Relation between Distance-Based Topological Indices

Ivan Gutman* and Željko Tomović

Faculty of Science, University of Kragujevac, P.O. Box 60, YU-34000 Kragujevac, Yugoslavia

The π-index of the line graph of the molecular graph of an alkane is shown to be related to the π, Harary, and reciprocal Wiener indices of that alkane. Thus a mathematical relation between several distance-based topological indices is obtained.

1. INTRODUCTION

One of the most active fields of research in contemporary chemical graph theory[1] is the study of *topological indices*[2] (TIs)—graph invariants that can be used for describing and predicting physicochemical and/or pharmacologic properties of organic compounds. So far over a hundred TIs have been considered in the chemical literature, and the growth of their number shows no sign of attenuation. It is, therefore, not without some value to demonstrate that, sometimes, mathematical relations exist between certain TIs, implying that they are not mutually independent and, therefore, (at least) one of them is redundant.

From a practicioner's point of view,[3] redundant TIs are those between which a linear relation exists, i.e., those that give the regression equations with the same statistical parameters: the coefficient of regressions, the standard error, and the Fisher ratio. From our, more formal, point of view, if between p distinct TIs there exists a *mathematically exact* relation, either linear or nonlinear, enabling that one TI can be computed from the remaining $p - 1$ TIs, then we may consider one of these p TIs as redundant. In all hitherto known cases,[4−12] the mathematical relations between TIs are either linear or can be made linear, thus satisfying both the formal and the practical redundancy requirements. As seen below, the same applies to the relations deduced in this paper.

Relations of this kind were discovered between the Wiener and molecular topological indices (in the case of alkanes),[4,5] between the quasi-Wiener and Kirchhoff indices,[6,7] as well as some other.[8−10] The relation between the Wiener indices of a tree and of its line graph is also long known,[11] and was recently rediscovered in a chemical context.[12]

In this article we describe another connection of the same kind: We show that the π-index of the line graph of a tree is related to the π-index of the tree itself, as well as to an infinite sequence of Wiener-type indices among which the dominant contributions come from the Harary and the reciprocal Wiener indices.

Let G be a connected graph on n vertices, possessing m edges. Denote by $d(u,v|G)$ the distance[13,14] between the vertices u and v of G. Recall that the distance between adjacent vertices is unity. Therefore G has exactly m pairs of vertices with the property $d(u,v|G) = 1$. For the remaining $\binom{n}{2} - m$ vertex pairs, $d(u,v|G) \geq 2$.

Now, if G is a molecular graph, then a class of distance-based TIs associated with it is defined as:[10]

$$W_k(G) = \sum_{u<v} d(u,v|G)^k \qquad (1)$$

where the summation goes over all pairs of distinct vertices of G. For $k = 1$ the quantity W_k reduces to the ordinary Wiener index;[15] for $k = -2$ and $k = -1$ it coincides with the recently introduced Harary index[16,17] and reciprocal Wiener index,[18,19] respectively. For $k = +2$ and $k = +3$ the right-hand side of eq 1 is related to the hyper-Wiener[20] (*WW*) and Tratch-Stankevich-Zefirov[21] (*TSZ*) indices, respectively:[10]

$$WW = \frac{1}{2}W_1 + \frac{1}{2}W_2$$

$$TSZ = \frac{1}{3}W_1 + \frac{1}{2}W_2 + \frac{1}{6}W_3$$

In other words, *WW* reduces to a linear combination of W_k, $k = 1, 2$, whereas *TSZ* is a linear combination of W_k, $k = 1, 2, 3$. More details on this matter can be found elsewhere.[10]

For the purpose of the present work we rewrite formula 1 as

$$W_k(G) = W_k'(G) + m \qquad (2)$$

where

$$W_k'(G) = \sideset{}{'}\sum_{u<v} d(u,v|G)^k \qquad (3)$$

with $\sum'$ indicating summation over only those vertex pairs for which $d(u,v|G) \geq 2$.

