Chemical applications of the Laplacian spectrum. VII. Studies of the Wiener and Kirchhoff indices

Ivan Gutman*, Dušica Vidović & Boris Furtula

Faculty of Science, University of Kragujevac, P.O.Box 60, YU-34000 Kragujevac, Serbia & Montenegro

Some further chemical applications of the Laplacian spectra are reported. The Kel'mans theorem for the calculation of the coefficients of the Laplacian characteristic polynomial is stated and exemplified. By means of this theorem a (previously known) formula for the Wiener and Kirchhoff index is deduced. It is shown that the Wiener index is correlated with the "algebraic connectivity", namely, the smallest positive Laplacian eigenvalue. Lower and upper bounds for the Kirchhoff index are obtained.

Introduction

Contrary to the ordinary graph spectrum, which found numerous applications in chemistry (see, for instance, the books[1,2]), the Laplacian graph spectrum was for a long time on the periphery of interest of mathematical chemists. Yet, in a number of recent works[3-8] (which are parts I-VI of the series "Chemical applications of the Laplacian spectrum") we communicated results showing that by means of Laplacian eigenvalues of molecular graphs also, it is possible to acquire chemically useful conclusions. Other authors also did research along the same lines[9-13].

In this paper we continue our studies[3,4] of the connections between the Laplacian eigenvalues and the Wiener and Kirchhoff indices. In Section 2 we state and exemplify the application of the Kel'mans theorem from year 1967 (a result not widely known among mathematical chemists) and show how from it the famous formula (1) for the Wiener index W is deduced:

$$W = n \sum_{i=1}^{n-1} \frac{1}{\mu_i} \qquad \qquad \dots (1)$$

For cycle-containing molecular graphs the right-hand side of (1) differs from the Wiener index and is equal to another structure-descriptor, initially called "quasi-Wiener index"[14], but later identified[15] with the Kirchhoff index (Kf); for details see Section 4.

Formula (1) suggests that the gross part of the Wiener and Kirchhoff indices is determined by the smallest positive Laplacian eigenvalue (the so-called "algebraic connectivity"). The quality of this assumption is tested in Section 3.

In Section 4 formula (1) is used to deduce lower and upper bounds for W and Kf.

The notation used throughout this paper is the following: G denotes a (molecular) graph, possessing n vertices and m edges. Its adjacency matrix is $A(G)$. (Recall[16] that the eigenvalues of A form the ordinary spectrum of G.) The diagonal matrix, whose i-th diagonal element is equal to the degree (= number of first neighbours) of the vertex i, is denoted by $D(G)$. Then the Laplacian matrix of G is $L(G) = D(G)-A(G)$. The eigenvalues of $L(G)$ form the Laplacian spectrum of the graph G.

The Laplacian eigenvalues are denoted by μ_i, $i=1,2,...,n$, and are labelled so that

$$\mu_1 \geq \mu_2 \geq ... \geq \mu_{n-1} \geq \mu_n$$

A general result of the theory of Laplacian spectra is that $\mu_n = 0$ for all graphs, and that μ_{n-1} is equal to zero if and only if G is not connected. For connected graphs (such as all molecular graphs), μ_{n-1} is positive-valued and is called "the algebraic connectivity" (of the graph G).

The Laplacian characteristic polynomial $\Psi(G,x)$ is defined as

$$\Psi(G,x) = det [x I - L(G)]$$

Recall that the Laplacian eigenvalues of the graph G are just the zeros of the polynomial $\Psi(G,x)$, i.e., the roots of the equation $\Psi(G,x)=0$. More details on the Laplacian spectrum can be found in the reviews[17,18].

The Kel'mans theorem

In 1967, the Russian mathematician A. K. *Kel'mans* published an article *"On the properties of the characteristic polynomial of a graph"*. This article appeared on pp. 27-41 of a booklet entitled "Cybernetics in the Service of Communism", and was in Russian language. This may be the reason why many of the contemporary mathematical chemists are not familiar with Kel'mans' result.

The Laplacian characteristic polynomial of the graph G is a polynomial of order n that can be written as

$$\Psi(G,x) = C_0 x^n + C_1 x^{n-1} + \dots + C_{n-1} x + C_n \qquad \dots (2)$$

The theorem discovered by Kel'mans establishes the connection between the structure of the graph G and the coefficients C_0, C_1, ..., C_{n-1}, C_n of its Laplacian characteristic polynomial $\Psi(G,x)$. In order to state it we need to introduce a few simple graph-theoretic notions.

