Theoret. chim. Acta (Berl.) 26, 67—78 (1972)

Graph Theory and Molecular Orbitals

Application of Sachs Theorem

A. Graovac, I. Gutman, N. Trinajstić, and T. Živković
Institute "Rudjer Bošković", P.O.B. 1016, 41001 Zagreb, Croatia, Yugoslavia

Received March 1, 1972

The Sachs theorem [1] is discussed, and Kekulé and Sachs graphs are defined. The simple consequences of Sachs theorem which are of interest in chemistry are presented. Thus, it is shown that the pairing theorem, Hückel $(4m+2)$ rule, and Longuet-Higgins and Dewar rule [2] can be obtained easily with the use of Sachs theorem. The dependence of the number of Kekulé structures on the molecular topology is also shown.

Das Theorem von Sachs [1] wird diskutiert und Kekulé- und Sachs-Graphen werden definiert. Die einfachen Folgerungen aus dem Theorem von Sachs, die in der Chemie von Interesse sind, werden abgeleitet. So wird gezeigt, daß die Paarregel von Coulson und Rushbrooke, die Hückelsche $(4m+2)$-Regel und die Regel von Longuet-Higgins und Dewar [2] leicht mit dem Sachsschen Theorem erhalten werden können. Die Abhängigkeit der Zahl der Kekulé-Strukturen von der molekularen Topologie wird ebenfalls gezeigt.

Introduction

In this work our intention is to investigate the applicability of the mathematical apparatus of Graph theory (GT) to the simple molecular orbital theory. Without insisting on rigid mathematical formalism (which can be found elsewhere [3–5]), we define a graph as an ordered pair

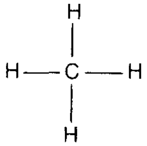
$$(\mathcal{N}, \mathcal{A}) \tag{1}$$

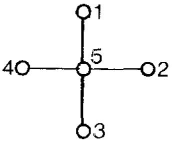
where $\mathcal{N}$ is a set of some elements ("nodes") and $\mathcal{A}$ is a relation defined on the set $\mathcal{N}$, which is symmetrical and antireflexive, i.e.:

$$(X, Y) \in \mathcal{A} \Rightarrow (Y, X) \in \mathcal{A}\,, \tag{2a}$$

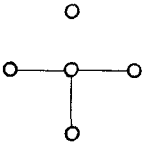
$$(X, Y) \in \mathcal{A} \Rightarrow X \neq Y. \tag{2b}$$

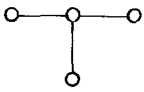
So, two nodes can either belong to the relation $\mathcal{A}$ (than we say that this nodes are adjacent), or not (then this nodes are not adjacent). A usual sketch representation of a graph is obtained when the nodes (elements of the set $\mathcal{N}$) are drawn as small circles and if two nodes belong to the relation $\mathcal{A}$ they are connected by a line (so called "edge"). The analogy between the nodes and edges of a graph and the atoms and bonds in a structural formula of a molecule is obvious.

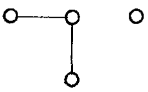




G


G'


G''

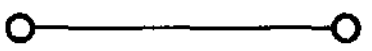

G'''

For example, the graph G which is a representation of the structural formula of methane has five nodes. The relation $\mathscr{A}$ is then the following set of pairs:

$$\{(1,5), (5,1), (2,5), (5,2), (3,5), (5,3), (4,5), (5,4)\}.$$

If some edges of a graph are removed a *subgraph* of the graph is obtained. If some nodes together with all adjacent edges are removed a *partial graph* is obtained. A subgraph of a partial graph is called a *partial subgraph* of the graph. G' is a subgraph, G'' is a partial graph and G''' is a partial subgraph of the graph G.

The graph with two nodes and one edge:



will be denoted as Γ.

