Graph Theory in Chemistry

II. Graph-Theoretical Description of Heteroconjugated Molecules

A. Graovac, O. E. Polansky, N. Trinajstić, and N. Tyutyulkov

(Z. Naturforsch. 30a, 1696–1699 [1975]; received October 25, 1975)

The Sachs’ formula which relates the structure of a (vertex- and edge-)weighted graph and its characteristic polynomial is given. Weighted graphs are used to represent conjugated molecules containing the variety of heteroatoms and heterobonds. Vertices and edges in such molecular graphs have different weights following the deviation of the corresponding (Hückel) Coulomb and resonance integrals from the standard (benzene) values.

Until very recently 1-7 graph theory, in the case of conjugated systems, has been applied (in the framework of HMO theory) to hydrocarbons only 8. In HMO theory graphs may be interpreted in two different ways: a) as the structural graph of the conjugated system considered 9-11 or b) as a graph in which vertices correspond to the basis functions used for the construction of the HMO’s (basis graph) 12. The concept of basis graphs may be extended to quantum chemical methods other than HMO theory. However, under the approximations of the simple HMO theory, the basis graph is the same as the structural graph. A molecular (carbon)-graph (or C-graph) \( G_N \) related to a conjugated molecule with \( N \) carbon atoms corresponds to its \( \tau \)-network. After labelling the vertices of \( G_N \) the adjacency matrix, \( A \), of \( G_N \) may be constructed following the definition 13:

\[
A_{rs} = \begin{cases} 
1 & \text{if vertices } r \text{ and } s \text{ are connected,} \\
0 & \text{otherwise.} 
\end{cases} \tag{1}
\]

The topological matrix of a conjugated molecule appearing in Hückel theory has been shown to be equivalent to the adjacency matrix of a corresponding molecular graph 9,14. Thus, using an appropriate scaling, the Hückel secular determinant may be expressed as follows:

\[
P(G_N; x) = \det (xI - A) = \sum_{n=0}^{N} a_n x^{N-n} \tag{2}
\]

and the Hückel problem is completely reduced to the adjacency matrix eigenvalue problem. According to the scale used the eigenvalues \( x_i \) \( (i = 1, 2, \ldots, N) \) are given by,

\[
x_i = (E_i - \alpha_0)/\beta_{CC} \tag{3}
\]

where \( E_i \) \( (i = 1, 2, \ldots, N) \) is the Hückel molecular orbital energy, \( \alpha_0 \) the Coulomb integral of a carbon atom, and \( \beta_{CC} \) is the standard carbon-carbon bond resonance integral 15.

The coefficients \( a_n \) of the characteristic polynomial \( P(G_N; x) \) may be found without going through the procedure of solving determinant but solely from the topology of the molecular graph. Such a procedure has been suggested first by Coulson 16 and has been put in the concise mathematical form by several mathematicians 17, but a form given by Sachs 18 is very convenient for our discussion. The Sachs’ formula which covers molecular graphs corresponding to conjugated hydrocarbons, is given by,

\[
a_n = \sum_{s \in S_n} (-)^{c(s)} 2^{r(s)} \quad 0 \leq n \leq N. \tag{4}
\]

Here \( s \) is a Sachs graph 19 (that is such a graph which has no other components 8 but complete graphs of valency one and cycles (rings)), \( S_n \) is the set of all Sachs graphs with \( n \) vertices, \( c(s) \) and \( r(s) \) are, respectively, the total number of components and the number of ring components in \( s \). Summation is over all Sachs graphs (with \( n \) vertices). \( N \) is the number of vertices in the graph \( G_N \). The Sachs’ formula as defined above may be very conveniently used for the evaluation of individual polynomial coefficients of the interest (e.g. \( a_N \)) and the structure-reactivity relationships are developed on this basis 19-22.

Recently, the Sachs’ formula was extended to vertex-weighted ** graphs, that is to say those graphs which may be used to represent some heteroconjugated molecules 1, 4, 5. A vertex-weighted graph is a graph which has one (or more) of its vertices distinguished from others. These vertices of a different “type” are weighted and the weight is related ** Originally 1, 4, 5 these graphs were named “rooted” graphs, a name which is not quite appropriately used there, because this type of graphs belong to trees 21, 24.

