CROATICA CHEMICA ACTA CCACAA 51 (2) 133—140 (1978)

CCA-1106

YU ISSN 0011-1643
539.19:516
Original Scientific Paper

The Determinant of the Adjacency Matrix of the Graph of a Conjugated Molecule

*A. Graovac and I. Gutman**

Received April 28, 1977

»Ruđer Bošković« Institute, P.O.B. 1016, 41001 Zagreb, Croatia, Yugoslavia

Various graph-theoretical techniques for the calculation of the determinant of the adjacency matrix of a molecular graph are discussed.

1. INTRODUCTION

The problem of the relations which exist between the structure of a graph G and the determinant D of its adjacency matrix was solved for the general case by Harary.[1] However, in chemical literature this result was hardly ever used. Moreover, Marcus[2] and Wilcox[3] have independently formulated rules for calculation of D of molecular graphs, which are rather close to Harary's result. Important contributions towards the understanding of the dependence of D on the structure of a graph were given in the chemical literature[4,5] much before Harary's paper appeared.

The theoretical chemistry interest in D of molecular graphs begins with the discovery by Longuet-Higgins[4] that the presence of non-bonding molecular orbitals (NBMO's) in Hückel theory is related to the absence of Kekulé structures in the molecule considered. The dependence of D on the number of Kekulé structures was established soon after that[5] and was elaborated in more detail in refs 3, 6 and 7.

In addition to the simple fact[4] that $D = 0$ is a necessary and sufficient condition for the presence of NBMO's, the actual numerical value of D is shown to parallel the thermodynamic stability of the corresponding molecule (as given within the HMO theory).[3,6,8] D is one of the most important topological factors in bounds[9] and approximate formulas[10] for total π-electron energy (as calculated within the HMO model). Evaluation of D is necessary also in connection with other problems of the topological theory of conjugated systems.[11]

In the present paper we shall present some graph-theoretical techniques for the calculation of the determinant of the adjacency matrix of a molecular graph.

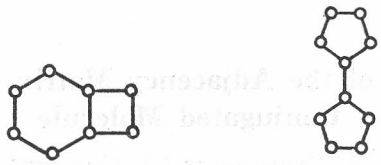
Necessary graph-theoretical concepts and definitions are introduced in section 2, where a convenient notation is also developed. Further details of graph theory and its application to chemistry can be found in refs 12—14.

---

* Present adress: Faculty of Sciences, University of Kragujevac, 34000 Kragujevac, R. Domanovića 12, Serbia, Yugoslavia

2. NOTATION AND DEFINITIONS

In the present work the graph representation of a molecular skeleton will be used.[13,15] Thus for example, benzcyclobutadiene and fulvalene are presented by molecular graphs G_1 and G_2, respectively.



G_1 G_2

Let G be a molecular graph with n vertices $v_1, v_2, \ldots, v_n$. Its adjacency matrix **A** is defined as usual.[11-14]

If $A_{pp} = h_p \neq 0$, we say that there is a loop of weight h_p on the vertex v_p of G. If $A_{pp} = 0$, there is no loop on the vertex v_p. The molecular graphs of hydrocarbons have no loops ($A_{pp} = 0$ for all p), while heteroatoms are represented by a loop on the corresponding vertex.[16]

The determinant of **A** wil be denoted by $D = D(G)$.

In the following we will consider linear graphs[1,17] of G. A linear graph consists of a number of mutually disjoint cycles of length 1, 2, 3, 4, ... Cycles of length 1 are called loops and cycles of length 2 are called edges.

Let the number of cycles of length t in a linear graph L be $c_t(L)$, and $c(L)$ be the total number of components* of L. Then

$$c_1(L) + c_2(L) + c_3(L) + \ldots = c(L)$$

$$c_1(L) + 2c_2(L) + 3c_3(L) + \ldots = n$$

where, of course, n is the number of vertices in L.

In the forthcoming discussion we shall need the following quantities:

$$c_3(L) + c_4(L) + c_5(L) + \ldots = r(L)$$

$$c_1(L) + c_5(L) + c_9(L) + \ldots = r_1(L)$$

$$c_2(L) + c_6(L) + c_{10}(L) + \ldots = r_2(L)$$

$$c_3(L) + c_7(L) + c_{11}(L) + \ldots = r_3(L)$$

$$c_4(L) + c_8(L) + c_{12}(L) + \ldots = r_4(L)$$

$$r_1(L) + r_3(L) = c_{odd}(L)$$

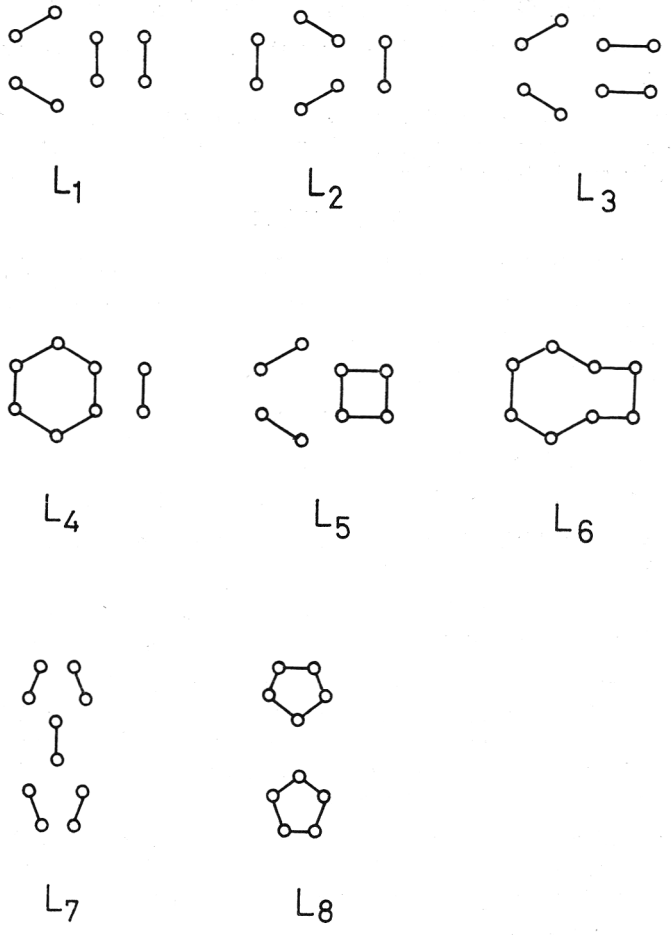
$$r_2(L) + r_4(L) = c_{even}(L)$$

Hence r is the number of cycles of length greater than two; r_1, r_2, r_3, r_4 are the total number of cycles of size $4m+1$, $4m+2$, $4m+3$, $4m$ (m = integer); c_{odd} and and c_{even} are the total number of odd and even cycles, respectively.

All linear graphs of graphs G_1 and G_2 are presented in Chart 1. The benzcyclobutadiene graph G_1 has 6 different linear graphs L (L_1 — L_6). The fulvalene graph G_2 has 2 linear graphs L (L_7, L_8).

* Component is a subgraph which is not connected to other subgraphs.

CHART 1



Linear graphs, the only components of which are isolated edges ($c_1 = c_3 = c_4 = \ldots = 0$, $c_2 = n/2$) are called[13,15,19] Kekulé graphs, since they are in an obvious manner related to the Kekulé structures of the corresponding molecule. For instance L_1, L_2, L_3 and L_7 in Chart 1 are Kekulé graphs.

Kekulé graphs, which play an important role in the subsequent considerations will be denoted by K_1, K_2, ..., K_k, with k being the number of Kekulé structures of the molecular skeleton G.

Let us define now certain sets of linear graphs. The set of all linear graphs of the molecular graph G is denoted by $\mathbf{L} = \mathbf{L}(G)$. The subsets of linear graphs from $\mathbf{L}$ with the property $c_{odd}(L) = 0$ and $c_{odd} > 0$ are denoted by $\mathbf{L}_{even}$ and $\mathbf{L}_{odd}$, respectively. Thus $\mathbf{L}_{even}$ contains the linear graphs which are composed entirely from even cycles and $\mathbf{L}_{odd}$ contains those with at least one odd cycle. Finally, the set of all Kekulé graphs of the molecular graph G is denoted by $\mathbf{K} = \mathbf{K}(G)$.

The *superposition*[7,15,18,20] $K_1 \oplus K_2$ of two Kekulé graphs K_1 and K_2 is the spanning subgraph of G which contains all edges from both K_1 and K_2.

Now we can formulate the following theorems, which are proved in refs 15, 18 and 20.

Theorem 1. The superposition of two Kekulé graphs K_i and K_j is a linear graph. Moreover, every linear graph $L \in \mathbf{L}_{\text{even}}$ can be presented in the form $L = K_i \oplus K_j$. In symbolic notation:

$$\mathbf{L}_{\text{even}} = \mathbf{K} \oplus \mathbf{K}$$

There are $2^{r(L)}$ distinct ordered pairs (K_i, K_j) such that $L = K_i \oplus K_j$.

