On the Definition of the Hyper-Wiener Index for Cycle-Containing Structures

D. J. Klein,*,† I. Lukovits,*,‡ and I. Gutman§,⊥

Texas A&M University at Galveston, Galveston, Texas 77553-1675, Central Research Institute for Chemistry, Hungarian Academy of Sciences, H-1525 Budapest, P.O. Box 17, Hungary, and Institute of Physical Chemistry, Attila József University, P.O. Box 105, H-6701 Szeged, Hungary

Received May 18, 1994®

The hyper-Wiener index was recently introduced by Randić. The original definition given by Randić can be used for acyclic structures only. In this paper the definition of Randić was extended in two different fashions so as to be applicable for any connected structure. The formula provides an easy method to calculate the hyper-Wiener index for any graph.

INTRODUCTION

The hyper-Wiener index R has been defined by Randić[1] for any connected acyclic structure as a sum over all "pair contributions" C_{ij}, where subscripts i and j denote vertices i and j in the underlying graph. The pair contributions C_{ij} are obtained by means of the following two-step algorithm.

0. Let T be the graph of an acyclic molecule. Then T itself is acyclic, i.e., T is a tree. Any two vertices i and j of a tree are connected by a unique path, which we denote by $\pi = \pi(i, j)$.

1. Remove from T all vertices and edges that belong to π, except the vertices i and j themselves. Delete any side branch attached to π, except those attached to the vertices i and j. Then we obtain a subgraph T_π of T that is disconnected and possesses two components. We denote the number of vertices in the components of T_π by $n_{1,\pi}$ and $n_{2,\pi}$.

2. C_{ij} is equal to the product of the number of vertices of the two components of T_π, i.e., $C_{ij} = n_{1,\pi(i,j)}\, n_{2,\pi(i,j)}$. The hyper-Wiener index R is the sum of all $N(N - 1)/2$ contributions from distinct pairs of vertices, when N denotes the number of vertices in the graph. Example: n-butane, where the carbons are numbered consecutively and hydrogens are neglected: $C_{12} = 1 \times 3 = 3$, $C_{13} = 1 \times 2 = 2$, $C_{14} = 1 \times 1 = 1$, $C_{23} = 2 \times 2 = 4$, $C_{24} = 2 \times 1 = 2$, $C_{34} = 3 \times 1 = 3$, and $R_{\text{butane}} = 3 + 2 + 1 + 4 + 2 + 3 = 15$.

The hyper-Wiener index may be obtained by using the "Wiener matrix" method[2] or by using formulas that were derived for several classes of structures.[3] Both Randić et al.[2] and Lukovits and Linert[4] have sought to extend the original definition of the hyper-Wiener index to cyclic structures.

Hundreds of distinct graph invariants have been proposed as descriptors in quantitative structure−property relationships, e.g., as discussed in ref 2, but only a few of them have proven to be useful in this respect. From its definition it seems that the hyper-Wiener index measures the "expansiveness" of a graph, weighting more expansive graphs even more so than does the Wiener index, and it will be shown in this paper that there is a relationship between these indices. Randić has stated[1] that the hyper-Wiener index already has passed definite tests to classify it as useful.

In this paper we derive two formulae for R, both of which are equivalent to the definition of Randić, but can also be applied to any cycle-containing connected structure. These formulae also allow an efficient calculation of R.

DERIVATION OF THE FORMULA

In this paper expressions "graph" and "molecular structure", "bond" and "edge", "vertex" and "atom" are used interchangeably. For tree T, the index R may be expressed by means of eq 1

$$R = \sum_{\pi}^{\pi\epsilon T} n_{1,\pi}\, n_{2,\pi} \tag{1}$$

where the summation goes over all paths of T. It is easy to see that

$$n_{1,\pi}\, n_{2,\pi} = \#_\pi \tag{2}$$

counts the number of paths π' contained in T, that are "external" with respect to π (i.e., such that $\pi \subseteq \pi'$). For any π' of length $|\pi'| = n$ there are $n = m + \overline{1}$ "internal" subpaths π of length m, and altogether there are $n + (n - 1) +, ..., + 1 = n(n + 1)/2 = \#'_{\pi'}$ "internal" paths of length $\leq n$. Since any path in a tree is uniquely determined by its endpoints i and j, the path $\pi = \pi(i,j)$ may be represented by ij in eq 1, which combined with eq 2 results in

