Estrada Index of Benzenoid Hydrocarbons

Ivan Gutman and Slavko Radenković

Faculty of Science, University of Kragujevac, P. O. Box 60, 34000 Kragujevac, Serbia

Reprint requests to Prof. I. G.; Fax: +381 34 335040; E-mail: gutman@kg.ac.yu

Z. Naturforsch. **62a,** 254 – 258 (2007); received December 11, 2006

A structure-descriptor EE, recently proposed by Estrada, is examined. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of the molecular graph, then $EE = \sum_{i=1}^{n} e^{\lambda_i}$. In the case of benzenoid hydrocarbons with n carbon atoms and m carbon-carbon bonds, EE is found to be accurately approximated by means of the formula $a_1 n \cosh\left(\sqrt{2m/n}\right) + a_2$, where $a_1 \approx 1.098$ and $a_2 = -0.64$ are empirically determined fitting constants. Within classes of benzenoid isomers (which all have equal n and m), the Estrada index is linearly proportional to the number of bay regions.

Key words: Estrada Index; Benzenoid Hydrocarbons; Molecular Graph; Spectrum (of Graph).

1. Introduction

In this paper we are concerned with a molecular structure-descriptor of benzenoid hydrocarbons, that we refer to as the *Estrada index*. It is defined as follows.

Let G be the molecular graph of a benzenoid hydrocarbon [1 – 5]. Let n and m be, respectively, the number of vertices and edges of G. Then the formula of the underlying hydrocarbon is C_nH_{3n-2m} and n must be even [3 – 5].

The eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of the adjacency matrix of G are said to be the eigenvalues of G and to form the spectrum of G. These will be labelled so that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n.$$

The basic properties of graph eigenvalues can be found in [6].

The Estrada index is defined as

$$EE = EE(G) = \sum_{i=1}^{n} e^{\lambda_i}. \tag{1}$$

Although introduced quite recently [7], the Estrada index has already found numerous applications. It was used to quantify the degree of folding of long-chain molecules, especially proteins [7 – 9]. Another, fully unrelated, application of EE was put forward by Estrada and Rodríguez-Velázquez [10, 11]. They showed that EE provides a measure of the average centrality of complex (communication, social, metabolic, etc.) networks. In addition to this, in a recent work [12] a connection between EE and the concept of extended atomic branching was suggested.

Until now only some straightforward mathematical properties of the Estrada index were established [10, 11, 13], but its dependence on molecular structure has not been investigated. The present paper is aimed at contributing towards filling this gap.

In what follows we use the functions *hyperbolic cosine* and *hyperbolic sine*, defined as usual as

$$\cosh(x) = \frac{e^x + e^{-x}}{2} \quad \text{and} \quad \sinh(x) = \frac{e^x - e^{-x}}{2},$$

respectively.

Our starting points are the well known relations for the eigenvalues of benzenoid graphs [2, 6]

$$\lambda_i + \lambda_{n+1-i} = 0 \ \text{ for } i = 1, 2, \ldots n, \tag{2}$$

known as the pairing theorem [2, 14, 15], and

$$\sum_{i=1}^{n/2} (\lambda_i)^2 = m. \tag{3}$$

Because of (2), half of the eigenvalues of a benzenoid graph G are positive (or zero) and the other half negative (or zero), implying [11]

$$EE(G) = 2 \sum_{i=1}^{n/2} \cosh(\lambda_i).$$

2. A McClelland-Type Bound for the Estrada Index

The method, by which a bound EE^* for EE is deduced in this section, is fully analogous to a method used long time ago [16] for estimating the total π-electron energy E. In the notation specified above, for a benzenoid graph G,

$$E = E(G) = 2\sum_{i=1}^{n/2} \lambda_i. \tag{4}$$

For a fixed value of m, i. e., assuming that condition (3) is obeyed, an extremal value E^* of E is obtained by employing the Lagrange multiplier technique: An auxiliary function F is constructed as

$$F := 2\sum_{i=1}^{n/2} \lambda_i - \alpha \left(\sum_{i=1}^{n/2} (\lambda_i)^2 - m \right)$$

and the condition

$$\frac{\partial F}{\partial \lambda_k} = 0$$

imposed for all $k = 1, 2, \ldots, n/2$. This leads to

$$2 - 2\,\alpha\,\lambda_k = 0, \quad \text{i. e.} \quad \lambda_k = \alpha,$$

which combined with (3) yields

$$\lambda_k = \sqrt{\frac{2m}{n}} \quad \text{for } k = 1, 2, \ldots, n/2,$$

and substituted back into (4) results in

$$E^* = \sqrt{2mn}.$$

This is just the famous McClelland upper bound for total π-electron energy [17 – 19].