* Two whom the reprint requests should be addressed.
to the value of $\alpha_X$. $\alpha_X$ is the Coulomb integral of a heteroatom $X$ and it is defined as follows:

$$a_X = a_C + a_{CC}.$$  

(5)

A non-zero value of $a$ describes the difference between the heteroatom (weighted vertex) and carbon atom (ordinary vertex). This is reflected on diagonal elements of the “adjacency” matrix. In order to avoid any confusion we call this type of “adjacency” matrix the weight matrix, $W$. The same term will be used also in the case of vertex- and edge-weighted graphs.

In the case where only ordinary vertices are present, the diagonal elements of adjacency matrix, $A$, of a graph $G_N$ are all zero [see Eq. (1)], that is $\text{Tr} A(G_N) = 0$. However, in the case of vertex-weighted graphs $G_{N,vw}$, $\text{Tr} W(G_{N,vw}) = \sum a_i$, where the sum is over all weighted vertices in $G_{N,vw}$. The characteristic polynomial $P(G_{N,vw}; x)$ of a vertex-weighted graph $G_{N,vw}$ (for the case of vertices being all of the same weight, $a$) is given by,

$$P(G_{N,vw}; x) = \sum_{n=0}^{N} \sum_{s \in S_n} (-1)^n c(s) 2^{r(s)} (x - a)^{t(G_N-s)} \cdot x^{N-n-t(G_N-s)}$$  

(6)

where the symbols already mentioned in Eq. (4) have their previous meaning. Graph $(G_N - s)$ is obtained by deletion of (Sachs’) graph $s$ from graph $G_N$ and $t(G_N - s)$ is the number of weighted vertices in (particular) vertex-weighted graph $(G_N - s)$. The use of formula (6) has been extensively described in our earlier work. However, it is limited to a class of hetero-conjugated molecules with the same type of heteroatom and the differences in the resonance integrals $\beta_{CC}$ and $\beta_{CX}$ are not considered.

Now, we have extended the Sachs’ formula to cover the vertex- and edge-weighted graphs $G_{N,vew}$, that is to graphs which represent heteroconjugated molecules containing bonds with the values of resonance integrals different from $\beta_{CC}$, and which contain several different heteroatoms. These graphs contain one (or more) edges which differ in some way from others. The resonance integral $\beta_{XY}$ of an edge X-Y connecting two differently weighted vertices X and Y may be written as:

$$\beta_{XY} = b_{XY} \beta_{CC}$$  

(7)

where $b_{XY}(\neq 1)$ describes the difference between the standard (benzene) carbon-carbon resonance integral, $\beta_{CC}$, and the resonance integral of X-Y bond, $\beta_{XY}$. This means that values other than 1 (denoted as $b$) have been introduced to the off-diagonal elements of the weight matrix.

The characteristic polynomial $P(G_{N,vew}; x)$ of a vertex- and edge-weighted graph $G_{N,vew}$ is given by,

$$P(G_{N,vew}; x) = \sum_{n=0}^{N} \sum_{s \in S_n} (-1)^n c(s) 2^{r(s)} \cdot \prod_{(G-s)} b_i \prod_{(x-a_j)} x^{N-n-t(G_N-s)}$$  

(8)

where the symbols already mentioned in Eqs. (4) and (6) have their previous meaning. Here we wish to point out that Eq. (8) is similar to Coates’ topological formula which is used in circuit theory. The product $\prod b_i$, which corresponds to (vertex- and edge-) weighted Sachs graph $s$, is obtained by multiplication of $b_i$’s in the following way. For a $\Gamma$ component of a Sachs graph $s$, $b$ is squared, and for a ring component of $s$, $b$’s are multiplied “around” the ring. As an example consider the product $\prod b_i$ for the following (vertex- and edge-) weighted Sachs graph $s$:

$$\prod_{(G-s)} b_i = b_0^2(b_1 b_2 b_3 b_4).$$

The product $\prod_{(G-s)} (x - a_j)$ may be obtained as follows. The factor $(x - a_j)$ must be taken into account as many times as there are (weighted) vertices with the weight $a_j$ in the graph $(G_N - s)$, and it needs to be done for every “type” of heteroatom present in the graph $(G_N - s)$. In the case that there is only one type of (weighted) vertex (with the weight $a$) present in $G_{N,vew}$, the product $\prod_{(G_N-s)} (x - a_j)$ reduces to $(x - a)_{(G_N-s)}$, as in Equation (6).

