The Range of the Wiener Index and Its Mean Isomer Degeneracy

I. Gutman
Faculty of Science, University of Kragujevac, Yugoslavia

L. Šoltés
Faculty of Mathematics and Physics, Comenius University, Bratislava, Czecho-Slovakia

Z. Naturforsch. **46a**, 865–868 (1991); received February 19, 1991

Sharp lower and upper bounds for the Wiener index (W) of a connected (n, m)-graph are reported; n = number of vertices, m = number of edges. The mean isomer degeneracy of W is estimated and is shown to unboundedly increase with increasing n. Thus the isomer-discriminating power of W is confirmed to be very low in the case of large molecules.

Introduction

The sum of all distances in a molecular graph is a structural characteristic of the carbon-carbon skeleton of organic molecules which attracted much attention of theoretical and mathematical chemists. This topological index was first introduced by Wiener in 1947 [1, 2] and, in a slightly different form, by Hosoya in 1971 [3] and by Rouvray in 1975 [4, 5]. It is usually called the *Wiener index* and is defined as

$$W = W(G) = \sum_{u, v} \mathrm{d}(u, v),$$

where $\mathrm{d}(u, v)$ is the distance (= the length of the shortest path) between the vertices u and v, and where the summation goes over all pairs of vertices of the (molecular) graph G. Recall that the graph G is assumed to be connected.

The Wiener index has been used to model a considerable number of physico-chemical properties of organic compounds, especially of alkanes. These are boiling points [1, 2], heats of isomerization and vaporization [6], critical constants [7], viscosity and surface tension [5], chromatographic retention times [8, 9] and ultrasonic sound velocity [10], to mention just a few. By means of the Wiener index it was possible to establish the mean conformation of long-chain alkanes near their boiling point [11]. Application of W in polymer science [12] and in the physical chemistry of solid state [13, 14] were also reported. Further details of the chemical applications of the Wiener index can be found in the reviews [15–17].

In the theory of topological indices [17] it is desirable to have a high isomer-discriminating power, namely that the topological index has different numerical values for isomeric species. Because none of the numerous topological indices put forward to date [17, 18] can completely distinguish between isomers, the so-called *mean isomer degeneracy* μ was proposed [19] as a measure of the isomer-discriminating power. Let N be the number of chemical isomers that possess non-isomorphic molecular graphs, and let t be the number of different values assumed by a certain topological index I. Then [17, 19]

$$\mu = N/t\,. \tag{1}$$

Since the molecular graphs of all isomers have equal numbers of vertices (n) and edges (m), it is evident that N can be considered as a function of the graph invariants n and m. Consequently, μ is fully determined by n, m, and the actual form of the topological index I.

If the topological index I is capable of fully characterizing the molecular (n, m)-graphs, then $\mu = \mu(n, m, I)$ is equal to unity. Otherwise, the mean isomer degeneracy of I is greater than one. In the case when $I = W$, empirical studies revealed that $\mu(n, m, W)$ assumes quite large values [18–22]. In the present work we support these empirical findings by appropriate mathematical arguments. In particular, we demonstrate that with increasing value of n, the mean isomer degeneracy of the Wiener index becomes arbitrarily large and tends to infinity.

In what follows μ will always denote the mean isomer degeneracy of the Wiener index, for the class of (molecular) graphs with n vertices and m edges.

Reprint requests to Prof. Dr. Ivan Gutman, Faculty of Science, P.O. Box 60, YU-34000 Kragujevac, Yugoslavia.

0932-0784 / 91 / 1000-0865 $ 01.30/0. – Please order a reprint rather than making your own copy

Bounds for the Wiener Index

The basic mathematical properties of the Wiener index were established in a number of recently published works [23–27]. It is well known (and easy to prove) that for any connected graph G with n vertices

$$W(K_n) = \binom{n}{2} \leq W(G) \leq W(P_n) = \binom{n+1}{3}, \quad (2)$$

where K_n and P_n denote the complete graph and the path, respectively, with n vertices. Equality signs in the relations (2) occur only if $G = K_n$ and $G = P_n$, respectively.

The above bounds hold for all connected graphs with n vertices, irrespective of the actual number of edges. If the consideration is restricted to classes of graphs with n vertices and m edges (with m being a fixed number lying between $n-1$ and $n(n-1)/2$), then the following two results have been obtained [23, 28]. Recall that the diameter of a (connected) graph is the maximum distance between any two of its vertices.

(a) *Among connected (n, m)-graphs, the graphs having minimal diameter have minimal Wiener indices* [23]. (Note that if $m < n(n-1)/2$, then the minimal diameter is equal to 2, whereas in the case of the complete graph, $m = n(n-1)/2$, the diameter is equal to 1.)

(b) *Among connected (n, m)-graphs, the graph having maximal Wiener index has also maximal diameter* [28]. (The (n, m)-graph having maximal diameter is the so-called path-complete graph, $PK_{n,m}$ [29], whose structure is described below.)