If the molecule considered is alternant, the molecular graph G is bipartite and has no odd cycles. Then $\mathbf{L}_{\text{even}} = \mathbf{L}$.

Theorem 2. If G is bipartite,

$$\mathbf{L} = \mathbf{K} \oplus \mathbf{K}$$

Corollaries. Let G be a bipartite graph.

a) If $k = 0$, there exists no linear graph of G and $\mathbf{L} = \emptyset$.

b) If $k = 1$, the Kekulé graph is the unique linear graph of G, that is $\mathbf{L} = \{K_1\}$.

c) The existence of two Kekulé graphs in G is a necessary and sufficient condition for the existence of a linear graph L, such that $r(L) > 0$.

From Theorem 1 we can see that some cycles of G can be obtained by superposition of Kekulé graphs. These cycles are called by Randić[21] »*conjugated circuits*« and recently it was shown by the same author[21,22] that the number and size of conjugated circuits in a molecular graph are responsible for aromaticity (antiaromaticity) of the corresponding conjugated compound.

Of course, all cycles of a molecular graph are not conjugated circuits (e. g. odd circuits are never conjugated). In this respect the following statement summarizes an important property of benzenoid systems.[21,22]

Theorem 3. All conjugated circuits of benzenoid graphs are of size $4m+2$ (m = integer).

Another formulation of this theorem is

$$r_1(L) = r_3(L)) = r_4(L) = 0$$

$$\text{for all} \quad L \in \mathbf{L}$$

It is not possible to deduce the above statement without precisely knowing the conditions the molecule has to fulfill in order to be called »benzenoid«. Therefore, the proof of a result (Theorem 2 of ref 18), which is seemingly equivalent to our Theorem 3 cannot be assumed as satisfactory for the present approach. In the chemical literature there seems to be no generally accepted and mathematically consequent definition of benzenoid systems.[23] Moreover, according to Randić,[21,22] benzenoid graphs can be *defined* as those fulfilling the requirements of Theorem 3. We note, however, that in the great majority of cases there is no doubt whether a conjugated system is benzenoid or not.

3. HARARY'S THEOREM

The following result of Harary[1] gives a general answer to the problem of dependence of D on the structure of a graph G.

Theorem 4. For an arbitrary graph G,

$$D(G) = \sum_{\mathbf{L}} (-1)^{c_{even}(L)} \, 2^{r(L)} \tag{1}$$

where the summation in the above formula goes over the set of all linear graphs of the graph G.

Theorem 4 holds for G being an arbitrary graph including the case of graphs with loops of unit weight. However, if the loops in G have weights $h \neq 1$, Theorem 4 does not apply and the related formula is derived.

Theorem 5. Let the linear graph L contain t loops of weight $h_1, h_2, \ldots h_t$ and let $p(L)$ be the product of $h_1, h_2, \ldots, h_t$.

Then,

$$D(G) = \sum_{\mathbf{L}} (-1)^{c_{even}(L)} \, 2^{r(L)} \, p(L) \tag{2}$$

3.1. *Alternative Formulations of Harary's Theorem*

Theorem 6. For an arbitrary graph G,

$$D(G) = (-1)^n \sum_{\mathbf{L}} (-1)^{c(L)} \, 2^{r(L)} \tag{3}$$

$$D(G) = (-1)^{n/2} \sum_{\mathbf{L}} (-1)^{r_4(L) + \frac{r_3(L) - r_1(L)}{2}} \, 2^{r(L)} \tag{4}$$

The formulation of Harary's theorem in the form (3) was first given by Sachs[24], and later widely applied in chemistry.[13,15,18] From formula (4) the following corollary can be deduced.

If G is a bipartite graph, $r_3 = r_1 = 0$ for all linear graphs of G. Then eq 4 reads

$$D(G) = (-1)^{n/2} \sum_{\mathbf{L}} (-1)^{r_4(L)} \, 2^{r(L)} \tag{5}$$

3.2. *A Recursion Relation for D*

In this section we develop an algorithm for calculation of D based on a recursion relation. Let consider a graph G and the set $\mathbf{L}$ of its linear graphs. Let e be an arbitrary edge of G. Then the following three cases can occur: either a linear graph $L \in \mathbf{L}$ does not contain e, or e is an isolated edge of L, or e belongs to a cycle C_i of L. Accordingly, the set $\mathbf{L}$ can be partitioned as

$$\mathbf{L} = \mathbf{L}(\ddot{e}) \cup \mathbf{L}(e) \cup \bigcup_i \mathbf{L}(C_i)$$

where

- $\mathbf{L}(\ddot{e})$ is the set of those linear graphs of G which do not contain e,
- $\mathbf{L}(e)$ is the set of those linear graphs of G which contain e as an isolated edge,
- $\mathbf{L}(C_i)$ is the set of those linear graphs of G which contain e within the cycle C_i.