By examining the behavior of W_k in the vicinity of $k = 0$, we have recently conceived the π-index,[22] defined as:
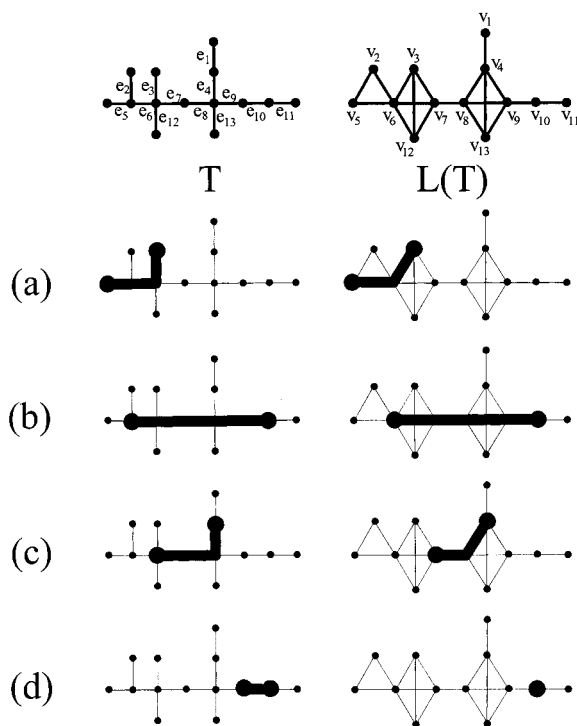
$$\pi(G) = \prod_{u<v} d(u,v|G) \qquad (4)$$

In (4) the multiplication goes over all pairs of distinct vertices of the graph G. Pairs of adjacent vertices play, of course, no role, and therefore we may also write (4) as

$$\pi(G) = \sideset{}{'}\prod_{u<v} d(u,v|G) \qquad (5)$$

where the meaning of $\prod'$ is analogous to that of $\sum'$.

* Corresponding author. E-mail: gutman@knez.uis.kg.ac.yu.

1334 *J. Chem. Inf. Comput. Sci., Vol. 40, No. 6, 2000*

GUTMAN AND TOMOVIC

Figure 1. Molecular graph *T* of 2,3,3,5-tetramethyl-5-ethyloctane and its line graph *L(T)*. Heavy lines indicate paths in *T* and the corresponding paths in *L(T)*.

2. LINE GRAPHS OF MOLECULAR GRAPHS AND THE TOPOLOGICAL INDICES BASED ON THEM

The *line graph* is a standard notion of (mathematical) graph theory. Let the edges of the graph *G* be e_1, e_2, ..., e_m. Then the line graph *L(G)* of the graph *G* has *m* vertices, say v_1, v_2, ..., v_m, such that the vertices v_i and v_j of *L(G)* are adjacent if and only if the edges e_i and e_j of *G* touch each other (i.e., have a common endpoint). An example is shown in Figure 1.

Bertz[23] in 1981 was the first who explicitly proposed the usage of line graphs for designing quantitative structure−property relations (QSPR). However, extensive research along these lines started only a few years ago, mainly due to the efforts of Estrada.[24−33] A survey of the chemical applications of line graphs, including details on the work prior to the 1980s, as well as an exhaustive bibliography is found in the article.[30]

The main strategy in using line graphs in QSPR studies is the following.[32] If $TI^g = f(G)$ is a topological index, obtained from the molecular graph *G* by some operation *f*, then the topological index $TI^l = f(L(G))$ may also have useful properties, different from those of TI^g.

In the case of alkanes, the Wiener index is found to be an exception.[12] Namely, for *G* being an *n*-vertex tree,

$$W_1(L(G)) = W_1(G) - \binom{n}{2} \qquad (6)$$

an identity discovered by Buckley.[11] Because of (6) the Wiener indices of both the graph *G* and the line graph *L(G)* have precisely the same structure dependency. (Clearly, eq 6 implies a linear relation between the Wiener indices of *G* and *L(G)* only within classes of isomeric alkanes, having equal values of *n*.)

