# Topological Indices Based on the Line Graph of the Molecular Graph

Ivan Gutman*,†,§ and Ernesto Estrada‡

Institute of Physical Chemistry, Attila Jozsef University, P.O. Box 105, H-6701 Szeged, Hungary, and Departamento de Diseño de Fármacos, Centro de Bioactivos Químicos Universidad Central de Las Villas, Santa Clara 54830, Cuba

Received September 30, 1995⊗

The topological index $\epsilon$ recently proposed by one of the authors [E. Estrada, *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 31−33] is shown to be identical to the connectivity index of the line graph of the molecular graph. This observation makes it possible to conceive a whole class of novel, line-graph-based topological indices. The Wiener index provides an exception: in the case of acyclic molecules, the Wiener index of the molecular graph and the Weiner index of the line graph differ by a constant.

## 1. INTRODUCTION

In a recent article[1] one of the present authors put forward a novel topological index $\epsilon = \epsilon(G)$, based on the so-called edge-adjacency matrix $\mathbf{E} = \mathbf{E}(G)$, where $G$ stands for the molecular graph. This matrix was defined[1] in analogy to the usual vertex-adjacency matrix $\mathbf{A} = \mathbf{A}(G)$. Then the index $\epsilon$ is calculated from $\mathbf{E}(G)$ in precisely the same way as the classical Randić's, connectivity index[2] $\chi = \chi(G)$ is computed from $\mathbf{A}(G)$.

The aim of this paper is to point out that $\mathbf{E}(G)$ can be viewed as the vertex-adjacency matrix of the line graph $L(G)$ of the molecular graph $G$. If so, then $\epsilon$ is just Randić's connectivity index of the line graph, $\epsilon(G) = \chi(\mathrm{L}(G))$. Furthermore, $\epsilon$ occurs to be just the first representative of an entire series of possible novel topological indices, based on the line graph.

## 2. EDGE-ADJACENCY MATRIX AND THE CONCEPT OF LINE GRAPH

Employing the same notation as in our previous work[1] we denote the molecular graph by $G = \{V,E\}$, where $V$ is the vertex set and $E$ is the edge set. This graph is assumed to be connected, possessing $n$ vertices and $m$ edges. If a vertex $v$ is incident to an edge $e$ then we write $v \sim e$ or $e \sim v$. The vertex-adjacency matrix (or, as usually called,[3−6] the adjacency matrix) of $G$ is a symmetric square matrix $\mathbf{A} = \mathbf{A}(G) = [a_{ij}]$ of order $n$, the elements of which are defined as

$$a_{ij} = \begin{cases} 1 & \text{if } e \in E;\ e \sim v_i,\ v_j \\ 0 & \text{otherwise} \end{cases} \qquad (1)$$

The edge-adjacency matrix of $G$ is a symmetric square matrix $\mathbf{E} = \mathbf{E}(G) = [g_{ij}]$ of order $m$, the elements of which are given by

$$g_{ij} = \begin{cases} 1 & \text{if } v \in V;\ v \sim e_i,\ e_j \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

The edge-adjacency matrix is often encountered in graph theory[3] and its chemical applications.[4,5] There is an evident analogy between (1) and (2).

In graph theory[3] one defines the so-called line graph $L(G)$ of a graph in the following manner. Let $e_1, e_2, ..., e_m$ be the edges of the graph $G$. Then $e_1, e_2, ..., e_m$ are the vertices of $L(G)$; two vertices of $L(G)$ are adjacent if the corresponding edges in $G$ are incident (i.e., have a vertex in common). The construction of line graphs is illustrated in Figure 1.

It is now straightforward to see that (2) defines precisely the vertex-adjacency matrix of the line graph of $G$. In other words, the following simple relation holds

$$\mathbf{E}(G) = \mathbf{A}(L(G)) \qquad (3)$$

An immediate consequence of (3) is that the topological index $\epsilon = \epsilon(G)$ introduced in our previous paper[1] is equal to the connectivity index[2] of the line graph: $\epsilon(G) = \chi(L(G))$. Thus the results communicated in ref 1 can be viewed as one more novel application of Randić's index.[7,8]

The theory of line graphs is well elaborated.[3] Of the numerous results known on $L(G)$ we mention just a few, which may be directly applicable in chemical considerations.