$$R = \sum_{i<j} \#_{ij} \tag{3}$$

Similarly $\#'_{\pi'} \equiv \#'_{ij}$ and

$$\#'_{ij} \equiv d_{ij}(d_{ij} + 1)/2 \tag{4}$$

where d_{ij} denotes the distance between the vertices i and j. It is easy to see that

$$\sum_{i,j} \#_{ij} = \sum_{i,j} \#^{r}_{ij'} \tag{5}$$

because each "external" path π' of length d_{ij} appears d_{ij} $(d_{ij} + 1)/2$ times, so that the counts of all "internal" paths and "external" paths are equal. From eqs 3, 4, and 5 it then follows that

† Texas A&M University at Galveston.
‡ Hungarian Academy of Sciences.
§ Attila József University.
⊥ On leave from Faculty of Science, Kragujevac.
® Abstract published in *Advance ACS Abstracts*, November 15, 1994.

0095-2338/95/1635-0050$09.00/0

$$R = \sum_{i<j} d_{ij}(d_{ij}+1)/2 = 1/2 \sum_{i<j}(d_{ij}^2 + d_{ij}) \qquad (6)$$

The summation on the second term in the parentheses gives rise to the Wiener index[5] W. The summation on the first term gives a (unnormalized) second moment of distance, so that denoting this by D_2, we obtain

$$R = (D_2 + W)/2 \qquad (7)$$

Therefore R is the average of the sum of the entries and the sum of the squared entries of the upper triangle of the distance matrix.

Equation 6 indicates a natural extension for the definition of the hyper-Wiener index for the case of connected cycle-containing structures.

In connection with the above more general definition of R, we wich to remind the reader that originally[5] the Wiener index W was also conceived only for trees. Hosoya[6] extended the Wiener-index-concept to all connected cycle-containing structures by (re)defining W as the sum of the entries of the upper triangle of the respective distance matrix. Equation 7 is a quite similar (and presumably equally natural) extension of the hyper-Wiener-index-concept.

With regard to eq 7 it is worth mentioning that some time ago Balaban[7] introduced and examined an invariant D,

$$D = \left[\binom{N}{2}^{-1} \sum_{i<j} d_{ij}^2\right]^{-1} \qquad (8)$$

that is closely related to D_2.

For the Wiener number for general cycle-containing graphs there is another extension which also applies to R. The so called "resistance distance"[8] Ω_{ij} between two vertices i and j is just the effective electrical resistance between i and j when unit resistors are placed on each bond. For trees $\Omega_{ij} = d_{ij}$, so that replacing d_{ij} by Ω_{ij} in eq 6 leads to a different extension (for general graphs).

For either case (of graph or resistance distances) the result of eq 6 further suggests a whole sequence of indices, namely the nth moments of the distribution of distances.

NUMERICAL EXAMPLES

Table 1 lists several values for simple cycles as obtained by using eq 6, by using the method by Lukovits and Linert[4] and by using the "resistance" distances.[8] It can be seen that the values by ref 4 are at least as great as those obtained by using eq 6, while that based on eq 6 with d_{ij} replaced by Ω_{ij} are even smaller (as must be the case for all cycle-containing graphs). The method given in ref 4, however, has been developed (thus far) only for trees and single cycles. All three methods, at least within this limitation, are legitimate extensions of the definition by Randić,[1] and the two methods developed here are perfectly general (with limitation only to connectedness). Based on eq 7 a formula for R of simple cycles could be derived. It is well-known that for cycles $W = N^3/8$ if N is even, and $W = (N^3 - N)/8$, if N is odd. Therefore it is enough if formulas for D_2 could be derived in simple cycles. In fact one may readily show that if N is odd

$$D_2 = N\{1^2 + 2^2 + ... + [(N-1)/2]^2\} = (N^4 - N^2)/24 \qquad (9)$$

and if N is even

Table 1. Hyper-Wiener Index of Simple N-Chains and N-Cycles

	chain	cycle		
N	R	R^a	R^b	R^c
1	0			
2	1			
3	5	3	3	$5/3 \simeq 1.67$
4	15	10	10	$37/8 \simeq 4.63$
5	35	25	20	$51/5 \simeq 10.20$
6	70	51	42	$469/24 \simeq 19.5$
7	126	94.5	70	34
8	210	160	120	$441/8 \simeq 55.1$