A spanning subgraph of a graph G is a subgraph of G containing all vertices of G. Hence, any spanning subgraph of G is obtained by deleting some edges from G, but keeping all vertices of G.

An acyclic spanning subgraph of a graph G is called a spanning forest of G. A connected acyclic spanning subgraph is called a spanning tree of the respective graph.

Examples of spanning forests and spanning trees are found in Figs 1 and 2.

Let F be a spanning forest of the (molecular) graph G. In the general case F is disconnected and consists of several components F_1, F_2, ..., F_p. Each component F_i, $i=1,2,...,p$, is a graph for itself and is connected.

The number of vertices of the component F_i is denoted by $n(F_i)$, $i=1,2,...,p$. The product of the numbers $n(F_i)$, $i=1,2,...,p$, is denoted by $\gamma(F)$.

The set of all spanning forests of the graph G, containing exactly p components, is denoted by $F(G,k)$. For illustrative examples see Figs 1 and 2.

With this notation, the Kel'mans theorem reads as follows: For $k=0,1,...,n$, the k-th coefficient of the Laplacian characteristic polynomial of the graph G, Eq. (2), is related to the structure of G via

$$C_k = (-1)^k \sum_{F \in F(G,n-k)} \gamma(F) \qquad \dots (3)$$

Hence, in order to obtain the k-th coefficient of $\Psi(G,x)$ we simply have to sum the γ-values of all spanning forests of G, possessing $n-k$ components.
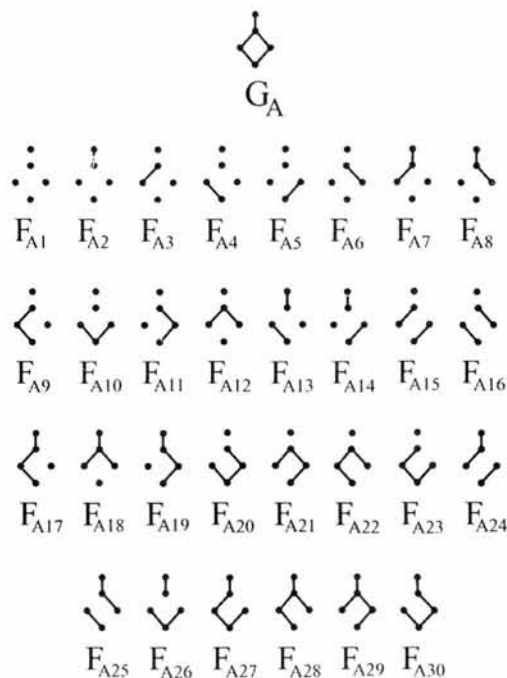

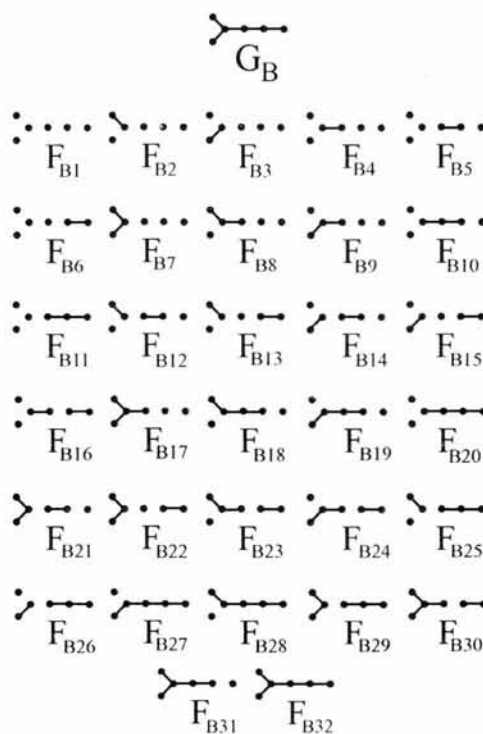Fig. 1—The molecular graph of methylcyclobutane (G_A) and its spanning forests.

Fig. 2—The molecular graph of 2-methylpentane (G_B) and its spanning forests.