In view of the earlier established fact[22,34] that in the case of alkanes there is a very good linear correlation between the logarithm of the π-index, eq 4 or 5, and the Wiener index, eq 1, $k = 1$, one might expect that an identity analogous to (6) will hold also for the π-index. This, however, is not the case. In the subsequent section we show that the π-indices of a tree *T* and of *L(T)* are related as

$$\pi(L(T)) = \pi(T) \sum_{k=1}^{\infty} e^{-\Theta_k} \qquad (7)$$

where

$$\Theta_k = \frac{1}{k} W'_{-k}(T) = \frac{1}{k} [W_{-k}(T) - n + 1] \qquad (8)$$

or, in logarithmic form as:

$$
\begin{aligned}
\ln \pi(L(T)) &= \ln \pi(T) - \sum_{k=1}^{\infty} \Theta_k \\
&= \ln \pi(T) - W'_{-i}(T) - \\
&\quad \frac{1}{2} W'_{-2}(T) - \frac{1}{3} W'_{-3}(T) - ... \quad (9)
\end{aligned}
$$

3. PROOF OF EQ 7

In what follows *T* denotes an *n*-vertex tree, that is, a connected acyclic graph, possessing *n* vertices and $n - 1$ edges. The proof of eq 7 is based on the fact that in both *T* and *L(T)* between any two vertices there exists a unique shortest path. (Recall that the distance between two vertices is just the number of edges in a shortest path connecting these vertices.)

Now, there is a one-to-one correspondence between shortest paths of length greater than 1 in *T* and the shortest paths in *L(T)*. To see this, consider two distinct vertices, say *r* and *s*, of *T*. (In Figure 1 these are indicated by heavy dots.) If e_i and e_j are the first and last edges in the path of *T*, connecting *r* and *s*, and if $e_i \neq e_j$, then this path corresponds to the shortest path in *L(T)*, connecting the vertices v_i and v_j. Furthermore, $d(v_i,v_j|L(T)) = d(r,s|T) - 1$. This is illustrated in Figure 1 in diagrams (a), (b), and (c).

Exceptionally, if the vertices *r* and *s* are adjacent (as in diagram (d) in Figure 1), then $e_i \equiv e_j$ and a single vertex of *L(T)* corresponds to the path between *r* and *s*.

We thus see that

$$\pi(L) = \prod_{i<j} d(v_i,v_j|L(T)) = \prod_{r<s}{}' [d(r,s|T) - 1]$$

i.e.,

$$
\begin{aligned}
\ln \pi(L(T)) &= \sum_{r<s}{}' \ln [d(r,s|T) - 1] \\
&= \sum_{r<s}{}' \ln d(r,s|T) + \sum_{r<s}{}' \ln \left[1 - \frac{1}{d(r,s|T)}\right] \\
&= \ln \pi(T) + \sum_{r<s}{}' \ln \left[1 - \frac{1}{d(r,s|T)}\right]
\end{aligned}
$$

Combining the relations (2) and (3) with the power-series expansion

DISTANCE-BASED TOPOLOGICAL INDICES

J. Chem. Inf. Comput. Sci., Vol. 40, No. 6, 2000 **1335**

Table 1. The Approximations $\ln \pi(L(T))^{(p)}$, Eq 10 for the First Few Values of the Parameter p, and the Exact $\ln \pi(L(T)) \equiv \ln \pi(L(T))^{(\infty)}$ of the Octane Isomers[a]