[a] Reference 4. [b] This work, eq 6. [c] This work via eq 6 with d_{ij} replaced by Ω_{ij}.

$$D_2 = N(2\{1^2 + 2^2 + ... + [(N-2)/2]^2\} + N^2/4)/2 = (N^4 + 2N^2)/24 \qquad (10)$$

Therefore by using eqs 7, 9, 10, and the expressions for W, we obtain

$$R_{\text{cycle}} = (N^4 + 3N^3 - N^2 - 3N)/48 \quad (N \text{ is odd}) \qquad (11)$$

and

$$R_{\text{cycle}} = (N^4 + 3N^3 + 2N^2)/48 \quad (N \text{ is even}) \qquad (12)$$

The formulas aid a search for pairs of structures with identical values of R. It was found that $R = 275$ for cyclodecane ($N = 10$), and this value is identical with R of 3-methyloctane ($N = 9$). Similarly the hyper-Wiener index of n-hexane ($N = 6$) is identical with the hyper-Wiener index of cycloheptane ($N = 7$), namely $R = 70$. This is also the value of the hyper-Wiener index of a star $K_{1,7}$. The hyper-Wiener index seems to be the measure of the "expansiveness" of the underlying molecule, as is the original Wiener index. However, through eq 7, large distances dominate in R. This is the reason why the values of R in cycles are much smaller than the values of the corresponding chains. Therefore R and W, even if they are related functionally (eq 7), account for different aspects of the structure of a molecule.

For the hyper-Wiener index based on the resistance distance[8] one can sometimes express the resistance distance Ω_{ij} between vertices i and j in terms of the ordinary graphical distance d_{ij}. For the N-cycle

$$\Omega_{ij} = d_{ij}(N - d_{ij})/N \qquad (13)$$

as may be readily obtained from the standard result for resistances d_{ij} and $N - d_{ij}$ in parallel. Thus the computation of the resistance distance hyper-Wiener index R' devolves (for an N-cycle) also to the determination of moment sums of the d_{ij}. We find

$$R' = 1/2\left(D_1 + \frac{N-1}{N}D_2 - \frac{2}{N}D_3 + \frac{1}{N^2}D_4\right) \qquad (14)$$

where $D_1 = W$ and D_2 are given in eqs 9 and 10. The higher moment sums D_3 and D_4 may be developed much as was done for D_2, whence for odd N we obtain

$$D_3 = N(N^2 - 1)^2/64 \qquad (15)$$

$$D_4 = N^2(N^2 - 1)(3N^2 - 7)/480 \qquad (16)$$

For even N, the quantity D_i/N is just the average of the preceding and succeeding odd N quantities D_i/N.

ACKNOWLEDGMENT

One of the authors (D.J.K.) acknowledges support from The Welch Foundation of Houston, Texas. I.G. thanks the Mathematical Institute in Belgrade for financial support.

REFERENCES AND NOTES

(1) Randić, M. Novel Molecular Description for Structure-Property Studies. *Chem. Phys. Lett.* **1993**, *211*, 478–483.
(2) Randić, M.; Guo, X.; Oxley, T.; Krishnapriyan, H. J. Wiener Matrix: Source of Novel Graph Invariants. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 709–716.
(3) Lukovits, I. Formulas for the Hyper-Wiener Index. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1079–1081. Lukovits, I. A Note on a Formula for the Hyper-Wiener Index of Some Trees. *Comput. Chem.*, in press.
(4) Lukovits, I.; Linert, W. A Novel Definition of the Hyper-Wiener Index for Cycles. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 899–902.
(5) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, *69*, 17–20. Wiener, H. Correlation of Heats of Isomerization and Differences in Heats of Vaporization of Isomers. *J. Am. Chem. Soc.* **1947**, *69*, 2636–2638.
(6) Hosoya, H. Topological Index. A newly Proposed Quantity Characterizing the Topological Nature of Structural Isomers of Saturated Hydrocarbons. *Bull. Chem. Soc. Jpn.* **1971**, *44*, 2332–2339.
(7) Balaban, A. T. Topological Indices Based on Topological Distances in Molecular Graphs. *Pure Appl. Chem.* **1983**, *55*, 199–206.
(8) Klein, D. J.; Randić, M. Resistance Distance. *J. Math. Chem.* **1993**, *12*, 81–95.

CI940058V