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

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

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 $[2, 6]$

$$\lambda_i + \lambda_{n+1−i} = 0 \quad \text{for} \quad i = 1, 2, \ldots, n; \quad (2)$$

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

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

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 $\pi$-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 $\pi$-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

$$2 \sinh(\lambda_k) - 2 \alpha \lambda_k = 0.$$
3. An \((n,m)\)-Type Approximation for the Estrada Index

The McClelland upper bound \(E^*\) for a total \(\pi\)-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 \(a_1 E E^* + a_2\) 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

\[
EE = a_1 E E^* + a_2,
\]

\[
a_1 = 1.098 \pm 0.001, \quad a_2 = -0.64 \pm 0.08,
\]

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

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

which implies

\[
EE(G) \approx n + m + \frac{1}{24} M_4 + \frac{1}{720} M_6.
\]

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

\[
M_4(G) = 18m - 12n,
\]

\[
M_6(G) = 158m - 144n + 48 + 6b,
\]

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

\[
EE(G) = \frac{1}{720} (1418m + 216n + 48 + 6b) + \text{higher order terms},
\]
The respective correlation coefficient is $R$. All sets considered contain all possible isomers, equal to $N.I.$ The respective line is (almost) independent of $n$ 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 $a_i^*$ 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.

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 influence 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 $a_i^*$ in Table 1).

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.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$m$</th>
<th>$N.I.$</th>
<th>$a_1^*$</th>
<th>$a_2^*$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>21</td>
<td>5</td>
<td>0.01072 ± 0.00008</td>
<td>46.9326 ± 0.0001</td>
<td>0.9992</td>
</tr>
<tr>
<td>22</td>
<td>26</td>
<td>12</td>
<td>0.01072 ± 0.00005</td>
<td>58.0114 ± 0.0001</td>
<td>0.9990</td>
</tr>
<tr>
<td>24</td>
<td>29</td>
<td>13</td>
<td>0.01085 ± 0.00004</td>
<td>64.5537 ± 0.0001</td>
<td>0.9991</td>
</tr>
<tr>
<td>26</td>
<td>31</td>
<td>36</td>
<td>0.01080 ± 0.00003</td>
<td>69.0900 ± 0.0001</td>
<td>0.9987</td>
</tr>
<tr>
<td>26</td>
<td>32</td>
<td>9</td>
<td>0.01070 ± 0.00008</td>
<td>71.0966 ± 0.0002</td>
<td>0.9980</td>
</tr>
</tbody>
</table>

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.

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.