Using GT, different combinatorial problems of the Organic chemistry can be solved: The number of structural isomers [6], the number of conformations [7], the number of Kekulé structures [8]. However, using a graph theory representation of a molecule, all geometric properties (bond lengths, angles) are neglected. The only molecular property considered is the existence or non-existence, respectively, of the chemical bond between two atoms. This molecular property is usually called "topology" [9] and, although the expression is not the most precise one, it is in common use.

Relation between the Graph Theory and Hückel Theory

There is some scepticism towards the attempts to obtain any useful conclusion about chemical and physical properties of compounds only from the knowledge of their topology [10]. However, in the last few years, various modifications of the HMO method enabled to predict chemical behaviour of conjugated molecules as well as by using more complicated SCF calculations [11]. Similarly, the Hückel formalism using additional elements of GT was successfully applied to different classes of inorganic compounds [12].

The relation between GT and HMO theory can be presented in the following manner.

After a numeration of the nodes is performed, there is one to one correspondence between a graph and so called "adjacency matrix" A, defined as:

$$A_{pq} = \begin{cases} 1 \text{ if the nodes } p \text{ and } q \text{ are adjacent} \\ 0 \text{ if they are not.} \end{cases} \tag{3}$$

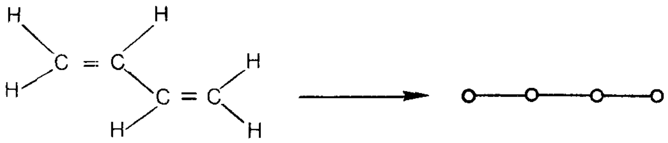
The Hamiltonian matrix in the Hückel approximation can be now written as:

$$\boldsymbol{H} = \alpha \boldsymbol{I} + \beta A \tag{4}$$

where $\boldsymbol{I}$ is a unit matrix; α and β are the Coulomb and the resonance integrals.

When conjugated hydrocarbons are examined in the HMO approach, the related graphs correspond to the carbon-carbon σ-bond skeleton, while the

π-bonds and the C–H σ-bonds are neglected:



In this work all graphs are assumed to be of the above type.

The secular equation:

$$\det|\boldsymbol{H}-\varepsilon\boldsymbol{S}|=0 \tag{5}$$

can be written, after adopting the zero-overlap approximation [13], as:

$$\det|x\boldsymbol{I}-\boldsymbol{A}|=0 \tag{6}$$

where:

$$x=\frac{\varepsilon-\alpha}{\beta}. \tag{7}$$

In other words, the problem of the Hückel orbital energies can be completely reduced to the adjacency matrix eigenvalue problem. Albeit this fact should have been known, at least implicitly, to everyone who once solved a Hückel problem, its fundamental meaning was discovered relatively recent [4, 5, 12c, 14].

Eq. (6) was exhaustively studied within GT [1, 15, 16]. Many useful results were obtained. Some of them we shall comment here.

In the GT the following nomenclature is used. $\det|x\boldsymbol{I}-\boldsymbol{A}|\equiv P_G(x)$ is the "characteristic polynomial". The set of its roots:

$$\{x_1, x_2, \ldots, x_N\}$$

being called the "spectrum" of the corresponding graph.

So, let

$$P_G(x)=\sum_{n=0}^{N} a_n x^{N-n} \tag{8}$$

where N is the number of nodes in the graph. Since the characteristic polynomial $P_G(x)$ is uniquely defined by the graph G, the coefficients a_n can be found without going through the procedure of solving the determinant, but solely knowing the topological structure of the graph.

