MATCH
Communications in Mathematical and in Computer Chemistry

no. 50, February 2004
MATCDY (50) 83- 92 (2004)

ISSN 0340 - 6253

THE FIRST ZAGREB INDEX 30 YEARS AFTER

Ivan Gutman[a] and Kinkar Ch. Das[b]

[a] *Faculty of Science, University of Kragujevac, Serbia & Montenegro*
e-mail: gutman@knez.uis.kg.ac.yu

[b] *Department of Mathematics, Indian Institute of Technology, Kharagpur 721302, W. B., India*
e-mail: kinkar@lycos.com

(Received August 26, 2003)

Abstract

The graph invariant M_1, known under the name *first Zagreb index*, equal to the sum of the squares of the degrees of the vertices of the respective (molecular) graph, was first considered by Trinajstić and one of the present authors in 1972. We show that M_1 is related to a number of other quantities of interest in chemical graph theory, and point out some of its general mathematical properties. In particular, the trees with minimal and maximal M_1 are the path and the star, respectively. Several lower and upper bounds for M_1 are reported.

INTRODUCTION

In a paper [1], published more than 30 years ago, Trinajstić and one of the present authors examined the dependence of total π-electron energy on molecular structure. They found that in the approximate expressions for total π-electron energy two terms occur:

$$M_1 = \sum_{\text{vertices}} (\delta_u)^2$$

$$M_2 = \sum_{\text{edges}} \delta_u \cdot \delta_v$$

with δ_u standing for the degree (= number of first neighbors) of the vertex u of the molecular graph. It was immediately recognized that M_1 and M_2 reflect the extent of branching of the molecular skeleton (and are thus responsible for the decrease of total π-electron energy with increasing branching). This viewpoint was then elaborated in the paper [2]. Eventually, M_1 and M_2 were named [3] the *first Zagreb-Group index* and the *second Zagreb-Group index*, respectively. These names were later abbreviated into *first Zagreb index* and *second Zagreb index* [4]; note that some authors call M_1 the *Gutman index* [4].

Recently, on the occasion of the 30th anniversary of the Zagreb indices, a paper [5] was published in which their main properties were summarized. However, the authors of [5] failed to mention a number of known results, especially for M_1. The purpose of the present article is to fill this gap. We also offer a few hitherto unpublished results for M_1.

NOTATION AND BASIC DEFINITIONS

Let G be a graph, possessing n vertices and m edges. In what follows we assume that $n \geq 2$. The vertices of G are denoted by $v_1, v_2, \ldots, v_n$. The number of first neighbors of the vertex v_i is its degree, and will be denoted by δ_i. The vertices of G are labelled so that $\delta_1 \geq \delta_2 \geq \cdots \geq \delta_n$. The average of the degrees of the vertices adjacent to v_i is denoted by m_i.

The first Zagreb index, pertaining to the graph G, is defined as

$$M_1 = M_1(G) = \sum_{i=1}^{n} (\delta_i)^2 \tag{1}$$

an expression which should be compared with the well known result

$$\sum_{i=1}^{n} \delta_i = 2m \ . \tag{2}$$

Let n_k denote the number of vertices of G, possessing degree k. Then, of course,

$$\sum_{k \geq 0} n_k = n \tag{3}$$

and, in view of (2),

$$\sum_{k \geq 0} k\, n_k = 2m \ . \tag{4}$$

Then the first Zagreb index conforms to the relation

$$M_1 = \sum_{k \geq 0} k^2 n_k \ . \tag{5}$$

A graph in which $\delta_1 = \delta_2 = \cdots = \delta_n = r$ is said to be regular, of degree r. For regular graphs, $M_1 = n\, r^2$ and therefore these graphs are of little interest as far as the first Zagreb index is concerned. Yet, it should be mentioned that in recent studies in chemical graph theory regular graphs of degree 3 were much investigated, because these provide a representation for fullerenes, nanotubes and similar carbon-atom clusters.

An acyclic connected graph is called a tree. For all trees, $\delta_{n-1} = \delta_n = 1$.

The tree with $\delta_1 = 2$ is the path P_n; the tree with $\delta_1 = n - 1$ is the star S_n. For the path, $n_1 = 2$, $n_2 = n - 2$, and $n_k = 0$ for $k \neq 1, 2$. Therefore,

$$M_1(P_n) = 2 \times 1^2 + (n-2) \times 2^2 = 4n - 6 \ .$$

For the star, $n_1 = n - 1$, $n_{n-1} = 1$, and $n_k = 0$ for $k \neq 1, n - 1$. Therefore,

$$M_1(S_n) = (n-1) \times 1^2 + 1 \times (n-1)^2 = n(n-1) \ .$$

In molecular graphs it must be $\delta_1 \leq 4$. Trees with the property $\delta_1 \leq 4$ are called chemical trees. For chemical trees, as a special case of (5),

$$M_1 = n_1 + 4\, n_2 + 9\, n_3 + 16\, n_4$$

as pointed out in [5].

