Approximating total π-electron energy of phenylenes in terms of spectral moments

Svetlana Marković
Faculty of Science, University of Kragujevac, P.O.B. 60, YU-34000 Kragujevac, Serbia and Montenegro

Received 24 January 2003

The total π-electron energy of phenylenes is approximated by means of a linear combination of the first few spectral moments of both molecular and line graphs. The two sets of moments produce very similar results, with very high accuracy. It is found that over 99.8% of the HMO total π-electron energy of phenylenes is determined by the number of carbon atoms. Number of bay regions plays a significant role in the dependence of E on molecular topology of phenylenes.

Chemical graphs are the basic tools used in applying the techniques of mathematical graph theory to the specific problems of chemistry. The molecular graph of a conjugated hydrocarbon is a graph, representing the carbon-atom skeleton of the molecule. Consider a molecular graph G with n vertices and m edges. The line graph of G (denoted by L) can be constructed in the following manner: a vertex of the line graph is associated with each edge of G. The edges of L are now obtained by connecting those vertices of L which represent the adjacent edges in G. Molecular graphs reflect the adjacency of carbon atoms (existence of chemical bonds), whereas line graphs reflect the adjacency of these bonds. The graphs G and L can be represented with their adjacency matrices A and E, respectively.

An important theoretical feature of conjugated molecules is the HMO total π-electron energy (E). Within the Hückel theory it is defined as twice the sum of the positive eigenvalues of the adjacency matrix of the respective molecular graph. E is one of the most extensively studied π-electron characteristics of conjugated hydrocarbons, and particularly of benzenoid hydrocarbons. The dependence of E on molecular structure of benzenoid hydrocarbons is nowadays understood to a great extent. In this paper, an analogous quantitative investigation for a class of non-benzenoid hydrocarbons, known as phenylenes has been reported.

Phenylenes are a class of conjugated hydrocarbons composed of four- and six-membered rings (hexagons), where the hexagons are adjacent only to four-membered rings, and each four-membered ring is adjacent to a pair of hexagons. Recently, the experimental chemistry of phenylenes has made a remarkable development, owing mainly to the work of the Volhardt’s research group. Some examples of phenylenes are presented with their structural formulas and respective molecular and line graphs in Fig. 1.

It is well known that the physico-chemical behavior of conjugated hydrocarbons is highly dependent on molecular topology, and it can be successfully described by means of topological invariants, such as spectral moments. Spectral moments of G and L, \( M_k \), and \( \mu_k \), are equal to the number of self-returning walks of length k contained in molecular and line graphs, respectively. They are defined as:

\[
M_k = \text{Tr}[A^k] \quad \text{and} \quad \mu_k = \text{Tr}[E^k]
\]

where Tr denotes the traces of the respective matrix. Spectral moments of molecular graphs have been employed in the physical chemistry of solid state and theoretical chemistry of conjugated hydrocarbons. Spectral moments of line graphs have found remarkable applications in the quantitative structure-property (QSPR) and quantitative structure-activity (QSAR) studies, and in the theoretical chemistry of aromatic compounds. Recently, spectral moments of both molecular and line graphs of polycyclic aromatic hydrocarbons were applied to kinetics of extraction with supercritical carbon dioxide.

Spectral moments of phenylenes

In all applications of spectral moments it is necessary to know how they depend on molecular
structure. For this reason, efforts have been made to establish topological formulas for spectral moments of various classes of molecules. Particular attention has been devoted to benzenoid hydrocarbons\textsuperscript{1,6,16-22} and phenylenes\textsuperscript{23,24}. Since each self-returning walk embraces a certain structural fragment of the molecular graph, it is possible to express $M_k$ as a linear combination of the number of these fragments. For phenylenes the expressions for the first few spectral moments can be significantly simplified. Thus, for a phenylene with $h$ hexagons it holds\textsuperscript{23,24}:

\begin{align}
M_0 &= 6h \quad \ldots \quad (1) \\
M_1 &= 16h - 4 \quad \ldots \quad (2) \\
M_2 &= 80h - 44 \quad \ldots \quad (3) \\
M_6 &= 484h - 352 + 6h_{A_3} \quad \ldots \quad (4) \\
M_8 &= 3136h - 2636 + 128b + 16h_{A_2} \quad \ldots \quad (5) \\
M_{10} &= 20976h - 19504 + 1680B + 3370C 
+ 5060F + 6750G + 580h_{A_3} \quad \ldots \quad (6)
\end{align}

The meanings of the symbols $b$, $h_{A_3}$, $B$, $C$, $F$ and $G$ are explained by the example given in Fig. 2. An $A_3$-type hexagon is a structural detail of branched phenylenes; it is connected to three four-membered rings. $h_{A_3}$ represents the numbers of hexagons of this type. In Fig. 2 some structural fragments of the perimeter of phenylenes are also depicted: bay, cove, fjord and gulf. $B$, $C$, $F$ and $G$ denote the numbers of such fragments, respectively. Then, the number of bay regions ($b$) is determined as $b = B + 2C + 3F + 4G$.