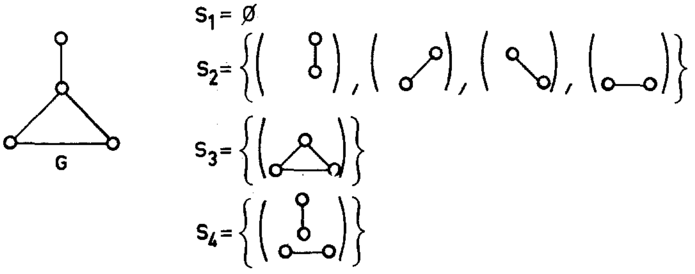
If we define a Sachs graph as such graph which has no other components but graphs Γ and rings, a following formula, originally given by Sachs [1], determins those coefficients:

$$\begin{aligned} a_n &= \sum_{s\in S_n} (-)^{c(s)} 2^{r(s)} \quad \text{for} \quad 0<n\leqq N \\ a_0 &= 1\,. \end{aligned} \tag{9}$$

Here s is a Sachs graph and S_n is the set of all Sachs graphs with n nodes. $c(s)$ and $r(s)$ denote respectively the number of components and the number of ring components in s. The summation in (9) is over all elements of the set S_n. If there is $S_n=\emptyset$, where $\emptyset$ is an empty set, then $a_n=0$. We notice although that for odd n every $s\in S_n$ must contain at least one ring with odd number of nodes.

Let us illustrate Eq. (9) by two examples:

I



s are drawn in small brackets. Now it can be seen that:

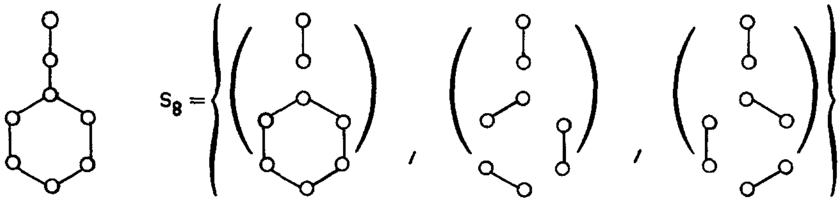
$$a_1 = 0$$
$$a_2 = (-)^1 2^0 + (-)^1 2^0 + (-)^1 2^0 + (-)^1 2^0 = -4$$
$$a_3 = (-)^1 2^1 = -2$$
$$a_4 = (-)^2 2^0 = +1$$

and hence:

$$P_G(x) = x^4 - 4x^2 - 2x + 1 .$$

The use of Eq. (9) often requires a same amount of work as a direct solving of the determinant. But, some coefficients, which prove to be important for different reasons (see discussion later on), can be obtained in a rather simple way:

II



$$a_8 = (-)^2 2^1 + (-)^4 2^0 + (-)^4 2^0 = 4 .$$

Application of Sachs Formula to Conjugated Hydrocarbons

We will give now a survey of some simple consequences of Sachs formula which are of particular interest in chemistry.

A) According to the definition of S_n there is always:

$$S_1 = \emptyset , \tag{10a}$$

i.e.

$$a_1 = 0 . \tag{10b}$$

Since

$$a_1 = \sum_{j=1}^{N} x_j$$

it follows that:

$$\sum_{j=1}^{N} x_j = 0 . \tag{11}$$

The total π-electron energy of the molecule in the singlet ground state is:

$$E_\pi = 2\beta \sum_{j=1}^{N/2} x_j \tag{12}$$

where x_j are the roots given in decreasing order.

B) It is simple to see that:

$$-a_2 = \text{number of edges.} \tag{13}$$

As a consequence of the Eq. (11), the following relation holds:

$$a_2 = -\frac{1}{2} \sum_{j=1}^{N} x_j^2 . \tag{14}$$

This can be proved in a simple manner:

$$a_2 = \sum_{i<j} x_i x_j = \frac{1}{2}\left(\sum_{i=1}^{N} \sum_{j=1}^{N} x_i x_j - \sum_{j=1}^{N} x_j^2 \right)$$

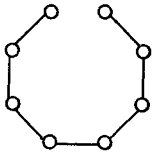
$$= -\frac{1}{2} \sum_{j=1}^{N} x_j^2 + \frac{1}{2} \sum_{i=1}^{N} x_i \sum_{j=1}^{N} x_j = -\frac{1}{2} \sum_{j=1}^{N} x_j^2 ,$$

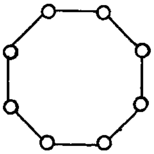
the first step being known from algebra.