As an example, the characteristic polynomial of $I$ is evaluated below.

$$\prod_{(G-s)} (x - a_j) = (x - a)_{(G_N-s)}.$$

In the corresponding molecular graph, $G_{N,vew}(I)$, there are three different “types” of vertices and
Table 1. The characteristic polynomial of a molecular graph representing $I$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$S_n$</th>
<th>$P(S_n; x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\emptyset$</td>
<td>$x^2(x-h_0^2) (x-h_1) = x^4 - x^3(h_0 + h_1) + x^2 h_0 h_1$</td>
</tr>
<tr>
<td>2</td>
<td>${ (o-o-o), (o-h_0-o), (o-h_1-o) }$</td>
<td>$-(-)^{12} b_0^2 (x-h_0) (x-h_1) x^2 + (+)^{12} b_0^2 (x-h_0) x^2 + + (-)^{12} b_0^2 x^2 = -x^2(1+b_0^2+b_1^2) + x(h_0+h_1) + b_0 b_1 h_1 - h_0 h_1$</td>
</tr>
<tr>
<td>4</td>
<td>${ (o-o-o-o), (h_0-o-o), (h_1-o-o) }$</td>
<td>$(-)^{30} b_1^2 x^2 = b_1^2$</td>
</tr>
</tbody>
</table>

$P(G_{vew}(I); x) = \sum_{n} P(S_n; x) = x^4 - x^3(h_0 + h_1) - x^2(1+b_0^2+b_1^2 - h_0 h_1) + x(h_0+h_1) + b_0 b_1 h_1 + b_1^2-h_0 h_1$

Three differently weighted edges.

Nitrogen and oxygen atoms are weighted differently ($a_0$ and $a_1$) than carbon atoms ($a=0$). Similarly, bonds C – N and N – O are indicated with weight $b_0$ and $b_1$, respectively. Using $\beta_{cc}$ as the unity, the C – C bond, of course, has a weight of unity 8. The evaluation of $P(G_{vew}(I); x)$ is demonstrated in Table 1. The characteristic polynomial of $G_{vew}(I)$ for $a_1 = 0$ and $b_0 = b_1 = 1$ reduces to the polynomial which corresponds to a system $G_{vew}(II)$:

$$P(G_{vew}(II); x) = x^4 - a_0 x^3 - 3 x^2 + a_0 x + 1$$

and for $a_0 = 0$ reduces, of course, to the polynomial which corresponds to butadiene system:

$$P(G_{vew}(II); x) = x^4 - 3 x^2 + 1$$

The study of molecules by use of quantum chemical methods other than HMO theory, as in the studies on the energy spectrum of substituted polyenes 26, 27, lead also to the use of (vertex- and edge-) weighted (complete) graphs 12 (but in that case they are interpreted as basis graphs). The formula for the characteristic polynomial of uniformly weighted complete graph (all vertices and all edges have the same weight) is derived 12, in the manner different than the present one, as in the formula for the case of regularly and irregularly weighted complete graphs (i.e. vertex- and edge-weighted complete graphs). The latter graphs can be interpreted as a joint 28 of (some number of) uniformly weighted subgraphs. The formula for the characteristic polynomial is derived only for the case when the weights of all edges, connecting any pair of such subgraphs, have the same value (depending on the pair under consideration). Although, all these formulas 12 can be understood as the special case of formula (B), they are more convenient in treatment of the removal of degeneracy under substitution, under inclusion of enlarged basis set, under inclusion of those molecular integrals which are usually neglected in the simple MO theories, and so on.

The special case of formula (B) is arising 6, 7 when the off-diagonal elements can take values: +1, 1, or 0, which is the case in the so called Möbius or anti-Hückel systems 6, 7, 29, 30. The similar graphs appear also in our attempts to estimate the energy gap of hydrocarbon polymers 31.
2 J. V. Knop, N. Trinajstić, I. Gutman, and L. Klasinc, Naturwiss. 60, 475 [1973].
5 R. B. Mallion, A. J. Schwenk, and N. Trinajstić, in Recent Advances in Graph Theory, Edited by M. Fiedler, in press.
6 A. Graovac and N. Trinajstić, J. Mol. Structure, in press.
12 O. E. Polansky, Match, in press.
25 C. L. Coates, IRE Trans. Circuit Theory CT-6, 170 [1959].
28 A. A. Zykov, Mat. Sbornik 24, 163 [1949].

a The Rugjer Bošković Institute, P.O.B. 1016, 41001 Zagreb, Croatia, Yugoslavia.
b Institut für Strahlenchemie im Max-Planck-Institut für Kohlenforschung, D-4330 Mülheim a. d. Ruhr, B.R.D.
c Institute of Organic Chemistry, Bulgarian Academy of Sciences, Sofia 13, Bulgaria.

* To whom the reprint request should be addressed.