We first illustrate the application of the Kel'mans theorem on the example of the molecular graph G_A of methylcyclobutane. This graph and its 30 spanning forests are depicted in Fig. 1. We have

$F(G_A, 5) = \{ F_{A1} \}$

$F(G_A, 4) = \{ F_{A2}, F_{A3}, ..., F_{A6} \}$

$F(G_A, 3) = \{ F_{A7}, F_{A8}, ..., F_{A16} \}$

$F(G_A, 2) = \{ G_{A17}, F_{A18}, ..., F_{A26} \}$

$F(G_A, 1) = \{ G_{A27}, G_{A28}, G_{A29}, G_{A30} \}$

$F(G_A, 0) = \emptyset \ (empty \ set)$

For any n-vertex graph the set $F(G,n)$ has a unique element, for which $\gamma=1$. Therefore, for any graph, $C_0=1$.

All elements of $F(G,n-1)$ possess one edge and $n-2$ isolated vertices. Therefore, each such element has $\gamma=2$, and by Eq. (3)

$$C_1 = (-1)^1 [2m] = -2m$$

where m is the number of edges. In our case $m=5$ and therefore $C_1 = -10$.

There are two kinds of 3-component spanning forests of G_A : for six spanning forests (F_{A7}, F_{A8}, F_{A9}, F_{A10}, F_{A11} and F_{A12}) it is $\gamma=3\cdot1\cdot1=3$, whereas for four spanning forests (F_{A13}, F_{A14}, F_{A15} and F_{A16}) the value of γ is $2\cdot2\cdot1=4$. Application of Eq. (3) gives then

$$C_2 = (-1)^2 [6 \cdot 3 + 4 \cdot 4] = 34$$

Similarly, there are two kinds of 2-component spanning forests of G_A : for seven of them (F_{A17}, F_{A18} ,..., F_{A23}) it is $\gamma=4\cdot1=4$, whereas for three spanning forests (F_{A24}, F_{A25} and F_{A26}) the value of γ is $2\cdot3=6$. Application of Eq. (3) gives then

$$C_3 = (-1)^3 [7 \cdot 4 + 3 \cdot 6] = -46$$

The four 1-component (i. e., connected) subgraphs of G_A are its spanning trees. These are F_{A27}, F_{A28}, F_{A29}, F_{A30}. Thus,

$$C_4 = (-1)^4 [4 \cdot 5] = 20$$

In the general case, the coefficient C_{n-1} is equal to (plus or minus) the number of spanning trees times the number of vertices.

There cannot be spanning subgraph with 0 components, implying that the set $F(G,0)$ is always empty. Consequently, the coefficient C_n is always equal to zero.

In summary, the Laplacian characteristic polynomial of the methylcyclobutane graph is

$$\Psi(G_A, x) = x^5 - 10x^4 + 34x^3 - 46x^2 + 20x$$

From this example we established two general results:

$$C_1 = -2m \qquad \qquad ...(4)$$

and

$$C_{n-1} = (-1)^{n-1} nt \qquad \qquad ...(5)$$

with n, m and t denoting the number of vertices, edges and spanning trees, respectively, of the underlying molecular graph. Relations (4) and (5) are tantamount to

$$\sum_{i=1}^{n-1} \mu_i = 2m \qquad \qquad ...(6)$$

and

$$\prod_{i=1}^{n-1} \mu_i = nt \qquad \qquad ...(7)$$

Formulae (4)-(7) hold for all graphs.

In order to arrive at another application of the Kel'mans theorem we consider the special case of trees, exemplified by the molecular graph G_B of 2-methylpentane, see Fig. 2. Here we shall be interested only in the 2- and 1-component spanning forests. The respective sets are:

$$F(G_B, 2) = \{ F_{B27}, F_{B28}, F_{B29}, F_{B30}, F_{B31} \}$$

and

$$F(G_B, 1) = \{ F_{B32} \}$$

One should note that each member of the set $F(G_B, 2)$ is obtained by deleting an edge from G_B. This is true for any tree: each 2-component spanning forest F of a tree T is obtained by deleting an edge from T. Therefore the numbers $n(F_1)$ and $n(F_2)$ of vertices of the two components of such a spanning forest count the vertices lying on the two sides of the respective edge of T. According to a long known result in the theory of Wiener index[19,20], the summing the products of these two numbers over all edges of T results in the Wiener index W of T. On the other hand, the product of $n(F_1)$ and $n(F_2)$ is just the γ-value of F. Therefore, from the Kel'mans theorem (3) it immediately follows,

$$C_{n-2} = (-1)^{n-2} W \qquad \ldots (8)$$

and this results holds for all trees. In terms of Laplacian eigenvalues, formula (8) is rewritten as

$$W = \left(\frac{1}{\mu_1} + \frac{1}{\mu_2} + \ldots + \frac{1}{\mu_{n-1}} \right) \prod_{i=1}^{n-1} \mu_i \qquad \ldots (9)$$

For trees, as a special case of (7),

$$\prod_{i=1}^{n-1} \mu_i = n \qquad \ldots (10)$$

because for trees, evidently, $t=1$.