isomer	$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 6$	$p = 7$	$p = 8$	$\ln \pi$
a	18.812	17.561	17.217	17.102	17.059	17.042	17.035	17.032	17.030
b	17.202	15.837	15.451	15.321	15.272	15.252	15.244	15.241	15.238
c	16.342	14.935	14.539	14.406	14.356	14.336	14.328	14.324	14.322
d	16.069	14.651	14.252	14.118	14.068	14.049	14.040	14.037	14.034
e	14.338	12.750	12.285	12.124	12.062	12.038	12.027	12.023	12.019
f	13.050	11.391	10.906	10.740	10.677	10.651	10.641	10.636	10.633
g	14.266	12.712	12.266	12.115	12.058	12.036	12.026	12.022	12.019
h	14.637	13.108	12.669	12.520	12.464	12.441	12.432	12.428	12.425
i	15.555	14.072	13.645	13.499	13.444	13.422	13.413	13.409	13.405
j	13.622	12.033	11.577	11.423	11.366	11.343	11.333	11.329	11.326
k	15.209	13.749	13.340	13.203	13.152	13.132	13.124	13.121	13.118
l	12.406	10.711	10.217	10.047	9.984	9.958	9.948	9.943	9.940
m	11.463	9.674	9.143	8.959	8.890	8.862	8.850	8.845	8.841
n	12.576	10.860	10.351	10.173	10.105	10.078	10.066	10.061	10.057
o	11.092	9.279	8.741	8.555	8.485	8.456	8.445	8.440	8.436
p	9.205	7.205	6.594	6.379	6.296	6.263	6.249	6.243	6.238
q	13.349	11.748	11.290	11.135	11.078	11.055	11.046	11.041	11.038
r	12.035	10.316	9.814	9.643	9.579	9.553	9.542	9.538	9.534

[a] a = n-octane, b = 2-methylheptane, c = 3-methylheptane, d = 4-methylheptane, e = 2,2-dimethylhexane, f = 3,3-dimethylhexane, g = 2,3-dimethylhexane, h = 2,4-dimethylhexane, i = 2,5-dimethylhexane, j = 3,4-dimethylhexane, k = 3-ethylhexane, l = 2,3,4-trimethylpentane, m = 2,2,3-trimethylpentane, n = 2,2,4-trimethylpentane, o = 2,3,3-trimethylpentane, p = 2,2,3,3-tetramethylbutane, q = 2-methyl-3-ethylpentane, r = 3-methyl-3-ethylpentane.

$$\ln(1 - x) = - \sum_{k=1}^{\infty} \frac{1}{k} x^k$$

which is convergent for $|x| < 1$, we calculate

$$\sum_{r<s}' \ln\left[1 - \frac{1}{d(r,s|T)}\right] = - \sum_{k=1}^{\infty} \frac{1}{k} W'_{-k}(T)$$

from which eq 9 follows directly.

4. DISCUSSION

With increasing k the quantities Θ_k, eq 8, rapidly tend to zero, faster than $n^2/(k2^{k+1})$. For instance, for 2-methylbutane, $\Theta_1 = 2.6667$, $\Theta_2 = 0.6111$, $\Theta_3 = 0.1914$, $\Theta_4 = 0.0687$, $\Theta_5 = 0.0266$, $\Theta_6 = 0.0109$, $\Theta_7 = 0.0046$, $\Theta_8 = 0.0020$, $\Theta_9 = 0.0009$, $\Theta_{10} = 0.0004$, $\Theta_{11} = 0.0002$, and $\Theta_{12} = 0.0001$. The analogous values for 2,5-dimethyloctane are 11.4024, 2.0953, 0.5745, 0.1910, 0.0709, 0.0282, 0.0118, 0.0051, 0.0022, 0.0010, 0.0004, and 0.0002, respectively.