We now point at an interesting property of the minimal-diameter (n, m)-graphs: *All such graphs have equal Wiener indices.*

Really, in a minimal-diameter (n, m)-graph there are m pairs of adjacent vertices (whose distance is 1) and $\binom{n}{2} - m$ pairs of non-adjacent vertices (whose distance is 2). Therefore the Wiener index of such a graph is equal to

$$1 \times m + 2 \times \left[\binom{n}{2} - m\right] = n(n-1) - m \quad (3)$$

and is thus fully determined by n and m.

The maximum-diameter (n, m)-graph $PK_{n,m}$ is constructed according to the following recipe [29]. Let $c = m - n + 1$ be the cyclomatic number of the (n, m)-graphs considered. Denote by P_n the path with n vertices and label its vertices by $1, 2, \ldots, n$ so that the vertices i and $i+1$ are adjacent, $i = 1, 2, \ldots, n-1$. Recall that $W(P_n) = \binom{n+1}{3}$.

If $c = 0$, then $PK_{n,m} = P_n$.

If $c > 0$, then $PK_{n,m}$ is obtained by adding c new edges to P_n, connecting the first c pairs of vertices in the below scheme:

(1, 3)
(1, 4) (2, 4)
(1, 5) (2, 5) (3,5)
...
(1, i) (2, i) ... ($i-2$, i)
(1, n) (2, n) ... ($n-2$, n).

For instance, if $c = 5$, then $PK_{n,m}$ is constructed by connecting in P_n the vertex pairs (1, 3), (1, 4), (2, 4), (1, 5) and (2, 5).

In order to calculate $W(PK_{n,m})$, observe that by introducing the edge (1, 3) the Wiener index of P_n will be diminished by $n-2$. By introducing either the edge (1, 4) or (2, 4) the W-value will be further diminished by $n-3$. In the general case, the introduction of each of the edges in the i-th row of the above scheme diminishes the W-value by $n-i-1$. Consequently,

$$W(PK_{n,m}) = \binom{n+1}{3} - [1 \cdot (n-2) + 2 \cdot (n-3) + \ldots + k \cdot (n-k-1)] - j \cdot (n-k-2),$$

where k and j are the (unique) integer solutions of the equation

$$k(k+1)/2 + j = c; \quad 0 \leq j \leq k.$$

A lengthy, but elementary algebraic manipulation leads then to the formula

$$W(PK_{n,m}) = \binom{n+1}{3} - \binom{k+2}{3} - (n-k-2)(m-n+1), \quad (4)$$

where

$$k = \text{int}\,[(\sqrt{8m-8n+9} - 1)/2] \quad (5)$$

and where int $[x]$ denotes the integer part of the number x.

Combining the formulas (3) and (4) we arrive at one of the main results of this paper: *Let G be a connected graph with n vertices and m edges and let the parameter k be given by means of* (5). *Then the Wiener index of G is bounded as follows:*

$$n(n-1) - m \leq W(G) \leq \binom{n+1}{3} - \binom{k+2}{3} - (n-k-2)(m-n+1). \quad (6)$$

The bounds (6) are sharp because there exist (n, m)-graphs for which the equality signs apply, namely the above described minimal- and maximal-diameter graphs.

Estimating the Mean Isomer Degeneracy of the Wiener Index

The bounds (6) imply that the number of distinct values which the Wiener index may assume cannot exceed t^*,

$$t^* = \left[\binom{n+1}{3} - \binom{k+2}{3} - (n-k-2)(m-n+1)\right] - [n(n-1)-m]+1 \,. \quad (7)$$

Evidently, t^* is an upper bound for the quantity t in (1).

In order to estimate the mean isomer degeneracy of the Wiener index we have to find also a lower bound for the number N of connected (n, m)-graphs. A very rough such bound is obtained by means of the following reasoning.

By Cayley's celebrated theorem, there are precisely n^{n-2} labeled trees with n vertices [30]. Since every tree can be labeled in at most $n!$ distinct ways, there are at least $n^{n-2}(n!)^{-1}$ (unlabeled) trees with n vertices.

Let G be a connected (n, m)-graph and $c = m-n+1$ be its cyclomatic number. The spanning trees of G are obtained by deleting c edges from G. Since c edges in G can be chosen in $\binom{m}{c} = \binom{m}{n-1}$ ways, the graph G has at most $\binom{m}{n-1}$ spanning trees. On the other hand, every n-vertex tree is a spanning tree of some (n, m)-graph. Consequently, there are at least $[n^{n-2}(n!)^{-1}]\binom{m}{n-1}^{-1}$ nonisomorphic connected (n, m)-graphs. In other words, the number N^*,

$$N^* = [n^{n-2}(n!)^{-1}]\binom{m}{n-1}^{-1} = n^{n-3}(m-n+1)!\,(m!)^{-1} \,, \quad (8)$$

is a lower bound for the quantity N in (1).