Combining (9) and (10) one immediately arrives at the formula (1).

Formula (1) was simultaneously discovered by several mathematicians nearly in the same time (in the late 1980s); bibliographic details can be found in the article[3]. Chemists became aware of it after the publication of the papers in refs 3 and 10.

Relation between Wiener index and the algebraic connectivity

Formula (1) may be viewed as a decomposition of the Wiener index in terms of Laplacian eigenvalues. Each Laplacian eigenvalue has a certain contribution, equal to n/μ_i, $i=1,2,\ldots,n-1$, and the Wiener index is equal to the sum of these contributions.

As already mentioned, the second smallest Laplacian eigenvalue (μ_{n-1}) is called "algebraic connectivity". Evidently, it will have the greatest contribution to the Wiener index. Within the theory of Laplacian spectra the dependence of the algebraic connectivity on graph structure was much investigated[17,18,21-23].

In view of this we were interested to see if there exists some correlation between the Wiener index W and the term n/μ_{n-1}. Indeed, such a correlation does exist.

In Fig. 3 we show the correlation between the Wiener index and the term n/μ_{n-1} for the chemical trees with 10 vertices. In Fig. 4 we show the analogous correlation for all (chemical and non-chemical) trees with 10 vertices. Analogous correlations were found to exist in the case of n-vertex trees and chemical trees also for other values on n. The respective correlation coefficients are given in Table 1.
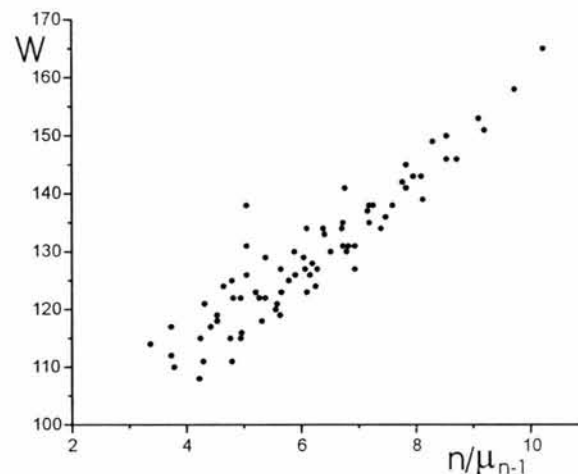

Fig. 3—Correlation between the Wiener index (W) and the term n/μ_{n-1} for chemical trees with $n=10$ vertices; μ_{n-1} is the algebraic connectivity. The outlier corresponds to the molecular graph of 4-propylheptane (cf. Fig. 5).
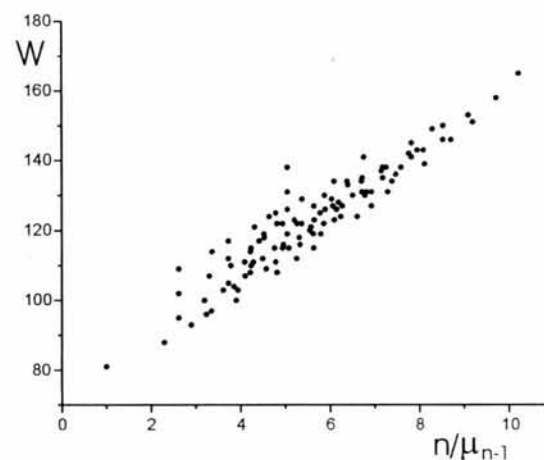

Fig. 4—Analogous data as in Fig. 3 for general trees with $n=10$ vertices.