In the case of phenylenes the expressions for the first few spectral moments of the line graph bear remarkable resemblance to the spectral moments of the molecular graph. These were found to be as follows\textsuperscript{26}:

\begin{align}
\mu_0 &= 8h - 2 \quad \ldots \quad (7) \\
\mu_1 &= 0 \quad \ldots \quad (8) \\
\mu_2 &= 28h - 16 \quad \ldots \quad (9) \\
\mu_3 &= 24h - 24 \quad \ldots \quad (10) \\
\mu_4 &= 180h - 144 + 4b \quad \ldots \quad (11) \\
\mu_5 &= 360h - 360 + 20b \quad \ldots \quad (12) \\
\mu_6 &= 1600h - 1480 + 108b + 12h_{A_3} \quad \ldots \quad (13) \\
\mu_7 &= 4480h - 4536 + 504b + 56h_{A_3} \quad \ldots \quad (14) \\
\mu_8 &= 16708h - 16608 + 2268B + 4544C 
+ 6820F + 9096G + 416h_{A_3} \quad \ldots \quad (15) \\
\mu_9 &= 53160h - 55176 + 9792B + 19656C 
+ 29520F + 39384G + 2268h_{A_3} \quad \ldots \quad (16)
\end{align}

The notation in Eqs. (7)-(16) is the same as in the case of Eqs. (1)-(6). It should be pointed out that another more general approach to the spectral moments of the line graph has also been elaborated\textsuperscript{27-29}.

The significance of the formulas (1)-(16) lies in the fact that they provide a quantitative insight into the structural features that determine the numerical values of spectral moments of phenylenes. This fact was recently used to express the heat of formation of phenylenes in terms of their structural fragments\textsuperscript{26}. An interesting observation was made in this work: $M_2$ is
It was also observed that the change in the standard set of 106 benzenoid hydrocarbons\textsuperscript{32}. On the basis of these benzenoids, corresponding phenylenes have been constructed. The molecular graphs of these phenylenes are given elsewhere\textsuperscript{26}.

The coefficients, $a_0$, $a_2$, ..., $a_{p+2}$, as well as $b_0$, $b_2$, $b_4$, ..., $b_{q+1}$, have been determined by least squares fitting, and may be obtained from the author upon request. The correlation coefficients, average relative errors and sample dispersions for Eqs. (17) and (18) have also been calculated. These results, showing the quality of the approximations (17) and (18), are presented in Table 1, except those where linear dependency is encountered (see the text below).

The introduction of a new fitting parameter, $a_q$, into (17) or $b_q$ into (18), necessarily increases the correlation coefficients and decreases the average errors and sample dispersions. In order to check whether this improvement is statistically significant or not, the standard F-test was used\textsuperscript{33}. The results of the F-test are presented in Table 1. The quantity $f$ for Eq. (17) is equal to the ratio of the sample dispersion for a certain $p$ and that for $p+2$, divided by the respective value of the F-distribution at the confidence level of 99%. Therefore, if $f$ is greater than unity, then the change $p \rightarrow p+2$ significantly improves the quality of the respective approximation; if $f$ is below unity, then, from a statistical point of view, the previous and new equations are equivalent, and preference should be given to the former (since it contains fewer empirical parameters). The quantity $f$ for Eq. (18) is defined analogously.

By increasing the parameters $p$ and $q$ the accuracies of Eqs. (17) and (18) are necessarily enhanced, implying that the correlation coefficients are increased, whereas the relative average errors and sample dispersions are decreased. For the first few values of $p$ and $q$ the quality of the approximations significantly increases, until the relative errors fall to 0.01%, and the correlation coefficients reaches the value of 0.99999988. With further increase in the values of $p$ and $q$, these improvements fade, and for $p > 16$ and $q > 12$, they become immaterial.

As a consequence of the mentioned linear dependence of some $M_k$ of phenylenes, the change $p=0 \rightarrow p=2$, as well as $p=2 \rightarrow p=4$, will not cause any improvement in the accuracy of Eq. (17). This implies that the correlation coefficients, average relative errors and sample dispersions will remain completely unchanged. Similarly, the change

<table>
<thead>
<tr>
<th>Eqs. (17) and (18)</th>
<th>R</th>
<th>ARE (%)</th>
<th>$10^6$SD</th>
<th>$f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>p=5, q=5</td>
<td>0.99999987</td>
<td>0.032</td>
<td>423.7</td>
<td>11.52</td>
</tr>
<tr>
<td>p=10, q=10</td>
<td>0.99999988</td>
<td>0.030</td>
<td>319.3</td>
<td>0.66</td>
</tr>
<tr>
<td>p=15, q=15</td>
<td>0.99999992</td>
<td>0.026</td>
<td>286.0</td>
<td>0.56</td>
</tr>
<tr>
<td>p=20, q=20</td>
<td>0.99999998</td>
<td>0.020</td>
<td>253.7</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Since the molecular graphs of phenylenes are bipartite, their odd spectral moments are equal to zero. The line graphs of phenylenes contain triangles and larger odd-membered rings, so that their odd spectral moments (except $\mu_p$) are non-zero. For these reasons, $p$ in Eq. (17) is necessarily even, and the term $b_1\mu_1$ in Eq. (18) is left out.