For additional concepts from graph theory and chemical graph theory see the books [6, 7, 8, 9].

CONNECTION OF M_1 TO OTHER GRAPH INVARIANTS

1. Number of pairs of incident edges.

Denote by $\ell(G,2)$ the number of pairs of edges which have a common endpoint. There are $\binom{\delta_i}{2}$ pairs of edges meeting at vertex v_i. Therefore,

$$\ell(G,2) = \sum_{i=1}^{n} \binom{\delta_i}{2} = \frac{1}{2} M_1 - m$$

where we have taken into account relation (2).

2. Number of pairs of independent edges.

Denote by $m(G,2)$ the number of pairs of edges which have no vertex in common. Then, because the total number of pairs of edges is $\binom{m}{2}$, we get

$$m(G,2) = \binom{m}{2} - \ell(G,2) = \frac{1}{2} m(m+1) - \frac{1}{2} M_1 \ .$$

Note that $m(G,2)$ is also the second coefficient of the matching polynomial.

3. Fourth coefficient of the characteristic polynomial.

Using the Sachs theorem [7] we immediately get

$$a_4 = m(G,2) - 2\,Q = \frac{1}{2} m(m+1) - \frac{1}{2} M_1 - 2\,Q$$

where Q denotes the number of 4-membered cycles contained in G. This expression was first time reported in [1] and was eventually deduced by many other authors.

4. Fourth spectral moment.

The sum of the fourth powers of the eigenvalues of G, denoted here by S_4, can be calculated from a_4, using the Newton identities [1, 9]. A more direct way is to notice that S_4 is equal to the number of self-returning walks of length 4. There are three

types of such self-returning walks: (u,v,u,v,u), involving two vertices, (u,v,w,v,u), involving three vertices, and walks along a four-membered cycle: (u,x,y,z,u). The number of self-returning walks of the first type is equal to twice the number of edges because (u,v,u,v,u) and (v,u,v,u,v) both pertain to the same edge uv. The number of self-returning walks of the second type is equal to four times the number of pairs of incident edges because (u,v,w,v,u), (v,w,v,u,v), (v,u,v,w,v), and (w,v,u,v,w) all pertain to the same pair of incident edges uv and vw. The number of self-returning walks of the third type is $8Q$, because we may start at any vertex of the four-membered cycle and go in two opposite directions. Thus,

$$S_4 = 2m + 4\ell(G,2) + 8Q = 2M_1 - 2m + 8Q \ .$$

This expression was first time reported in [10].

5. Second coefficient of the Laplacian characteristic polynomial.

Using the Kel'mans theorem [7, 11, 12] we get

$$c_2 = 3\ell(G,2) + 4m(G,2) = 2m^2 - m - \frac{1}{2}M_1 \ .$$

6. Second Laplacian spectral moment.

If $\mu_1, \mu_2, \ldots, \mu_n$ are the Laplacian eigenvalues of G, then

$$c_2 = \sum_{i<j}\mu_i\mu_j = \frac{1}{2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\mu_i\mu_j - \sum_{i=1}^{n}(\mu_i)^2\right] = \frac{1}{2}\left[\left(\sum_{i=1}^{n}\mu_i\right)\left(\sum_{j=1}^{n}\mu_j\right) - \sum_{i=1}^{n}(\mu_i)^2\right] =$$

$$= \frac{1}{2}\left[(2m)^2 - \sum_{i=1}^{n}(\mu_i)^2\right]$$

which together with the previously deduced expression for c_2 yields

$$\sum_{i=1}^{n}(\mu_i)^2 = M_1 + 2m \ .$$

7. Variance of vertex degrees.

The variance $Var(\delta)$ of vertex degrees was recently examined in the mathematical literature [13]. In view of the fact that by definition,

$$Var(\delta) = \frac{1}{n}\sum_{i=1}^{n}\left(\delta_i - \frac{2m}{n}\right)^2$$

we obtain by direct calculation

$$Var(\delta) = \frac{M_1}{n} - \left(\frac{2m}{n}\right)^2 .$$

LOWER AND UPPER BOUNDS FOR M_1

Bearing in mind the relations between M_1 and numerous other graph invariants, it is not fully surprising that the quantity M_1 attracted the attention of mathematicians [14, 15, 16]. They, however, studied M_1 unrelated to its (earlier) chemical applications.