Table 1—Correlation coefficients for the correlation between the Wiener index and the term n/μ_{n-1} for n-vertex chemical trees and n-vertex trees. Chemical trees are trees in which no vertex has degree greater than 4.

n	Chemical trees	General trees
6	0.993	0.993
7	0.918	0.958
8	0.945	0.966
9	0.937	0.952
10	0.931	0.943
11	0.938	0.941
12	0.939	0.938
13	0.936	0.933
14	0.937	0.932
15	0.936	0.929
16	0.935	0.928
17	0.934	0.926
18	0.934	0.925
19	0.933	0.924
20	0.932	0.922

From the data given in Table 1 (as well as from the examples shown in Figs 3 and 4) we see that W and n/μ_{n-1} are linearly correlated, and that the quality of this correlation is essentially the same for both chemical and general trees. Although the correlation is not particularly good, in all cases studied the correlation coefficients are well above 0.9. The correlation slowly weakens with increasing number n of vertices.

In Figs 3 and 4 an outlier can be seen. It corresponds to 4-propylheptane. By scrutinizing our data for higher values of n we noticed that the points for two molecular graphs significantly deviate from the regression line. We denote them by T_n and Q_n. The first of these exceptional chemical trees (T_n) was identified as the n-vertex tree possessing a single vertex of degree 3 to which three linear branches are attached, having length as equal as possible. The second exceptional chemical tree (Q_n) is the n-vertex tree possessing a single vertex of degree 4 to which four linear branches are attached, having length as equal as possible. For instance, for $n=15$, T_n and Q_n are the molecular graphs of 5-butylundecane and 5,5-dipropylnonane. Another example is found in Fig. 5.

The exceptional nature of the chemical trees T_n and Q_n is hardly at all seen for smaller values of n. Therefore, in order to test their exceptional nature we constructed a set of 100 randomly chosen trees with very large ($n=48$) number of vertices, and added to it the trees T_{48} and Q_{48}. The respective plot is shown in Fig. 5. The two outliers are now clearly visible, thus corroborating our hypothesis.
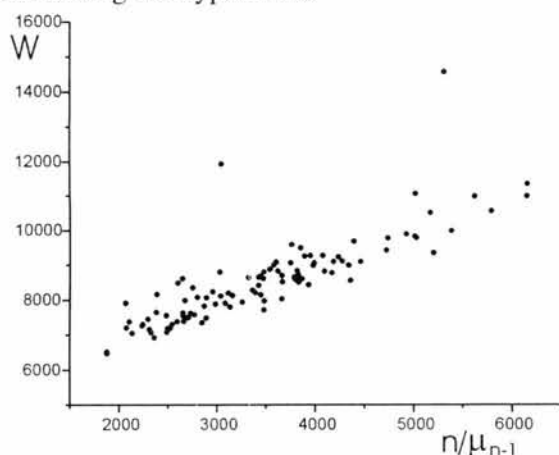

Fig. 5—Analogous data as in Figs. 3 & 4 for a set of 100 randomly constructed trees with $n=48$ vertices, plus two specially chosen trees: one with a single vertex of degree 3 to which linear branches of length 15, 16 and 16 are attached (T_{48}), and another with a single vertex of degree 4 to which linear branches of length 12, 12, 12 and 11 are attached (Q_{48}). These trees pertain to the two obvious outliers.

For reasons which we do not fully understand, in the exceptional trees T_n and Q_n the term n/μ_{n-1} reproduces a much smaller fraction of the Wiener index than in the case of other trees. Thus, in the trees T_n and Q_n the algebraic connectivity has an unusually small effect on the value of the Wiener index.

Bounds for the Wiener and Kirchhoff indices

As well known[19,20], the Wiener index is equal to the sum of distances between all pairs of vertices of the molecular graph, where "distance" means the length of the shortest path connecting the respective two vertices. In 1993 Klein and Randic[24] conceived another type of distance, the so-called "resistance distance". It is equal to the electrical resistance between the respective two vertices, assuming that any two adjacent vertices are connected by a resistance of 1 Ohm. The sum of resistance distances between all pairs of vertices of the molecular graph was named the "Kirchhoff index"[25]. In the case of trees the Wiener and the Kirchhoff indices coincide, but for cycle-containing graphs their values differ.

As a kind of surprise it was shown[15] that the Kirchhoff index is precisely equal to the right-hand side of formula (1). This, in particular, means that formula (1) is meaningful not only for trees, but also for all connected (molecular) graphs.

In this section we use formula (1), together with the conditions (6) and (7) to deduce lower and upper bounds for the Kirchhoff index, in terms of the parameters n, m and t. As before, n is the number of vertices, m the number of edges and t the number of spanning trees of the molecular graph G considered.