The calculations have been performed on the set that is obtained in the following way: Forty nine catacondensed benzenoids have been selected from

Results

Here we report the results of investigation of the dependence of $E$ of phenylenes on both spectral moments of $G$ and $L$. For this purpose $E$ is approximated by means of the following polynomials:

$$E = a_0 M_0 + a_2 M_2 + \ldots + a_{p+2} M_{p+2} \quad p=0, 2, 4, \ldots, 16$$

...(17)

$$E = b_0 \mu_0 + b_2 \mu_2 + b_4 \mu_4 + \ldots + b_{q+1} \mu_{q+1} \quad q=0, 2, 3, \ldots, 12$$

...(18)

In this work, a similar quantitative approach is used to express the total $\pi$-electron energy of phenylenes in terms of their structural details.
q=0 \rightarrow q=2$, as well as $q=2 \rightarrow q=3$, does not affect the accuracy of Eq. (18). For this reason, the respective data concerning Eqs. (17) and (18) are not given in Table 1.

**Discussion**

Due to the mutual linear dependence between lower members of $M_k$ and $\mu_k$, Eqs. (17) & (18) produce identical results up to $p=12$ and $q=10$. This implies that one can achieve the same accuracy of the approximation for $E$ of phenylenes by using, for example, either

$$E = a_0 M_0 + a_6 M_6 + a_8 M_8 + a_10 M_{10} + a_{12} \quad \text{or} \quad \ldots \quad (19)$$

$$E = b_0 \mu_0 + b_4 \mu_4 + b_6 \mu_6 + b_{10} \mu_{10} + b_{12} \quad \ldots \quad (20)$$

For higher values of $p$ and $q$ ($p \geq 14$, $q \geq 11$), Eqs. (17) and (18) produce different but very similar results, which are only slightly better for Eq. (18). These findings are in good accord with the results of investigation of the total $\pi$-electron energy of benzenoids based on linear combination of spectral moments.

Substituting Eqs (1), (4), (5) and (10) with the respective values for $p$ into Eq. (19), as well as Eqs. (7), (10), (13) and (15) with corresponding values for $q$ into Eq. (20), the following quantitative expressions for $E$ of phenylenes are obtained:

$$E = 8.5752h+4.1559b-0.0086a_6-4.0824B-8.1891C-12.2958F-16.4025G-0.5860 \quad \ldots \quad (21)$$

$$E = 8.5322h+6.9656b-0.0098a_6-6.8947B-13.8138C-20.7328F-27.6518G-0.5472 \quad \ldots \quad (22)$$

These approximations are, certainly, very similar, and they reproduce about 99.97% of $E$ of phenylenes. As an illustration, a plot of the HMO values for $E$ versus those calculated via Eq. (21), is given in Fig. 3. Since the $E$-values of isomers are very similar, the points form groups in the plot, where each group of points corresponds to a certain number of hexagons (from $h=2$ to $h=9$).

Both Eqs. (17) and (18) reproduce $E$ of phenylenes with very high accuracy. They also elucidate the dependence of $E$ of phenylenes on their structure.

When only $M_0$ or $\mu_0$ are used in the spectral-moment expansion, the relative error is less than 0.2%, which reveals that the major part of $E$ (more than 99.8%) is determined by the size of the molecules, expressed by the number of hexagons. Bearing in mind that $n = 6h$ (as usual, $n$ stands for the number of vertices) it turns out that more than 99.8% of $E$ of phenylenes is determined by the number of carbon atoms. On introducing $M_6$ into (17), as well as $\mu_4$ or $\mu_5$ into (18) the quality of the approximations is drastically improved. This is clearly revealed by the results of the F-test, yielding 11.52 in both cases. Since the number of bay regions is a structural feature influencing $M_6$, $\mu_4$, and $\mu_5$, one can conclude that $b$ plays a significant role in the dependence of $E$ on molecular topology of phenylenes. Another significant improvement of the approximations is achieved for $p=12$ or $q=11$. These approximate formulae reproduce $E$ of phenylenes with the accuracy of about 99.99%. Unfortunately, the dependencies of $M_{14}$ and $\mu_{11}$ on molecular structure of phenylenes are not properly understood and we are unable to distinguish the respective structural features of phenylenes that make a significant contributions to $E$.

**Acknowledgement**

This work is supported by the Ministry of Science and Technology of Serbia, Project No. 1448.
References