Applying an analogous reasoning, we construct the auxiliary function

$$FF := 2\sum_{i=1}^{n/2} \cosh(\lambda_i) - \alpha \left(\sum_{i=1}^{n/2} (\lambda_i)^2 - m \right)$$

and impose

$$\frac{\partial FF}{\partial \lambda_k} = 0$$

for $k = 1, 2, \ldots, n/2$. This results in

$$2\sinh(\lambda_k) - 2\,\alpha\,\lambda_k = 0.$$

It is easily seen that for $\alpha > 0$ the equation $\sinh(x) - \alpha x = 0$ has a single positive-valued solution. Denote it by x_0. Then $\alpha = \sinh(x_0)/x_0$.

Now, from $\lambda_k = x_0$ for $k = 1, 2, \ldots, n/2$ and relation (3) we readily obtain

$$x_0 = \sqrt{\frac{2m}{n}},$$

and therefore

$$EE^* = n \cosh\left(\sqrt{\frac{2m}{n}} \right). \tag{5}$$

Note that $2m/n$ is the average vertex degree of the graph G . Therefore for benzenoid graphs, $2m/n > 2$, and thus $x_0 > \sqrt{2}$.

Curiously, however, in contrast to E^*, which is an upper bound for the total π-electron energy [17], the estimate EE^*, formula (5), is a lower bound for the Estrada index. To see this we examine the Hessian matrix $\mathcal{H}(FF)$ of the function FF.

Because of

$$\frac{\partial FF}{\partial \lambda_k} = 2\sinh(\lambda_k) - 2\,\alpha\,\lambda_k,$$

one has

$$\frac{\partial^2 FF}{\partial \lambda_k^2} = 2\cosh(\lambda_k) - 2\,\alpha$$

and

$$\frac{\partial^2 FF}{\partial \lambda_k\, \lambda_{k'}} = 0$$

for $k \neq k'$. Therefore $\mathcal{H}(FF)$ is a diagonal matrix whose all diagonal elements are equal to

$$2\cosh(x_0) - 2\,\alpha, \quad \text{i. e.} \quad 2\cosh(x_0) - 2\,\frac{\sinh(x_0)}{x_0}.$$

Now,

$$2\cosh(x_0) - 2\,\frac{\sinh(x_0)}{x_0} = \frac{(x_0 - 1)\,\mathrm{e}^{x_0} + (x_0 + 1)\,\mathrm{e}^{-x_0}}{x_0},$$

which for $x_0 > \sqrt{2}$ is evidently positive-valued. Thus all eigenvalues of $\mathcal{H}(FF)$ are positive-valued and, consequently, the extremal value EE^* is a minimum. We thus proved:

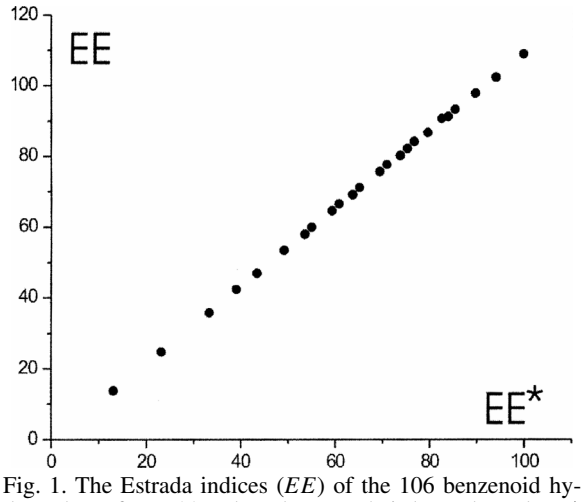

Fig. 1. The Estrada indices (EE) of the 106 benzenoid hydrocarbons from [22], plotted versus their lower bound EE^* according to (5). For details see text.

Theorem 1. The Estrada index of a benzenoid hydrocarbon with n carbon atoms and m carbon-carbon bonds is always greater than $n\cosh(\sqrt{2m/n})$.

3. An (n,m)-Type Approximation for the Estrada Index

The McClelland upper bound E^* for a total π-electron energy E provides an excellent approximate formula for E of conjugated molecules [17] of the form $E \approx a'_1 E^* + a'_2$. In fact, in the case of benzenoid hydrocarbons, this formula with $a'_1 = 0.898$ and $a'_2 = 0.45$ happens to be the best (n,m)-type approximation for E [20, 21]. In view of this, we examined how well an expression of the form $a_1 EE^* + a_2$ would approximate the Estrada index. The quality of this approximation is seen in Figure 1. Indeed, the correlation between EE and EE^* is almost perfectly linear.