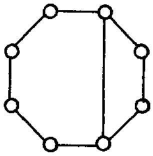
Eqs. (10), (11), (13) and (14) are already obtained by Günthard and Primas [5] using the theorem of Collatz and Sinogowitz [15], which is equivalent to Sachs theorem.

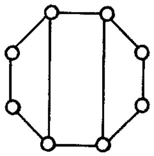
Since a_2 can be obtained from the graph it would be very useful if some correlation between E_π and a_2 could be established. Fortunately, such a correlation exists [17]. Almost in all cases when the number of edges is increased (the number of nodes being the same) the total π-electron energy increases too. In chemistry it corresponds to ring closure [18].

Example:









$-a_2$	7	8	9	10
E_π	9.52	9.66	10.38	10.47

C) We will now consider the important class of alternant hydrocarbons. The graphs which correspond to these molecules have the property that their nodes can be painted with two colors in such a manner that two adjacent nodes always have different color. Hence, they are called "bichromatic graphs". It can be proved [3] that a graph is bichromatic if and only if it does not have any odd membered ring as its partial subgraph. Hence

$$S_n = \emptyset \tag{15a}$$

and

$$a_n = 0 \tag{15b}$$

for all odd n. So, the characteristic polynomial of the graph is of the form:

$$P_G(x) = x^N + a_2 x^{N-2} + a_4 x^{N-4} + \cdots \tag{16}$$

and the spectre of the graph is symmetrically arranged with respect to $x = 0$.

It is interesting to note that this fact (the pairing theorem) was first proved by Coulson and Rushbrooke [19] long before and independent of the GT [20].

D) The even alternant hydrocarbons[1], as a consequence of the pairing theorem, have triplet ground state if there is at least one zero in the spectre of the corresponding graph [2a]. In other words, such molecules are expected to be unstable within the framework of the HMO theory. This important chemical implication of the HMO theory has therefore a topological background.

Since

$$a_N = x_1 x_2 \ldots x_N \equiv \prod_{j=1}^{N} x_j \tag{17a}$$

and, hence, for alternant hydrocarbons

$$a_N = (-)^{N/2} \prod_{j=1}^{N/2} x_j \tag{17b}$$

it is sufficient to investigate if

$$a_N = 0\,. \tag{18}$$

If Eq. (18) holds, there is at least one zero in the spectrum of $P_G(x)$ and, thus low chemical stability is to be expected. Further discussion about this problem will be reported in the subsequent publication [21].

a) One can show that Eq. (18) holds if there is no Kekulé structure for the given hydrocarbon.

First, if there is at least one Sachs graph $s \in S_N$ corresponding to the given molecule, it implies the existence of Sachs graphs without rings (see below).

The Sachs graphs without rings and with N nodes are in an obviously manner analogous to Kekulé structures of the molecule. We will call them "Kekulé graphs", and denote by

$$k_1, k_2, \ldots, k_K\,.$$

The set of all K Kekulé graphs will be denoted as $\mathscr{K}$. It is clear that $\mathscr{K} \subseteqq S_N$.

---

[1] We restrict our considerations only to the case of even number of π centers and electrons.

Therefore, the nonexistence of Kekulé graphs implicates the nonexistence of any Sachs graph with N nodes, and, hence, according to (9) there is

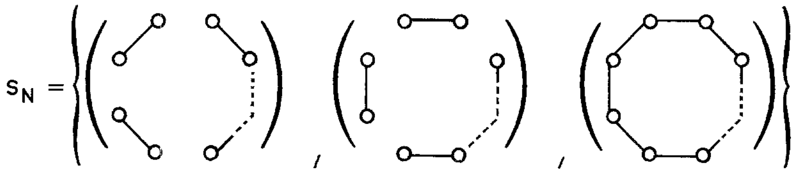
$$S_N = \emptyset\,.$$

b) Conversely, if there exists only one Kekulé structure (i.e. only one Kekulé graph and, hence, only one $s \in S_N$ as, for example, in the case of noncyclic molecules), then

$$a_N = (-)^{N/2} \tag{19}$$

and, of course, there is no zero in the spectre.