(i) The number of vertices of $L(G)$ is equal to $m$. If $d_i$ is the degree of the vertex $v_i$ of $G$, then the number of edges of $L(G)$ is equal to

$$\sum_{i=1}^{n} \binom{d_i}{2} = \sum_{i=1}^{n} (d_i)^2 - m = M_2 - m$$

where by $M_2 = \sum(d_i)^2$ is denoted the so-called Zagreb-group index.[4,9−12]

(ii) If $e_k$ is an edge of $G$ connecting the vertices $v_i$ and $v_j$, then the degree of the $k$th vertex of $L(G)$ is equal to $d_i + d_j - 2$.

(iii) Property (ii) implies that if $G$ is a regular graph of degree $d$, then $L(G)$ is also regular, of degree $2d - 2$. However, $L(G)$ may be regular also if $G$ is not; for example, the line graph of the $n$-vertex star is the $(n-1)$-vertex complete graph.

(iv) Suppose that $G$ has no isolated vertices. Then $G$ and $L(G)$ have an equal number of components. In particular, $L(G)$ is connected if and only if $G$ is connected.

(v) If $G$ is a connected graph, then $G$ uniquely determines $L(G)$, with the single exception of the graphs $G_3$ and $G_4$ depicted in Figure 1.

† Attila Jozsef University.
‡ Centro de Bioactivos Químicos Universidad Central de Las Villas.
§ On leave from the Faculty of Science, University of Kragujevac, 34000 Kragujevac, Yugoslavia.
⊗ Abstract published in *Advance ACS Abstracts,* January 1, 1996.

0095-2338/96/1636-0541$12.00/0 © 1996 American Chemical Society

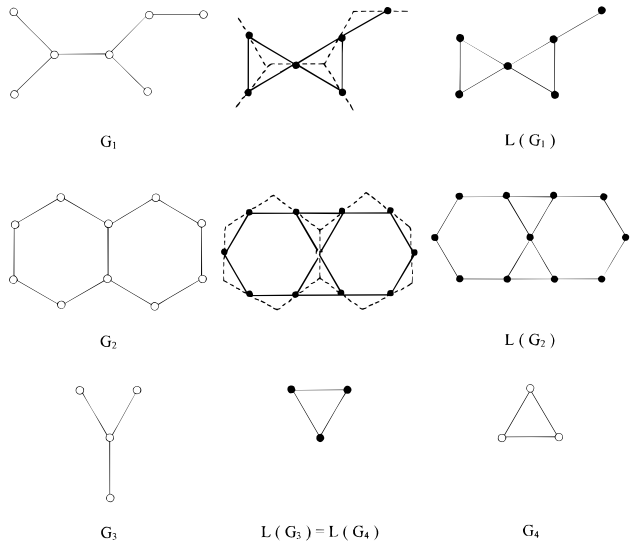



**Figure 1.** The molecular graphs of 2,3-dimethylpentane ($G_1$) and napthalene ($G_2$) and the corresponding line graphs; $G_3$ and $G_4$ are the only two connected graphs of which the line graphs coincide.

(vi) If $G$ is a connected graph, then $G$ and $L(G)$ coincide only in the case of circuits $C_n$, $n = 3, 4,...$

(vii) Line graphs of connected graphs are cyclic, with the only exception of paths $P_n$, $n = 2, 3,...$ For these latter graphs, $L(P_n) = P_{n-1}$.

## 3. ON POSSIBLE CHEMICAL APPLICATIONS OF LINE GRAPHS

The fact that Randić's connectivity index of the line graph of the molecular graph was found[1] to be highly correlated with certain physicochemical properties of alkanes indicates that analogous considerations could be purposeful also in the case of other topological indices.[11,12]

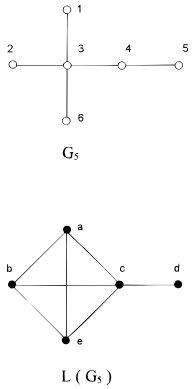
Thus, if $I(G)$ is a topological index that can be computed from the molecular graph $G$, we could image another topological index $i(G)$ which is computed in the same way as $I$ but from the line graph of the molecular graph. In other words, $i(G) = I(L(G))$. Because the number of currently used topological indices $I(G)$ is quite large,[11,12] the number of candidates for novel topological indices of the type $i(G)$ is large too. Their testing is presently under way and will be reported elsewhere.[13] Our preliminary results suggest that in some cases $i(G)$ is more suitable for QSPR and QSAR studies than $I(G)$.