Least-squares fitting, using the standard data set of 106 Kekuléan benzenoid hydrocarbons from [22] (same as employed in [20, 21] and elsewhere [23, 24]), yields the regression line

$$\begin{aligned} &EE = a_1 EE^* + a_2, \\ &a_1 = 1.098 \pm 0.001,\ a_2 = -0.64 \pm 0.08, \end{aligned} \tag{6}$$

with a remarkably high correlation coefficient 0.99994. In the sample examined the average relative error of the approximation (6) is 0.19%, and the maximal observed relative error 0.83%.

One evident conclusion from the above result is that the gross part of the Estrada indices of benzenoid systems is determined by the parameters n and m. In other words, the Estrada indices of benzenoid isomers differ very little.

In the subsequent section we examine these small differences of the EE values of isomeric benzenoids, and try to see which is the main structural feature that is responsible for them.

4. On Estrada Indices of Benzenoid Isomers

Expanding the function e^x into a power series, and using the definition (1) of the Estrada index, one readily arrives at [7, 10]

$$EE(G) = \sum_{k\geq 0} \frac{M_k(G)}{k!},$$

where M_k is the k-th spectral moment of the molecular graph G,

$$M_k = M_k(G) = \sum_{i=1}^{n} (\lambda_i)^k .$$

For all graphs, $M_0 = n$, $M_1 = 0$, and $M_2 = 2m$ [2, 6]. For all bipartite graphs (and thus also for the molecular graphs of benzenoid hydrocarbons), $M_k = 0$ for odd k. We thus have

$$EE(G) = n + m + \frac{1}{24}M_4 + \frac{1}{720}M_6 + \frac{1}{40320}M_8 + \cdots, \tag{7}$$

which implies

$$EE(G) \approx n + m + \frac{1}{24}M_4 + \frac{1}{720}M_6. \tag{8}$$

The dependence of the first few even spectral moments of benzenoid hydrocarbons on the molecular structure is known [25–27]. In particular,

$$\begin{aligned} M_4(G) &= 18m - 12n, \\ M_6(G) &= 158m - 144n + 48 + 6b, \end{aligned}$$

where b is the number of bay regions [3, 26, 27]. When these are substituted back into (7) and (8), we obtain

$$\begin{aligned} EE(G) = \frac{1}{720}(1418m + 216n + 48 + 6b) \\ + \text{higher order terms}, \end{aligned}$$

Table 1. Statistical data for correlations of the form $EE = a_1'' b + a_2''$ for sets of benzenoid isomers with n carbon atoms and m carbon-carbon bonds; b is the number of bay regions. All sets considered contain all possible isomers, equal to *N.I.* The respective correlation coefficient is R.

n	m	*N.I.*	a_1''	a_2''	R
18	21	5	0.01072 ± 0.00008	46.9326 ± 0.0001	0.99992
22	26	12	0.01072 ± 0.00005	58.0114 ± 0.0001	0.99990
24	29	13	0.01085 ± 0.00004	64.5537 ± 0.0001	0.99991
26	31	36	0.01080 ± 0.00003	69.0900 ± 0.0001	0.99987
26	32	9	0.01070 ± 0.00008	71.0966 ± 0.0002	0.99980

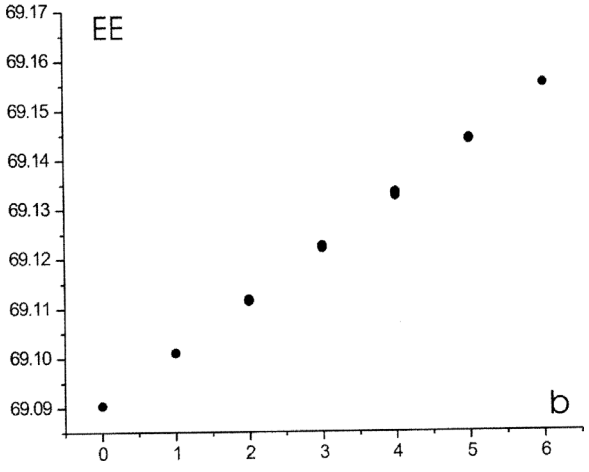

Fig. 2. The Estrada indices (EE) of the 36 benzenoid isomers with the formula $C_{26}H_{16}$, plotted versus their number of bay regions (b). For details see Table 1 and text.

i. e.,

$$EE(G) \approx \frac{1}{720}(1418m + 216n + 48 + 6b). \quad (9)$$

According to the approximation (9), within classes of isomeric benzenoids, the Estrada index is an increasing linear function of the parameter b, and the slope of the respective line is (almost) independent of n and m and (nearly) equal to $1/120 = 0.00833$. That this is indeed the case is seen from Fig. 2 and from the data given in Table 1.