E) For rings (annulenes) the Hückel $(4m+2)$rule can be proved by considering:



$$a_N = (-)^{N/2}\, 2^0 + (-)^{N/2}\, 2^0 + (-)^1\, 2^1 = \begin{cases} 0 \text{ for } N = 4m \\ -4 \text{ for } N = 4m+2\,. \end{cases}$$

It means that only the $(4m+2)$-annulenes have a singlet ground state (in HMO theory).

Recently, Goldstein and Hoffmann [22] presented an other approach to Hückel $(4m+2)$rule.

F) For polycyclic alternant hydrocarbons the rule of determining the value of a_N (and thus molecular stability) is given by Longuet-Higgins and Dewar [2]. This rule is interesting because it express a relationship between MO and VB theory (see also Ref. [23]).

Before the proof of this rule is given, we will point out an important special case.

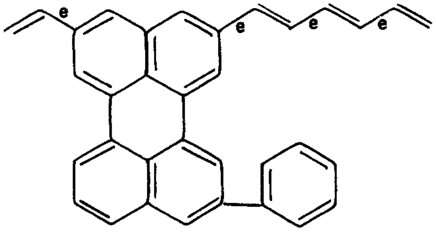
a) Let the graph G possess an edge $\underset{\sim}{e}$ such that by removing this edge the graph separates into two components G_1 and G_2, and let the number of nodes in these graphs: N, N_1 and N_2 be even. Than it is valid:

$$a_N(G) = a_{N_1}(G_1)\, a_{N_2}(G_2)\,. \tag{20}$$

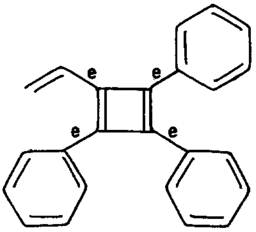
This is the consequence of the fact that no Sachs graph with N nodes possess the edge $\underset{\sim}{e}$.

The meaning of Eq. (20) is that by joining of two stable molecules (by one bond) we always obtain a stable molecule, and if, at least, one of the joined molecules is unstable (but have even number of atoms) the resulting is unstable too.

Example:



stable
(singlet ground state)



unstable
(triplet ground state)

It is of special chemical importance that the introduction of vinyl-, phenyl- and similar groups at arbitrary positions of molecule cannot cause the essential stabilisation of unstable molecules and the destabilisation of stable molecules.

b) Let the conjugated hydrocarbon have K Kekulé structures. A definite permutation can be assigned to every Kekulé structure (for details see [2]). The formula derived by Longuet-Higgins and Dewar is:

$$a_N = (-)^{N/2} \left(\sum_{j=1}^{K} p_j \right)^2 \tag{21}$$

where

$$p_j = \begin{cases} +1 \text{ if the permutation is even,} \\ -1 \text{ if the permutation is odd.} \end{cases} \tag{22}$$

Specially, if the parity of all permutations is the same (it occurs to molecules without $(4m)$-membered rings), then it is:

$$a_N = (-)^{N/2} K^2 . \tag{23}$$

In order to derive Eq. (21) from the Sachs formula (9) we will define the "summation" of two graphs with the same set of nodes as:

$$(\mathcal{N}, \mathcal{A}_1) \oplus (\mathcal{N}, \mathcal{A}_2) = (\mathcal{N}, \mathcal{A}_1 \cup \mathcal{A}_2) . \tag{24}$$

It can be easily seen that:

$$G_1 \oplus G_1 = G_1 , \tag{25}$$

$$G_1 \oplus G_2 = G_2 \oplus G_1 . \tag{26}$$

Then, the following theorem about the summation of Kekulé graphs is valid:

$$\mathcal{K} \oplus \mathcal{K} = S_N \tag{27}$$

where $\mathcal{K} \oplus \mathcal{K}$ is the set of all $k_A \oplus k_B$.