For one such index $i(G)$ no testing is needed whatsoever. Namely, if $W = W(G)$ is the Wiener number,[14−16] then the following simple relation holds

$$W(G) = W(L(G)) + \binom{n}{2} \tag{4}$$

showing that in this case $I(G)$ and $i(G)$ are not only linearly proportional but also their difference is a constant [$= n(n-1)/2$]. The identity (4) is obeyed whenever $G$ is a molecular graph of an alkane; as before, $n$ is the number of vertices of $G$, i.e., the number of carbon atoms in the respective molecule.

It should be mentioned here that line-graph-related topological indices were considered also in the past[17−20] but different types than the ones examined in the present paper.




| no. | corresponding paths of $G_5$ | and $L(G_5)$ |
|---|---|---|
| 1 | 1,3,2 | a,b |
| 2 | 1,3 | a |
| 3 | 1,3,4 | a,c |
| 4 | 1,3,4,5 | a,c,d |
| 5 | 1,3,6 | a,e |
| 6 | 2,3 | b |
| 7 | 2,3,4 | b,c |
| 8 | 2,3,4,5 | b,c,d |
| 9 | 2,3,6 | b,e |
| 10 | 3,4 | c |
| 11 | 3,4,5 | c,d |
| 12 | 3,6 | e |
| 13 | 4,5 | d |
| 14 | 4,3,6 | c,e |
| 15 | 5,4,3,6 | d,c,e |

**Figure 2.** The molecular graph of 2,2-dimethylbutane ($G_5$) and its line graph; in the paths of $G_5$ and $L(G_5)$ only the vertices are indicated; $W(G_5) = 28$, $W(L(G_5)) = 13$, $W(G_5) - W(L(G_5)) = \binom{6}{2} = 15$.

In order to verify the validity of (4) recall that a path in the graph $G$ is a sequence of vertices and edges[3−6]

$$v_0, e_1, v_1, e_2, ..., v_{k-1}, e_k, v_k \tag{5}$$

such that the edge $e_i$ is incident to the vertices $v_{i-1}$ and $v_i$ (and thus $v_{i-1}$ and $v_i$ are adjacent), $i = 1, 2, ..., k$. Further, no two among the vertices $v_0, v_1, v_2, ..., v_k$ coincide; therefore no two among the edges $e_1, e_2, ..., e_k$ coincide either. Such a path is said to connect the vertices $v_0$ and $v_k$ and to have length $k$. It may happen that the path (5) is composed of a single vertex and of no edge ($k = 0$); then its length is zero.

The length of the shortest path (or more precisely, the length of any among the several possible shortest paths) in a graph, connecting two vertices, is said to be the distance of these vertices. The Wiener number is equal to the sum of distances between all pairs of vertices.[14−16]

In connected acyclic graphs (trees) the path between any two vertices is unique. Hence, if $v_0$ and $v_k$ are any two distinct vertices of a tree, then the sequence (5) is unique. Then, however, the following sequence is unique too

$$e_1, v_1, e_2, ..., v_{k-1}, e_k \tag{6}$$

Bearing in mind the definition of the line graph we realize that (6) is just a path of $L(G)$, involving the vertices $e_1, e_2, ..., e_k$ of $L(G)$, connecting the vertices $e_1$ and $e_k$ and having length $k-1$. The sequence (5) is the unique path between the vertices $v_0$ and $v_k$ of $G$ and is therefore the shortest such path. The sequence (6) is not the unique path between the vertices $e_1$ and $e_k$ of $L(G)$ but is the shortest such path.