We thus arrived at a lower bound μ^* for the mean isomer degeneracy of the Wiener index:

$$\mu^* < \mu \,, \quad (9)$$

where $\mu^* = N^*/t^*$ and where t^* and N^* are given by (7) and (8), respectively.

Although the estimate (9) is quite inaccurate (and by no means sharp), it still can be used to demonstrate the following important property of the mean isomer degeneracy μ. Using elementary calculus it can be shown that for any fixed value of the cyclomatic number c,

$$\lim_{n\to\infty} \mu^* = \infty \,. \quad (10)$$

Bearing in mind the inequality (9), (10) implies: *For any fixed value of the cyclomatic number, the mean isomer degeneracy μ of the Wiener index becomes arbitrarily large as n tends to infinity.* Given integers $c \geq 0$, $d \geq 1$, there exist d-membered families of connected c-cyclic (n, m)-graphs (i.e., connected graphs on n vertices and $m = n+c-1$ edges) which all have the same Wiener index. Such families exist for all values of n greater than a certain critical value $n_0 = n_0(c, d)$.

We thus demonstrated that the isomer-discriminating power of the Wiener index becomes very low when the sizes of the molecular graphs considered are large. We hope that this finding (which by no means needs to be interpreted in a negative sense, but which certainly calls to some caution) will be taken into account in further physico-chemical studies based on the usage of the Wiener index.

[1] H. Wiener, J. Amer. Chem. Soc. **69,** 17 (1947).
[2] H. Wiener, J. Chem. Phys. **15,** 766 (1947).
[3] H. Hosoya, Bull. Chem. Soc. Japan **44,** 2332 (1971).
[4] D. H. Rouvray, Match **1,** 125 (1975).
[5] D. H. Rouvray and B. C. Crafford, South Afr. J. Sci. **72,** 47 (1976).
[6] H. Wiener, J. Amer. Chem. Soc. **69,** 2636 (1947).
[7] L. I. Stiel and G. Thodos, J. Amer. Inst. Chem. Ing. **8,** 527 (1962).
[8] D. Bonchev, O. Mekenyan, G. Protić, and N. Trinajstić, J. Chromatogr. **176,** 149 (1979).
[9] D. Papazova, M. Dimov, and D. Bonchev, J. Chromatogr. **188,** 297 (1980).
[10] D. H. Rouvray and W. Tatong, Z. Naturforsch. **41a,** 1238 (1986).
[11] D. H. Rouvray and R. B. Pandey, J. Chem. Phys. **85,** 2286 (1986).
[12] O. Mekenyan, S. Dimitrov, and D. Bonchev, Europ. Polymer J. **19,** 1185 (1983).
[13] O. Mekenyan, D. Bonchev, and H. G. Fritsche, Phys. Stat. Sol. **55A,** 181 (1979); **56A,** 607 (1979).

[14] D. Bonchev, O. Mekenyan, and H. G. Fritsche, J. Crystal Growth **49,** 90 (1980).
[15] D. H. Rouvray, Sci. Amer. **255,** 36 (1986).
[16] D. H. Rouvray, in: N. Trinajstić (ed.), Mathematical and Computational Concepts in Chemistry. Horwood, Chichester 1986, pp. 295–306.
[17] D. H. Rouvray, J. Comput. Chem. **8,** 470 (1987).
[18] A. T. Balaban, I. Motoc, D. Bonchev, and O. Mekenyan, Topics Curr. Chem. **114,** 21 (1983).
[20] M. Razinger, J. R. Chrétien, and J. E. Dubois, J. Chem. Inf. Comput. Sci. **25,** 689 (1985).
[21] I. Gutman and O. E. Polansky, Match **20,** 115 (1986).
[22] I. Gutman, S. Marković, D. Luković, V. Radivojević, and S. Rančić, Coll. Sci. Papers Fac. Sci. Kragujevac **8,** 15 (1987).
[23] J. Plesnik, J. Graph Theory **8,** 1 (1984).
[24] E. R. Canfield, R. W. Robinson, and D. H. Rouvray, J. Comput. Chem. **6,** 598 (1985).
[25] I. Gutman and O. E. Polansky, Mathematical Concepts in Organic Chemistry, Springer-Verlag, Berlin 1986, Chapter 11.1.
[26] O. E. Polansky and D. Bonchev, Match **21,** 133 (1986); **25,** 3 (1990).
[27] V. A. Skorobogatov and A. A. Dobrynin, Match **23,** 105 (1988).
[28] L. Šoltés, Math. Slovaka **41,** 11 (1991).
[29] F. Harary, Proc. Natl. Acad. Sci. US **48,** 1142 (1961).
[30] A. Cayley, Quart. Math. **23,** 376 (1889); cf. e.g., F. Harary and E. M. Palmer, Graphical Enumeration, Academic Press, New York 1973, Chapter 1.7.