First notice that every even membered ring (only such rings can occur in Sachs graphs of bichromatic graphs) can be obtained as:

$$\gamma_1 \oplus \gamma_2 \quad \text{and} \quad \gamma_2 \oplus \gamma_1$$

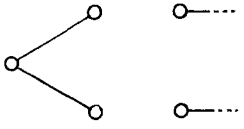
where γ_1 and γ_2 are the Kekulé graphs of the ring. Naturally these are the only two possibilities for obtaining this ring by summation of Kekulé garphs. Hence, for every Sachs graph $s \in S_N$ there exist at least one pair of Kekulé graphs k_A, k_B such that

$$k_A \oplus k_B = s\,.$$

Thus

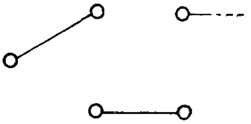
$$S_N \subseteqq \mathscr{K} \oplus \mathscr{K}\,. \tag{28}$$

Let us suppose that the summation of k_A and k_B gives a graph σ such that $\sigma \notin S_N$. For example, such a graph may be:



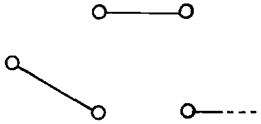
σ

The only way to obtain σ from Kekulé graphs is to sum:



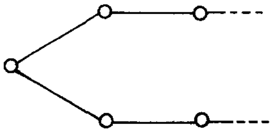
k_A

and



k_B

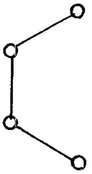
However, $k_A \oplus k_B$ gives not σ but σ':



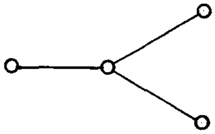
σ'

If $\sigma' \notin S_N$, we can repeat our considerations again.

It can be shown by a completely analogous argumentation that graphs:



and



cannot be obtained by summation of Kekulé graphs.

The above mentioned three cases cover all possible non-Sachs graphs. So we have shown that it is always

$$k_A \oplus k_B \in S_N\,, \quad \text{i.e.}$$

$$\mathscr{K} \oplus \mathscr{K} \subseteqq S_N \tag{29}$$

which completes the proof of Eq. (27).

An important corollary of the theorem (27) is that if a Sachs graph $s \in S_N$ contains r ring components, then there exist exactly 2^r different ordered pairs (k_A, k_B) with the property:

$$k_A \oplus k_B = s\,.$$

Since the set $\mathscr{K} \oplus \mathscr{K}$ contains K^2 elements it is therefore valid:

$$\sum_{s \in S_N} 2^{r(s)} = K^2\,. \tag{30}$$

This equation shows explicitly the dependence of the number of Kekulé structures on the molecular topology (see also Ref. [8]).

For bichromatic graphs Sachs formula (9) can be written in the form:

$$a_N = (-)^{N/2} \sum_{s \in S_N} (-)^{r'(s)} 2^{r(s)} \tag{31}$$

where $r'(s)$ is the number of $(4m)$-membered rings in the Sachs graph s, because the generating of a $(4m+2)$-membered ring from Γ graphs does not alter the parity of the number of components, while a $(4m)$-membered ring does.

By comparison of Eqs. (30) and (31) it can be seen that

$$a_N = (-)^{N/2} \sum_{A=1}^{K} \sum_{B=1}^{K} (-)^{r'(k_A \oplus k_B)}\,, \tag{32}$$

because $k_A \oplus k_B$ is an element of S_N, and there are exactly 2^r pairs which give by summation the same Sachs graph.