We thus reach the following conclusion: If $G$ is a tree, then there is a one-to-one correspondence between the paths of $G$ (having length one or greater) and the shortest paths of $L(G)$ (having length zero or greater). Futhermore, a path of $G$ of length $k$ corresponds to a shortest path of $L(G)$ of length $k-1$.

Consequently, if $G$ is a tree, then any pair of vertices of $G$ being at distance $k$, $k > 0$, is in a one-to-one correspondence with a pair of vertices of $L(G)$ being at distance $k-1$. Consequently, the sum of all distances in $G$ exceeds the sum of all distances in $L(G)$ by the number of vertex pairs in $G$, namely by $\binom{n}{2} = n(n-1)/2$.

By this we proved the identity (4).

In Figure 2 formula (4) as well as the underlying relations between the paths are illustrated in the example of the molecular graph of 2,2-dimethylbutane and its line graph.

## ACKNOWLEDGMENT

One author (I.G.) gratefully acknowledges the support by the Mathematical Institute, Belgrade and the Lady Davis Fellowship Trust, Jerusalem, Israel.

## REFERENCES AND NOTES

(1) Estrada, E. Edge Adjacency Relationships and a Novel Topological Index Related to Molecular Volume. *J. Chem. Inf. Comput. Sci*. **1995**, *35*, 31−33.
(2) Randić, M. On Characterization of Molecular Branching. *J. Am. Chem. Soc*. **1975**, *97*, 6609−6615.
(3) Harary, F. *Graph Theory*; Addison-Wesley: Reading, 1969.
(4) *Chemical Applications of Graph Theory*; Balaban, A. T., Ed.; Academic Press: London, 1976.
(5) Trinajstić, N. *Chemical Graph Theory*; CRC Press: Boca Raton, 1983; 2nd ed.; 1992.
(6) Gutman, I.; Polansky, O. E. *Mathematical Concepts in Organic Chemistry*; Springer-Verlag: Berlin, 1986.
(7) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Chemistry and Drug Research*; Academic Press: New York, 1976.
(8) Kier, L. B.; Hall, L. H. *Molecular Connectivity in Structure−Activity Analysis*; Wiley: New York, 1986.
(9) Gutman, I.; Ruščić, B.; Trinajstić, N.; Wilcox, C. F. Graph Theory and Molecular Orbitals. XII. Acyclic Polyenes. *J. Chem. Phys*. **1975**, *62*, 3399−3405.
(10) Balaban, A. T. Highly Discriminating Distance-Based Topological Index. *Chem. Phys. Lett*. **1982**, *89*, 399−404.
(11) Hansen, P. J.; Jurs, P. C. Chemical Applications of Graph Theory. Part I. Fundamentals and Topological Indices. *J. Chem. Educ*. **1988**, *65*, 574−580.
(12) Mihalić, Z.; Trinajstić, N. A Graph-Theoretical Approach to Structure−Property Relationships. *J. Chem. Educ*. **1992**, *69*, 701−712.
(13) Estrada, E.; Popović, L.; Gutman, I. Work in progress.
(14) Wiener, H. Structural Determination of Paraffin Boiling Points. *J. Am. Chem. Soc*. **1947**, *69*, 17−20.
(15) Gutman, I.; Yeh, Y. N.; Lee, S. L.; Luo, Y. L. Some Recent Results in the Theory of the Wiener Number. *Indian J. Chem*. **1993**, *32A*, 651−661.
(16) Nikolić, S.; Trinajstić, N.; Mihalić, Z. The Wiener Index: Development and Applications. *Croat. Chem. Acta* **1995**, *68*, 105−130.
(17) Bertz, S. H. On the complexity of graphs and molecules. *Bull. Math. Biol*. **1983**, *45*, 849−855.
(18) Bertz, S. H. A mathematical model of molecular complexity. *Stud. Phys. Theor. Chem*. **1983**, *28*, 206−221.
(19) Bertz, S. H.; Herndon, W. C. The similarity of graphs and molecules. *Am. Chem. Soc. Symp*. **1986**, *306*, 169−175.
(20) Herndon, W. C.; Bertz, S. H. Linear notation and molecular graph similarity. *J. Comput. Chem*. **1987**, *8*, 367−374.

CI950143I