19) Tratch, S. S.; Stankevich, M. I.; Zefirov, N. S. *J. Comput. Chem.* **1990**, *11*, 899–908.
(20) Barysz, M.; Jashavi, G.; Lall, R. S.; Srivastava, W. K.; Trinajstić, N. In *Studies in Physical and Theoretical Chemistry*; King, R. B., Ed.; Elsevier: Amsterdam, 1983; p 222.
(21) Canfield, E. R.; Robinson, R. W.; Rouvray, D. H. *J. Comput. Chem.* **1985**, *6*, 598–609.
(22) Harary, F. *Graph Theory*; Addison-Wesley: Reading, 1969; p 151.
(23) Bersohn, M. *J. Comput. Chem.* **1983**, *4*, 110–113.
(24) Müller, W. R.; Szymanski, K.; Knop, J. V.; Trinajstić, N. *J. Comput. Chem.* **1987**, *8*, 170–173.
(25) Warshall, S. *J. Assoc. Comput. Math.* **1962**, *9*, 11–15.
(26) Bonchev, D.; Mekenyan, O.; Knop, J. V.; Trinajstić, N. *Croat. Chem. Acta* **1979**, *52*, 361–362.
(27) Mekenyan, O.; Bonchev, D.; Trinajstić, N. *Croat. Chem. Acta* **1983**, *56*, 237–240.
(28) Mekenyan, O.; Bonchev, D.; Trinajstić, N. *Math. Chem.* **1979**, *6*, 93–95.
(29) Gutman, I.; Kennedy, J. W.; Quintas, L. V. *Chem. Phys. Lett.* **1990**, *173*, 403–408.
(30) Seybold, P. G. *Int. J. Quantum. Chem., Quantum Biol. Symp.* **1983**, *10*, 95–102.
(31) Lukovits, I. *Rep. Mol. Theory* **1990**, *1*, 127–131.