Therefore the product in (7) and the sum in (9) may be truncated without significantly affecting the values of $\pi(T)$ and $\ln \pi(T)$, respectively:
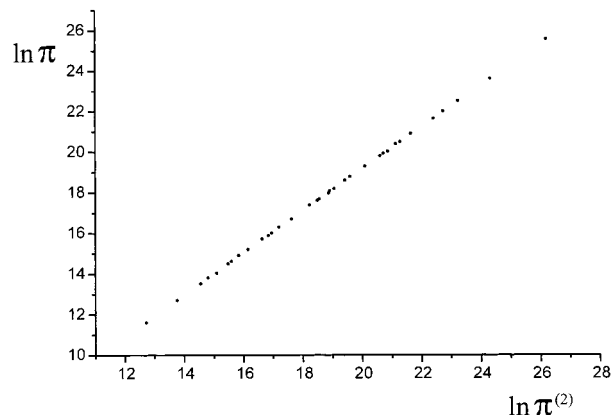
$$\pi(L(T)) \approx \pi(L(T))^p \equiv \pi(T) \prod_{k=1}^{p} e^{-\Theta_k}$$

$$\ln \pi(L(T)) \approx \ln \pi(L(T))^{(p)}$$
$$= \ln \pi(T) - W'_{-1}(T) - \frac{1}{2} W'_{-2}(T) -$$
$$\cdots - \frac{1}{p} W'_{-p}(T) \quad (10)$$

Clearly, for any value of p, the truncated quantity $\pi(T)^{(p)}$ is an upper bound for the π-index of the line graph of T. In addition, $\pi(L(T)) < \pi(T)$.

Numerical testing shows that already $p = 2$, and certainly $p = 3$, is sufficient to reach a chemically acceptable approximation. Examples are found in Table 1.

Figure 2. Logarithm of $\pi(L(T))$ versus $\ln \pi(L(T))^{(2)}$, eq 10, for the isomers of nonane. The slope of the regression line is 1.034 ± 0.002; correlation coefficient = 0.99997; note that $\pi(L(T))^{(2)}$ is a function of only the π, reciprocal Wiener, and Harary indices of the respective molecular graph.

Within classes of isomeric alkanes an excellent linear correlation between $\ln \pi$ and $\ln \pi^{(p)}$ exists already for $p = 1$. An illustration is given in Figure 2, revealing that the four TIs of alkanes, the π-indices of the graph and its line graph, the reciprocal Wiener indices, and the Harary indices, are mutually dependent. In particular, the π-index of the line graph can be a good approximation, expressed as a function of the other three TIs.

REFERENCES AND NOTES

(1) Trinajstić, N. *Chemical Graph Theory*; CRC Press: Boca Raton, FL, 1993.
(2) *Topological Indices and Related Descriptors in QSAR and QSPR*; Devillers, J.; Balaban, A. T.; Eds.; Gordon & Breach: New York, 1999.
(3) The statistical requirements for the redundancy of two topological indices were formulated by an anonymous referee.
(4) Klein, D. J.; Mihalić, Z.; Plavšić, D.; Trinajstić, N. Molecular Topological Index: A Relation with the Wiener Index. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 304−305.
(5) Gutman, I. Selected Properties of the Schultz Molecular Topological Index. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1087−1089.

(6) Gutman, I.; Mohar, B. The Quasi-Wiener and the Kirchhoff Indices Coincide. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 982−985.

(7) Zhu, H. Y.; Klein, D. J.; Lukovits, I. Extensions of the Wiener Number. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 420−428.

(8) Chan, O.; Gutman, I.; Lam, T. K.; Merris, R. Algebraic Connections between Topological Indices. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 62−65.

(9) Gutman, I.; Vidović, D.; Popović, L. On Graph Representation of Organic Molecules−Cayley's Plerograms vs His Kenograms. *J. Chem. Soc., Faraday Trans.* **1998**, *94*, 857−860.

(10) Klein, D. J.; Gutman, I. Wiener-Number-Related Sequences. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 534−536.

(11) Buckley, F. Mean Distance in Line Graphs. *Congr. Numer.* **1981**, *32*, 153−162.