Let further $G - e$, $G - (e)$ and $G - C_i$ denote the graphs obtained by deletion from G of the edge e, of the edge e and both vertices incident to it, and of the cycle C_i, respectively. Then from eq 3,

$$D(G) = (-1)^n \sum_{\mathbf{L}(G)} (-1)^{c(L)} \, 2^{r(L)} = (-1)^n \sum_{\mathbf{L}(\ddot{e})} + \sum_{\mathbf{L}(e)} + \sum_{i} \sum_{\mathbf{L}(C_i)} (-1)^{c(L)} \, 2^{r(L)}$$

$$= (-1)^n \left[\sum_{\mathbf{L}(G-e)} (-1)^{c(L)} \, 2^{r(L)} + \sum_{\mathbf{L}(G-(e))} (-1)^{c(L)+1} \, 2^{r(L)} + \right.$$

$$\left. \sum_{i} \sum_{\mathbf{L}(G-C_i)} (-1)^{c(L)+1} \, 2^{r(L)+1} \right]$$

By taking into account that the graphs $G - e$, $G - (e)$ and $G - C_i$ contain n, $n - 2$ and $n - |C_i|$ vertices, respectively, we get finally the recurrence formula

$$D(G) = D(G-e) - D(G-(e)) - 2 \sum_{i} (-1)^{|C_i|} D(G - C_i) \tag{6}$$

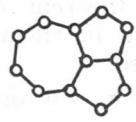
where $|C_i|$ is the size of the cycle C_i and $\sum_i$ denotes the summation over all cycles C_i which contain the edge e.

If C_i is a Hamiltonian cycle of G, $G - C_i$ is the graph $\emptyset$ without vertices. By definition, $D(\emptyset) = 1$.

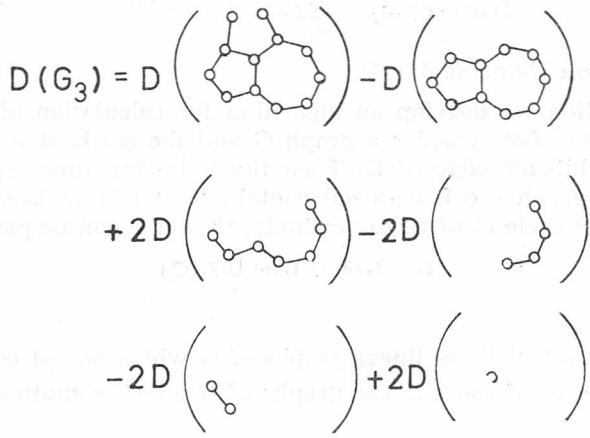
A simple special case of eq 6 is to be noted. Namely, if e is a terminal edge of G (that is an edge incident to a vertex of degree one), $G - e$ has an isolated vertex and therefore $D(G - e) = 0$. Moreover, e belongs to no cycles of G. Therefore, if e is a terminal edge

$$D(G) = -D(G - (e)) \tag{7}$$

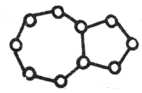
Let us now illustrate the applicability of eqs 6 and 7 on the example of aceazulene graph G_3.



G_3

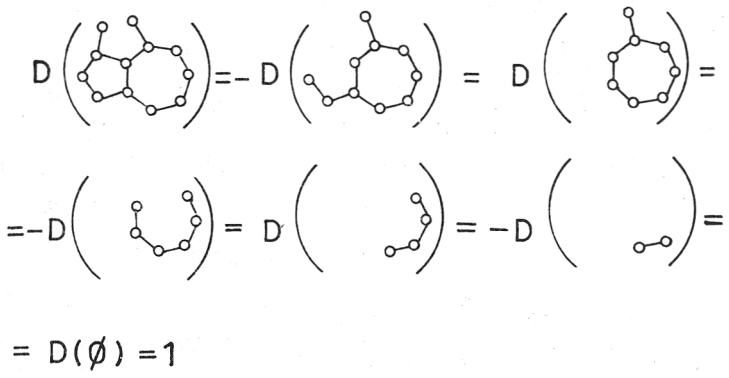


It can be established separately, that the determinant of the azulene graph G_4 is -4. Now we apply eq 7



G_4

several times. For instance,



Finally we have

$$D(G_3) = 1 - (-4) + 2 \cdot 0 - 2 \cdot 1 - 2 \cdot (-1) + 2 \cdot 0 = 5$$

Calculations of the above type are essentially simplified if one takes into account that the determinant of a path with n vertices fulfills the simple relation: $D = (-1)^{n/2}$ if n is even and $D = 0$ if n is odd.

