

TOPOLOGICAL PROPERTIES OF BENZENOID SYSTEMS. XLVIII.  
TWO CONTRADICTORY FORMULAS FOR TOTAL  $\pi$ -ELECTRON ENERGY AND  
THEIR RECONCILIATION

Ivan Gutman

Faculty of Science, University of Kragujevac, P.O. Box 60,  
YU-34000 Kragujevac, Yugoslavia

(Received: November 1986)

ABSTRACT

Hall has established that the total  $\pi$ -electron energy ( $E$ ) of benzenoid hydrocarbons is a linear function of the number of Kekulé structures ( $K$ ). Cioslowski recently came to the conclusion that  $E$  is a linear function of  $K^{2/n}$  where  $n$  denotes the number of carbon atoms. These two apparently contradictory statements can be reconciled, taking into account that both Hall's and Cioslowski's formula are approximations which only correlate with the exact  $E$  values.

## INTRODUCTION

The present paper is concerned with the total  $\pi$ -electron energy of benzenoid hydrocarbons, as calculated within the Hückel molecular orbital model and expressed in the units of  $\beta$  [1]. This quantity will be denoted by  $E$ .

The dependence of  $E$  on the structure of the molecule has attracted a considerable attention in the past few decades [1]. As a result of numerous investigations it has been established that the gross part of  $E$  is determined by only two topological invariants, namely  $n$  - the number of vertices and  $m$  - the number of edges of the molecular graph [1]. In the case of benzenoid systems "gross part" means something about 99.5 %.

The third topological invariant which is believed to significantly influence the value of  $E$  is  $K$  - the number of Kekulé structures. A variety of approximate formulas for  $E$ , depending on  $n$ ,  $m$  and  $K$  has been proposed in the literature [2-11]. In 1973 Hall observed [12] that when for a group of isomers  $E$  is plotted versus  $K$  an almost perfect straight line is obtained. This eventually resulted in the formula [6]:

$$E = An + Bm + CKD^{m-n} \quad (1)$$

where  $A$ ,  $B$ ,  $C$  and  $D$  are empirical constants determined by least squares fitting\*. The remarkable feature of formula (1) is that

---

\*Hall's original values were [6]  $A = 0.422$ ,  $B = 0.788$ ,  $C = 0.34$ ,  $D = 0.632$ . Cioslowski found [10]  $A = 0.5477$ ,  $B = 0.7027$ ,  $C = 0.2393$ ,  $D = 0.6658$ , whereas for catacondensed hydrocarbons the best values of these parameters were reported [7] to be  $A = 0.4044$ ,  $B = 0.8198$ ,  $C = 0.3206$ ,  $D = 0.6482$ . Hence, the parameters in (1) are to some extent sample-dependent.

its dependence on  $K$  is linear. Somewhat later Hall's formula was independently discovered in [7] where also a theoretical explanation of its validity was offered.

Recently Cioslowski [10] developed an approach by which he was able to show that approximate formulas for  $E$ , depending only on  $n$ ,  $m$  and  $K$  must be of the form

$$E = (2mn)^{1/2} F[K^{2/n}(2m/n)^{-1/2}] \quad (2)$$

where  $F$  is some function. In a later work [11] the condition  $F(1) = 1$  was imposed.

Cioslowski's result (2) is deduced under the assumption that the eigenvalues of the molecular graphs of benzenoid hydrocarbons obey a universal distribution law. This latter is, in the best case, a plausible approximation and therefore (2) must not be understood as an exact result.

The simplest non-trivial approximate expression based on (2) is [10]

$$E = A(2mn)^{1/2} + BnK^{2/n} \quad (3)$$

where  $A = 0.7676$  and  $B = 0.1775$  are obtained by least squares fitting.

The formulas (1) and (3) have similar accuracy and reliability. For a sample of 1030 singlet ground state benzenoids the average error of (1) and (3) is 0.06 % and 0.07 %, respectively. The maximal observed error is 0.39 % and 0.63 %, respectively. Hence Hall's formula (1) is found to be somewhat better [10].

In [10,11] a further argument in favour of (2) and (3) was

raised, namely their size-consistency. Really, (2) and (3) have the property

$$E(un,um,K^u) = uE(n,m,K) \quad (4)$$

i.e. the energy of a system composed of u identical benzenoid molecules is calculated to be u times the energy of a single such molecule. Hall's formula does not obey the requirement (4).

The true size-consistency property would be

$$E\left(\sum_{i=1}^u n_i, \sum_{i=1}^u m_i, \prod_{i=1}^u K_i\right) = \sum_{i=1}^u E(n_i, m_i, K_i) \quad (5)$$

which means that the energy of a system composed of u (not necessarily identical) benzenoid molecules is equal to the sum of their energies. This condition is, however, not fulfilled by either (1) or (2) or (3). The only expression satisfying (5) is

$$E = An + Bm + C \ln K$$

proposed by Carter as early as in 1949 [2]. Unfortunately, this latter approximation for E was found to be inferior compared to (1) and (3) [8,10].