In this section we state several bounds for M_1, existing in the mathematical literature and, most probably, not known to scholars active in chemical graph theory.

Recall first that for a graph with n vertices and m edges, the average value of the vertex degrees is $2m/n$. Let $p = \lfloor 2m/n \rfloor$, i. e., p is the integer part of $2m/n$. Then the first Zagreb index is bounded from both below and above by expressions depending solely on the parameters n and m:

$$2(2p+1)m - p(p+1)n \leq M_1 \leq m\left(\frac{2m}{n-1} + n - 2\right) . \qquad (6)$$

The right-hand side inequality (6) is due to de Caen [15] whereas the left-hand side inequality is obtained by one of the present authors [16].

The inequalities (6) hold for all graphs. In the case of trees (both chemical and non-chemical), $m = n - 1$ and $p = 1$ and (6) is significantly simplified:

$$n(n-1) \geq M_1 \geq 4n - 6$$

Earlier we have shown that $n(n-1) = M_1(S_n)$ and $4n - 6 = M_1(P_n)$. Bearing this in mind, we arrive at a noteworthy (yet not surprising) result:

Theorem 1. *Among n-vertex trees, the star S_n has maximum and the path P_n minimum value of the first Zagreb index. If T_n is an n-vertex tree, different from the star or path, then $M_1(S_n) > M_1(T_n) > M_1(P_n)$.*

In the case of molecular graphs which contain cycles (when $m \geq n$), the parameter p is always equal to 2, except in the case of molecular graphs of fullerenes, where it

is equal to 3. Therefore,

$$M_1 \geq 10m - 6n$$

except in the case of fullerene graphs, for which the lower bound in (6) is equal to $14m - 12n$, which is equal to $9n = 3^2 n$. This, in turn, is the M_1-value of any n-vertex regular graph of degree 3.

Another very simple upper bound for the first Zagreb index reads [16]:

$$M_1 \leq m(m+1) .$$

Knowing the value of d_1, the greatest vertex degree in the graph G, the bounds (6) can be somewhat sharpened [16].

Lower bound in parameters n, m, d_1:

Let $p^* = \lfloor 2(m - d_1)/(n-1) \rfloor$ and $t = 2(m - d_1) - p^*(n-1)$.

(a) If $d_1 > n - 1 - t$, then

$$M_1 \geq (2p^* + 1)(2m - d_1) + (d_1)^2 + 2(d_1 - n + 1 + t) - p^*(p^* + 1)(n-1)$$

(b) If $d_1 \leq n - 1 - t$, then

$$M_1 \geq (2p^* + 1)(2m - d_1) + (d_1)^2 - p^*(p^* + 1)(n-1) \qquad (7)$$

Upper bound in parameters n, m, d_1:

$$M_1 \leq m\left(\frac{2m}{n-1} + n - 2\right) - d_1\left(\frac{4m}{n-1} - 2d_1 - \frac{n+1}{n-1}d_1 + n - 1\right)$$

Note that for molecular graphs, for which $d_1 \leq 4$, only the case (7) is applicable. For acyclic, monocyclic and bicyclic molecular graphs, as well as tricyclic graphs without vertex of degree 4, $p^* = 1$. For all other molecular graphs, $p^* = 2$.

Some more complicated bounds for the first Zagreb index can be found in the paper [16]. Of them we mention here only the following:

Upper bound in parameters n, m, d_1, m_1:

$$M_1 \leq m\left(\frac{2m}{n-1} + n - 2\right) - d_1\left(\frac{4m}{n-1} - 2m_1 - \frac{n+1}{n-1}d_1 + n - 1\right)$$

AN ALTERNATIVE PROOF OF THEOREM 1

In this section we give a proof of Theorem 1 which uses arguments from linear programming. With pertinent modifications this method can be used to characterize unicyclic, bicyclic, etc. graphs with minimum and maximum first Zagreb index.

Consider a graph G without isolated vertices (i. e., $n_0 = 0$), in which the greatest vertex degree is D. Write Eqs. (3) and (4) in the form

$$n_1 + n_2 = n - \sum_{k=3}^{D} n_k$$

$$n_1 + 2n_2 = 2m - \sum_{k=3}^{D} k\, n_k$$

and solve them in variables n_1 and n_2. This gives

$$n_1 = 2n - 2m + \sum_{k=3}^{D} (k-2)\, n_k$$

$$n_2 = 2m - n - \sum_{k=3}^{D} (k-1)\, n_k$$

which substituted back into (5) results in

$$M_1 = 6m - 2n + \sum_{k=3}^{D} (k-1)(k-2)\, n_k \ . \tag{8}$$

Because all the terms in the summation on the right-hand side of equality (8) are positive-valued, we see that M_1 will attain its minimal value if all n_k, $k = 3, 4, \ldots, D$, are equal to zero (of course, under the assumption that a graph with such properties does exist).