The same way of reasoning which led to formula 6 can be applied also in the case of graphs containing (weighted) loops. Let the edge e be a loop of weight h on the vertex v of the graph $G(h)$. Let further G and $G - v$ be the graphs obtained by deletion of this loop from $G(h)$ (by formally seting $h = 0$) and by deletion of the vertex v from G, respectively. Then by applying Theorem 5, eq 6 becomes

$$D(G(h)) = D(G) - h\,D(G - v) \tag{8}$$

Acknowledgment. Professor M. Randić (Ames, Iowa) read this work in manuscript and gave numerous critical remarks and suggestions, which led to the improvement of our paper. His help, as well as the suggestions given by Professor D. Cvetković (Belgrade) and referees, are gratefully acknowledged.

This work was financially supported in part by the *Self-managing Authority for Scientific Research of SR Croatia* and also supported in part by the *National Science Foundation* (Grant No. F6F 006Y).

REFERENCES

1. F. Harrary, *SIAM Rev.* **4** (1962) 202.
2. See footnote 7 in: R. A. Marcus, *J. Chem. Phys.* **43** (1965) 2643.
3. C. F. Wilcox, *J. Amer. Chem. Soc.* **91** (1969) 2732.
4. H. C. Longuet-Higgins, *J. Chem. Phys.* **18** (1950) 265.
5. M. J. S. Dewar and H. C. Longuet-Higgins, *Proc. Roy. Soc.* **A 214** (1952) 482.
6. C. F. Wilcox, *Tetrahedron Lett.* (1968) 795.
7. D. Cvetković, I. Gutman, and N. Trinajstić, *J. Mol. Struct.* **28** (1975) 289 and references therein.
8. I. Gutman, N. Trinajstić, and C. F. Wilcox, *Tetrahedron* **31** (1975) 143; C. F. Wilcox, *Croat. Chem. Acta* **47** (1975) 87.

9. B. J. McClelland, *J. Chem. Phys.* **54** (1971) 640; I. Gutman, *Chem. Phys. Lett.* **24** (1974) 283.
10. I. Gutman, *Naturwissenschaften* **61** (1974) 216; *J. Chem. Phys.* **66** (1977) 1652.
11. I. Gutman, *Chem. Phys. Let.* **26** (1974) 355; **37** (1976) 475.
12. F. Harary, *Graph Theory*, Addison-Wesley, Reading 1969.
13. I. Gutman and N. Trinajstić, *Topics Curr. Chem.* **42** (1973) 49; *Croat. Chem. Acta* **47** (1975) 507; A. Graovac, I. Gutman, and N. Trinajstić, *Lecture Notes in Chemistry* **4** (1977), Springer, Berlin.
14. *Chemical Applications of Graph Theory* (Ed. A. Balaban), Academic Press, London 1976.
15. A. Graovac, I. Gutman, N. Trinajstić, and T. Živković, *Theor. Chim. Acta* **26** (1972) 67.
16. R. B. Mallion, A. J. Schwenk, and N. Trinajstić, *Croat. Chem. Acta* **46** (1974) 171; R. B. Mallion, N. Trinajstić, and A. J. Schwenk, *Z. Naturforsch.* **29a** (1974) 1481.
17. According to our previous terminology[15,18] linear graphs are in fact spanning Sachs graphs (i. e. Sachs graphs with n vertices).
18. D. Cvetković, I. Gutman, and N. Trinajstić, *J. Chem. Phys.* **61** (1974) 2700.
19. In graph-theoretical terminology[12], a Kekulé graph is called a 1-factor of the graph G.
20. I. Gutman and N. Trinajstić, *Croat. Chem. Acta* **45** (1973) 539.
21. M. Randić, *Chem. Phys. Lett.* **38** (1976) 68.
22. M. Randić, *J. Amer. Chem. Soc.* **99** (1977) 444; *Tetrahedron* **33** (1977) 1905.
23. The authors thank Professor M. Randić for helpful discussions on this problem.
24. H. Sachs, *Publ. Math.* (Debrecen) **11** (1964) 119.

SAŽETAK

Determinanta matrice susjedstva za graf konjugirane molekule

A. Graovac i I. Gutman

Diskutirane su različite graf-teorijske tehnike izračunavanja determinante matrice susjedstva molekulskog grafa.

INSTITUT »RUĐER BOŠKOVIĆ«,
P.O.B. 1016, 41001 ZAGREB

Prispjelo 28. travnja 1977.