5. Concluding Remarks

We deem to have established the main structural features of benzenoid hydrocarbons that determine the value of their Estrada indices. These are, first of all, the parameters n and m that are capable of reproducing some 99.8% of EE. The (n,m)-type approximation $a_1 EE^* + a_2$ with EE^* being given by (5) is maybe not the best possible, but is remarkably accurate. The quality of the linear relation between EE and EE^* is illustrated by Figure 1.

Anyway, the Estrada indices of benzenoid isomers (i. e., species having equal values of n and m) vary only to a very limited extent. The main structural feature influencing these variations is the number b of bay regions. Within sets of benzenoid isomers, EE is an increasing linear function of b. The slope of this function is practically independent of n and m (as seen from the data for a_1'' in Table 1). It is close, yet not equal to the slope predicted by means of a truncated expansion of EE in terms of spectral moments.

We dare to conclude that the structure dependence of the Estrada indices of benzenoid hydrocarbons is now almost completely understood.

[1] N. Trinajstić, Chemical Graph Theory, CRC, Boca Raton 1983.
[2] I. Gutman and O. E. Polansky, Mathematical Concepts in Organic Chemistry, Springer-Verlag, Berlin 1986.
[3] I. Gutman and S. J. Cyvin, Introduction to the Theory of Benzenoid Hydrocarbons, Springer-Verlag, Berlin 1989.
[4] M. Zander, Z. Naturforsch. **45a**, 1041 (1990).
[5] M. Zander, Topics Curr. Chem. **153**, 101 (1990).
[6] D. Cvetković, M. Doob, and H. Sachs, Spectra of Graphs – Theory and Application, Academic Press, New York 1980.
[7] E. Estrada, Chem. Phys. Lett. **319**, 713 (2000).
[8] E. Estrada, Bioinformatics **18**, 697 (2002).
[9] E. Estrada, Proteins **54**, 727 (2004).
[10] E. Estrada and J. A. Rodríguez-Velázquez, Phys. Rev. **71E**, 56103 (2005).
[11] E. Estrada and J. A. Rodríguez-Velázquez, Phys. Rev. **72E**, 46105 (2005).
[12] E. Estrada, J. A. Rodríguez-Velázquez, and M. Randić, Int. J. Quantum Chem. **106**, 823 (2006).
[13] I. Gutman, E. Estrada, and J. A. Rodríguez-Velázquez, Croat. Chem. Acta **79** (in press).
[14] I. Gutman, Z. Naturforsch. **39a**, 152 (1984).
[15] R. B. Mallion and D. H. Rouvray, J. Math. Chem. **5**, 1 (1990).
[16] I. Gutman, MATCH Commun. Math. Comput. Chem. **14**, 71 (1983).
[17] B. J. McClelland, J. Chem. Phys. **54**, 640 (1971).
[18] A. Graovac, I. Gutman, P. E. John, D. Vidović, and I. Vlah, Z. Naturforsch. **56a**, 307 (2001).
[19] H. Fripertinger, I. Gutman, A. Kerber, A. Kohnert, and D. Vidović, Z. Naturforsch. **56a**, 342 (2001).
[20] I. Gutman, Topics Curr. Chem. **162**, 29 (1992).

[21] I. Gutman and T. Soldatović, MATCH Commun. Math. Comput. Chem. **44**, 169 (2001).
[22] R. Zahradnik and J. Pancir, HMO Energy Characteristics, Plenum Press, New York 1970.
[23] J. Cioslowski and I. Gutman, Z. Naturforsch. **41a**, 861 (1986).
[24] I. Gutman, J. H. Koolen, V. Moulton, M. Parac, T. Soldatović, and D. Vidović, Z. Naturforsch. **55a**, 507 (2000).
[25] J. Cioslowski, Z. Naturforsch. **40a**, 1167 (1985).
[26] G. G. Hall, Theor. Chim. Acta **70**, 323 (1986).
[27] S. Marković and I. Gutman, J. Mol. Struct. (Theochem.) **235**, 81 (1991).