Using standard methods of calculus, we try to determine positive real numbers μ_1^*, μ_2^*,..., μ_{n-1}^* which satisfy the conditions

$$\sum_{i=1}^{n-1} \mu_i^* = 2m$$

and

$$\prod_{i=1}^{n-1} \mu_i^* = nt \text{ i.e., } \sum_{i=1}^{n-1} \ln \mu_i^* = \ln(nt)$$

and which minimize or maximize the expression

$$Kf^* = n \sum_{i=1}^{n-1} \frac{1}{\mu_i^*}$$

To achieve this goal we construct the auxiliary function Ξ

$$\Xi = n\sum_{i=1}^{n-1}\frac{1}{\mu_i^*} + \alpha\left(\sum_{i=1}^{n-1}\mu_i^* - 2m\right) + \beta\left(\sum_{i=1}^{n-1}\ln\mu_i^* - \ln(nt)\right)$$

find its derivatives with respect to μ_i^*, $i=1,2,...,n-1$, and set them equal to zero. This results in

$$-\frac{n}{(\mu_i^*)^2} + \alpha + \beta\frac{1}{\mu_i^*} = 0$$

i.e.

$$\alpha(\mu_i^*)^2 + \beta\mu_i^* - n = 0 \qquad \qquad ... (11)$$

which is just a quadratic equation, having two roots x and y. According to the nature of the problem we are examining, the roots x and y must be positive real numbers. They must not be equal, because $x=y$ would require that the condition

$$t = \frac{1}{n}\left(\frac{2m}{n-1}\right)^{n-1}$$

to be satisfied, which certainly cannot be obeyed by molecular graphs because they possess a relatively small number of spanning trees. (This condition is obeyed by the complete graph.)

Hence the roots x and y of Eq. (11) must be different. Without the loss of generality we assume that $x>y$.

Thus, the numbers μ_i^*, $i=1,2,...,n-1$, assume only two different values: x and y. Assuming that k of them are equal to x and the remaining $n-1-k$ are equal to y, we have

$$Kf^* = n\left(\frac{k}{x} + \frac{n-1-k}{y}\right)$$

where x and y are obtained by solving the system of equations:

$$kx + (n-1-k)y = 2m$$
$$k\ln x + (n-1-k)\ln y = \ln(nt)$$

Table 2—Lower (Kf_L) and upper (Kf_U) bounds for the Kirchhoff index, computed according to Eqs. (12) and (13), for some chemically realistic molecular graphs with n vertices, m edges and t spanning trees. A—alkanes; B—monocycloalkanes with a five- and a six-membered ring ($t=5$ and $t=6$, respectively); C—bicyclic systems with non-condensed rings; D—bicyclic systems with condensed rings. Bicyclic molecular graphs with two five-, a five- and a six-, and two six-membered rings have respectively $t=25$, 30 and 36 (if not condensed) and $t=24$, 29 and 35 (if condensed).

type	n	m	t	Kf_L	Kf_U	t	Kf_L	Kf_U	t	Kf_L	Kf_U
A	7	6	1	36.00	91.62						
	8	7	1	49.00	178.75						
	9	8	1	64.00	356.07						
	10	9	1	81.00	693.46						
	11	10	1	100.00	1369.79						
	12	11	i	121.00	3054.14						
	13	12	1	144.00	5480.85						
B	5	5	5	9.70	10.54	6	8.99	9.34			
	6	6	5	17.06	22.31	6	16.16	19.63			
	7	7	5	26.44	44.52	6	25.35	38.82			
	8	8	5	37.82	86.98	6	36.54	75.22			
	9	9	5	51.21	170.21	6	49.74	145.57			
	10	10	5	66.60	336.93	6	64.94	287.16			
	11	11	5	83.99	669.08	6	82.14	561.85			
C	9	10	25	40.72	81.35	30	39.52	70.88	36	38.33	62.30
	10	11	25	54.52	156.75	30	53.14	135.15	36	51.77	116.66
	11	12	25	70.32	310.82	30	68.75	260.78	36	67.21	223.34
	12	13	25	88.12	617.60	30	86.37	526.19	36	84.64	446.20
	13	14	25	107.91	1294.21	30	105.98	1079.77	36	104.06	868.67
	14	15	25	129.71	2983.42	30	127.59	2400.25	36	125.50	1733.74
	15	16	25	153.50	8408.82	30	151.20	4916.68	36	148.92	3487.69
D	8	9	24	29.15	43.87	29	28.08	38.62	35	27.04	34.33
	9	10	24	41.00	83.90	29	39.74	72.58	35	38.51	63.45
	10	11	24	54.84	162.17	29	53.39	138.86	35	51.98	119.17
	11	12	24	70.68	323.64	29	69.04	269.26	35	67.45	229.18
	12	13	24	88.52	646.17	29	86.69	546.18	35	84.91	446.20
	13	14	24	108.35	1294.21	29	106.34	1079.77	35	104.36	928.65
	14	15	24	130.19	2983.42	29	127.98	2400.25	35	125.82	2011.46