In other words, the graph with minimum first Zagreb index should possess only vertices of degree 1 and 2. In the case of trees, such graph does exist and is unique: it is the path P_n. Hence, the path P_n has minimum M_1-value among all n-vertex trees.

This proves the first half of Theorem 1. In order to verify its second half, write Eqs. (3) and (4) in the form

$$n_1 + n_D = n - \sum_{k=2}^{D-1} n_k$$

$$n_1 + D\, n_D = 2m - \sum_{k=2}^{D-1} k\, n_k$$

and solve them in variables n_1 and n_D. This gives

$$n_1 = \left(n - \frac{2m-n}{D-1}\right) + \sum_{k=2}^{D-1} \left(\frac{k-1}{D-1} - 1\right) n_k$$

$$n_D = \frac{2m-n}{D-1} - \sum_{k=2}^{D-1} \frac{k-1}{D-1} n_k$$

which substituted back into (5) results in

$$M_1 = n + (D+1)(2m-n) - \sum_{k=2}^{D-1} [(D-k)(k-1)]\, n_k \,. \tag{9}$$

Again, all the terms in the summation on the right-hand side of equality (9) are positive-valued. Because the entire summation has a negative sign, we see that M_1 will attain its maximal value if all n_k , $k = 2, 3, \ldots, D-1$, are equal to zero, provided that a graph with such properties does exist. In other words, the graph with maximum first Zagreb index should possess only vertices of degree 1 and D.

In the case of general trees, D should be set equal to $n-1$. If so, then for any n there exists a unique tree with the property $n_1, n_{n-1} \neq 0$ and $n_2, n_3, \ldots, n_{n-2} = 0$: it is the star S_n. Hence, the star S_n has maximum M_1-value among all n-vertex trees.

This completes the proof of Theorem 1.

Acknowledgement. This research was supported by the Ministry of Sciences, Technologies and Development of Serbia, within the Project no. 1389.

References

[1] I. Gutman, N. Trinajstic, Graph theory and molecular orbitals. Total π-electron energy of alternant hydrocarbons, *Chem. Phys. Lett.* **17** (1972) 535–538.

[2] I. Gutman, B. Ruščić, N. Trinajstić, C. F. Wilcox, Graph theory and molecular orbitals. XII. Acyclic polyenes, *J. Chem. Phys.* **62** (1975) 3399–3405.

[3] A. T. Balaban, I. Motoc, D. Bonchev, O. Mekenyan, Topological indices for structure-activity correlations, *Topics Curr. Chem.* **114** (1983) 21–55.

[4] R. Todeschini, V. Consonni, *Handbook of Molecular Descriptors*, Wiley-VCH, Weinheim, 2000.

[5] S. Nikolić, G. Kovačević, A. Milićević, N. Trinajstić, The Zagreb indices 30 years after, *Croat. Chem. Acta* **76** (2003) 113–124.

[6] F. Harary, *Graph Theory*, Addison-Wesley, Reading, 1969.

[7] D. Cvetković, M. Doob, H. Sachs, *Spectra of Graphs – Theory and Application*, Academic Press, New York, 1980.

[8] N. Trinajstić, *Chemical Graph Theory*, CRC Press, Boca Raton, 1983; 2nd revised ed. 1992.

[9] I. Gutman, O. E. Polansky, *Mathematical Concepts in Organic Chemistry*, Springer-Verlag, Berlin, 1986.

[10] I. Gutman, Some topological properties of benzenoid systems, *Croat. Chem. Acta* **46** (1974) 209–215.

[11] N. Trinajstić, D. Babić, S. Nikolić, D. Plavšić, D. Amić, Z. Mihalić, The Laplacian matrix in chemistry, *J. Chem. Inf. Comput. Sci.* **34** (1994) 368–376.

[12] I. Gutman, D. Vidović, B. Furtula, Chemical applications of the Laplacian spectrum. VII. Studies of the Wiener and Kirchhoff indices, *Indian J. Chem.* **42A** (2003) 1272–1278.

[13] F. K. Bell, A note on the irregularity of graphs, *Lin. Algebra Appl.* **161** (1992) 45–54.

[14] L. A. Székely, L. H. Clark, R. C. Entringer, An inequality for degree sequences, *Discr. Math.* **103** (1992) 293–300.

[15] D. de Caen, An upper bound on the sum of squares of degrees in a graph, *Discr. Math.* **185** (1988) 245–248.

[16] K. C. Das, Sharp bounds for the sum of the squares of the degrees of a graph, *Kragujevac J. Math.* **25** (2003) 000–000.