Now, it is not difficult to see that:

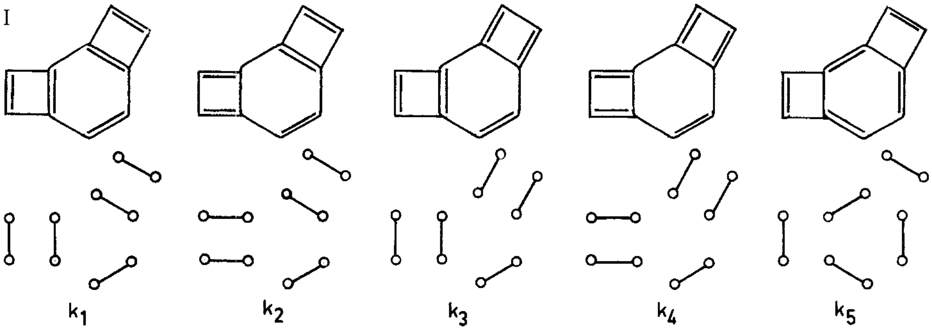
$$(-)^{r'(k_A \oplus k_B)} = p_A p_B \tag{33}$$

what leads straightforwardly to the Longuet-Higgins-Dewar formula (21).

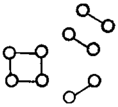
If the permutations of two Kekulé structures have opposite parity, then the sum of the corresponding two Kekulé graphs contains odd number of $(4m)$-membered rings. This enables us to estimate the parity of Longuet-Higgins-Dewar's permutations in a proper and simple manner.

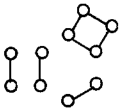
We will use as illustration the same two examples which were considered by Wilcox [2b].

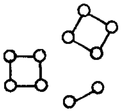
I

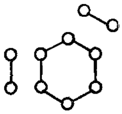


$p_1 \overset{\text{def}}{=} +1$

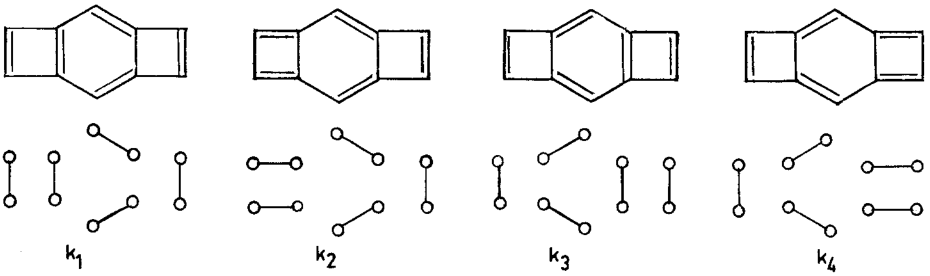
$k_1 \oplus k_2 =$  $\Longrightarrow p_2 = -1$

$k_1 \oplus k_3 =$  $\Longrightarrow p_3 = -1$

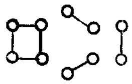
$k_1 \oplus k_4 =$  $\Longrightarrow p_4 = +1$

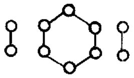
$k_1 \oplus k_5 =$  $\Longrightarrow p_5 = +1$

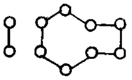
II



$p_1 \overset{\text{def}}{=} +1$

$k_1 \oplus k_2 =$  $\Longrightarrow p_2 = -1$

$k_1 \oplus k_3 =$  $\Longrightarrow p_3 = +1$

$k_1 \oplus k_4 =$  $\Longrightarrow p_4 = -1$

So, $a_{10}=(-)^5(1-1-1+1+1)^2=-1\neq 0$, and singlet ground state is to be expected.

So, $a_{10}=(-)^5(1-1+1-1)^2=0$ and triplet ground state, i.e. very low chemical stability or even nonexistence is to be expected.

References

1. Sachs, H.: Publ. Math. (Debrecen) **9**, 270 (1962).

2a. Longuet-Higgins, H. C.: J. chem. Physics **18**, 265 (1950); Dewar, M. J. S., Longuet-Higgins, H. C.: Proc. Royal Soc. (London) A **214**, 482 (1952).