It can be shown that W^* is a monotonically increasing function of k, implying that the choice $k=1$ results in a lower bound for the Kirchhoff index:

$$Kf_L = n\left(\frac{1}{x} + \frac{n-2}{y}\right) \qquad \cdots (12)$$

with x and y determined via

$$x + (n-2)y = 2m \quad \& \quad xy^{n-2} = nt \quad \& \quad x > y$$

whereas the choice $k=n-2$ yields an upper bound

$$Kf_U = n\left(\frac{n-2}{x} + \frac{1}{y}\right) \qquad \cdots (13)$$

with x and y determined via

$$(n-2)x + y = 2m \quad \& \quad x^{n-2}y = nt \quad \& \quad x > y$$

The numerical values of the lower and upper bounds for Kf, for some chemically relevant choices of the parameters n, m, t are given in Table 2. It should be noted that the lower bound for the Wiener index of acyclic systems (the case $m=n-1$, $t=1$) is the best possible[26] and is equal to $(n-1)^2$; the upper bound is somewhat above the best possible value[26] $n(n^2-1)/6$. For cyclic systems (the cases $m > n-1$, $t>1$) no bounds for the Kirchhoff index, better than those presently obtained, are known.

References

1 Graovac A, Gutman I & Trinajstić N, *Topological approach to the chemistry of conjugated molecules* (Springer-Verlag, Berlin), 1977.
2 Dias J R, *Molecular orbital calculations using chemical graph theory* (Springer-Verlag, Berlin), 1993.
3 Gutman I, Lee S L, Chu C H & Luo Y L, *Indian J Chem*, 33A (1994) 603.
4 Godsil C D & Gutman I, *ACH Models Chem*, 136 (1999) 503.
5 Gutman I, Gineityte V, Lepović M & Petrović M, *J Serb chem Soc*, 64 (1999) 673.
6 Merris R & Gutman I, *Z Naturforsch*, 55a (2000) 973.
7 Gutman I, Vidovic D, *Indian J Chem*, 41A (2002) 893.
8 Gutman I, Vidović D & Stevanović D, *J Serb chem Soc*, 67 (2002) 407.
9 Ivanciuc O, *Rev Roum Chim*, 38 (1993) 1499.
10 Mohar B, Babić D & Trinajstić N, *J chem Inf Comput Sci*, 33 (1993) 153.
11 Kunz M, *MATCH - Commun math Chem*, 32 (1995) 221.
12 Kunz M, *J chem Inf Comput Sci*, 36 (1996) 822.
13 Chan O, Lam T K & Merris R, *J chem Inf Comput Sci*, 37 (1997) 762.
14 Marković S, Gutman I & Bančević Ž, *J Serb chem Soc*, 60 (1995) 633.
15 Gutman I & Mohar B, *J chem Inf Comput Sci*, 36 (1996) 982.
16 Cvetković D, Doob M & Sachs H, *Spectra of graphs - theory and application* (Academic Press, New York), 1980.
17 Merris R, *Lin Algebra Appl*, 197-198 (1994) 143.
18 Merris R, *Lin multilin Algebra*, 39 (1995) 19.
19 Wiener H, *J Am chem Soc*, 69 (1947) 17.
20 Gutman I, Yeh Y N, Lee S L & Luo Y L, *Indian J Chem*, 32A (1993) 651.
21 Fiedler M, *Czech math J*, 23 (1973) 298.
22 Grone R & Merris R, *Czech math J*, 37 (1987) 660.
23 Grone R & Merris R, *Graphs Combin*, 6 (1990) 229.
24 Klein D J & Randić M, *J math Chem*, 12 (1993) 81.
25 Bonchev D, Balaban A T, Liu X & Klein D J, *Int J Quantum Chem*, 50 (1994) 1.
26 Gutman I, *Indian J Chem*, 36A (1997) 128.