(12) Gutman, I.; Estrada, E. Topological Indices Based on the Line Graph of the Molecular Graph. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 541−543.

(13) Buckley, F.; Harary, F. *Distance in Graphs*; Addison-Wesley: Redwood, 1990.

(14) Gutman, I.; Polansky, O. E. *Mathematical Concepts in Organic Chemistry*; Springer-Verlag: Berlin, 1986.

(15) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc.* **1947**, *69*, 17−20.

(16) Plavšić, D.; Nikolić, S.; Trinajstić, N.; Mihalić, Z. On the Harary Index for the Characterization of Chemical Graphs. *J. Math. Chem.* **1993**, *12*, 235−250.

(17) Mihalić, Z.; Trinajstić, N. A Graph-Theoretical Approach to Structure-Property Relationships. *J. Chem. Educ.* **1992**, *69*, 701−712.

(18) Diudea, M. V.; Ivanciuc, O.; Nikolić, S.; Trinajstić, N. Matrices of Reciprocal Distance, Polynomials and Derived Numbers. *Commun. Math. Chem. (MATCH)* **1997**, *35*, 41−64.

(19) Diudea, M. V. Indices of Reciprocal Properties or Harary Indices. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 292−299.

(20) Randić, M. Novel Molecular Descriptor for Structure−Property Studies. *Chem. Phys. Lett.* **1993**, *211*, 478−483.

(21) Tratch, S. S.; Stankevich, M. I.; Zefirov, J. Combinatorial Models and Algorithms in Chemistry. The Expanded Wiener Number−A Novel Topological Index. *J. Comput. Chem.* **1990**, *11*, 899−908.

(22) Gutman, I.; Linert, W.; Lukovits, I.; Tomović, Ž. The Multiplicative Version of the Wiener Index. *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 113−116.

(23) Bertz, S. H. The Bond Graph. *J. Chem. Soc., Chem. Commun.* **1981**, 818−820.

(24) Estrada, E. Edge Adjacency Relationships and a Novel Topological Index Related to Molecular Volume. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 31−33.

(25) Estrada, E. Edge Adjacency Relationships in Molecular Graphs Containing Heteroatoms: A New Topological Index Related to Molecular Volume. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 701−707.

(26) Estrada, E.; Ramírez, A. Edge Adjacency Relationships and Molecular Topographic Descriptors. Definition and QSAR Applications. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 837−843.

(27) Estrada, E. Spectral Moments of the Edge Adjacency Matrix in Molecular Graphs. 1. Definition and Applications to the Prediction of Physical Properties of Alkanes. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 844−849.

(28) Estrada, E. Spectral Moments of the Edge-Adjacency Matrix of Molecular Graphs. 2. Molecules Containing Heteroatoms and QSAR Applications. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 320−328.

(29) Estrada, E. Spectral Moments of the Edge Adjacency Matrix in Molecular Graphs. 3. Molecules Containing Cycles. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 23−27.

(30) Gutman, I.; Popović, L.; Estrada, E.; Bertz, S. H. The Line Graph Model. Predicting Physico-Chemical Properties of Alkanes. *ACH Models Chem.* **1998**, *135*, 147−155.

(31) Estrada, E. Generalized Spectral Moments of the Iterated Line Graphs Sequence. A Novel Approach to QSPR Studies. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 90−95.

(32) Estrada, E.; Rodríguez, L. Edge-Connectivity Indices in QSPR/QSAR Studies. 1. Comparison to other Topological Indices in QSPR Studies. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 1037−1041.

(33) Estrada, E. Edge-Connectivity Indices in QSPR/QSAR Studies. 2. Accounting for Long-Range Bond Contributions. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 1042−1048.

(34) Gutman, I.; Linert, W.; Lukovits, I.; Tomovic, Ž. On the Multiplicative Wiener Index and Its Possible Chemical Applications. *Monatsh. Chem.* **2000**, *131*, 421−427.

CI000014O