2b. Wilcox, C. F.: Tetrahedron Letters 795 (1968).

3. Berge, C.: Théorie des graphes et ses applications. Paris: Dunod 1958.

4. Schmidtke, H. H.: J. chem. Physics **45**, 3920 (1966).

5. Günthard, H. H., Primas, H.: Helv. chim. Acta **39**, 1645 (1956).

6. Polya, G.: Acta Math. **68**, 145 (1937). — Harary, F.: Uspekhi mat. Nauk (USSR) **24**, 179 (1969). — Ruch, E., Hásselbarth, W., Richter, B.: Theoret. chim. Acta (Berl.) **19**, 288 (1970).

7. Balaban, A. T.: Tetrahedron **27**, 6115 (1971) and references therein.

8. Gordon, M., Davison, W. H. T.: J. chem. Physics **20**, 428 (1952). — Teh Fu Yen: Theoret. chim. Acta (Berl.) **20**, 399 (1971).

9. Ruedenberg, K.: J. chem. Physics **34**, 1861 (1961).

10. Polansky, O. E.: Private communication.

11. Figeys, H. P.: Tetrahedron **26**, 4615, 5225 (1970). — Kruszewski, J., Krygowski, T. M.: Tetrahedron Letters 319 (1970). — Hess, B. A., Jr., Schaad, L. J.: J. Amer. chem. Soc. **93**, 305, 2413 (1971); J. org. Chem. **36**, 3418 (1971). — Milun, M., Sobotka, Ž., Trinajstić, N.: J. org. Chem. **37**, 139 (1972).

12a. Kettle, S. F. A.: Theoret. chim. Acta (Berl.) **3**, 211 (1965); **4**, 150 (1966). — Kettle, S. F. A., Tomlinson, V.: Theoret. chim. Acta (Berl.) **14**, 175 (1969).

12b. Schmidtke, H. H.: Coord. chem. Rev. **2**, 3 (1967); Int. J. quant. Chemistry **2s**, 101 (1968); J. chem. Physics **48**, 970 (1968).

12c. Schmidtke, H. H.: Theoret. chim. Acta (Berl.) **9**, 199 (1968).

13. *e.g.* Jug, K.: Theoret. chim. Acta (Berl.) **14**, 91 (1969).

14. Ruedenberg, K.: J. chem. Physics **34**, 1884 (1961).

15. Collatz, L., Sinogowitz, U.: Abh. math. Sem. Univ. Hamburg **21**, 63 (1957).

16. Cvetković, D.: Thesis, Beograd 1971.

17. McClelland, B. J.: J. chem. Physics **54**, 640 (1971). — Hakala, R. W.: Int. J. quant. Chemistry **1s**, 187 (1967).

18. Dewar, M. J. S.: Tetrahedron Suppl. **8**, 75 (1966). — Dewar, M. J. S., Pettit, R.: J. chem. Soc. 1617 (1954). — Peters, J.: J. chem. Soc. 1023, 1028 (1958). — Gutman, I., Milun, M., Trinajstić, N.: Croat. chem. Acta **44** (1972) in press. — Gutman, I., Trinajstić, N., Živković, T.: Chem. Physics Letters (in press).

19. Coulson, C. A., Rushbrooke, G. S.: Proc. Cambridge philos. Soc. **36**, 196 (1940).

20. Cvetković, D.: Matematička Biblioteka (Beograd) **41**, 193 (1969).

21. Cvetković, D., Gutman, I., Trinajstić, N.: Croat. chem. Acta submitted for publication.

22. Goldstein, M. J., Hoffmann, R.: J. Amer. chem. Soc. **93**, 6193 (1971).

23. Ham, N. S.: J. chem. Physics **29**, 1229 (1958). — Heilbronner, E.: Helv. chim. Acta **45**, 1722 (1962).

Dr. N. Trinajstić
Institute "Rudjer Bošković"
P.O.B. 1016
41001 Zagreb
Croatia, Yugoslavia