The most striking difference between Hall's and Cioslowski's formula is their dependence on K. In the subsequent sections we show that, from a practical point of view, the expressions (1) and (3) need not be considered as contradicting to each other.

CORRELATION BETWEEN K AND  $K^{2/n}$

In order to gain some information about the correlation between K and  $K^{2/n}$ , we shall consider the function  $y = x^{1/p}$  and assume that x belongs to a certain interval [a,b]. Suppose that we select uniformly by random N numbers  $x_1, x_2, \dots, x_N$  from the interval [a,b] and investigate whether there exists a linear correlation between the pairs  $(y_i, x_i)$ ,  $i = 1, 2, \dots, N$  where  $y_i = x_i^{1/p}$ . The measure of such a linear correlation will be the correlation coefficient

$$r = (\langle xy \rangle - \langle x \rangle \langle y \rangle) [(\langle x^2 \rangle - \langle x \rangle^2) (\langle y^2 \rangle - \langle y \rangle^2)]^{-1/2}$$

where  $\langle f \rangle$  denotes the mean value of the quantity f. If N is sufficiently large, one may take

$$\langle f \rangle = (b-a)^{-1} \int_a^b f \, dx \quad .$$

Then we immediatly get

$$\begin{aligned} \langle x \rangle &= (a+b)/2 \\ \langle y \rangle &= (1+1/p)^{-1} (b^{1+1/p} - a^{1+1/p}) (b-a)^{-1} \\ \langle x^2 \rangle &= (a^2 + ab + b^2)/3 \\ \langle y^2 \rangle &= (1+2/p)^{-1} (b^{1+2/p} - a^{1+2/p}) (b-a)^{-1} \\ \langle xy \rangle &= (2+1/p)^{-1} (b^{2+1/p} - a^{2+1/p}) (b-a)^{-1} \end{aligned}$$

which after some algebraic manipulations lead to the following results.

If  $a = 0$  and b has an arbitrary positive value, then

$$r = [(3p^2 + 6p) / (4p^2 + 4p + 1)]^{1/2} \tag{6}$$

and thus  $r$  is independent of  $b$ .

If  $a \neq 0$  and  $t = b/a$ , then

$$\begin{aligned}
 r &= (3p)^{1/2} (t-1)^{-1} [(t^{2+1/p-1}) (2p+1)^{-1} - \\
 &\quad - t(t^{1/p-1})] [(t-1) (t^{1+2/p-1}) (p+2)^{-1} - \\
 &\quad - pt(t^{1/p-1})^2]^{-1/2}
 \end{aligned} \tag{7}$$

and thus  $r$  depends only on the ratio of  $b$  and  $a$ .

A detailed examination of formula (6) shows that it monotonously decreases with increasing  $p$ ,  $p \geq 1$ . Its maximal value is for  $p = 1$  when, of course,  $r = 1$ . Its minimal value is  $(3/4)^{1/2}$  which is its limit when  $p$  tends to infinity.

The analysis of eq. (7) is somewhat more difficult. It can be shown that

$$\lim_{t \rightarrow 1} r = (1+p) (9p^2 - 10p + 5)^{-1/2} ,$$

$$\lim_{p \rightarrow \infty} r = (3/4)^{1/2} (t^2 - 2t \ln t - 1) (t-1)^{-1} [(t-1)^2 - t(\ln t)^2]^{-1/2} ,$$

$$\lim_{t \rightarrow 1} (\lim_{p \rightarrow \infty} r) = 1 .$$

In addition to this, if  $p = 1$ , then  $r = 1$  for all values of  $t$ . For any fixed value of  $t$ ,  $r$  monotonically decreases with increasing  $p$ ,  $p \geq 1$ .

DISCUSSION

In Table 1 we present as a realistic example the correlation coefficients corresponding to the case when the sample examined contains non-branched catacondensed benzenoid hydrocarbons with  $h$  hexagons. Then  $a = K_{\min} = h+1$ ,  $b = K_{\max} =$  the  $(h+1)$ -th Fibonacci number and  $p = 2h+1$ .

Table 1

$h$	$r$ , eq. (7)
3	0.9997
4	0.9986
5	0.9960
6	0.9918
7	0.9860
8	0.9786
10	0.9604
12	0.9405
15	0.9137
20	0.8866
$\infty$	$(3/4)^{1/2} = 0.8660$

It can be seen that for benzenoid systems which are usually considered in chemistry (and, in particular, for the sample used by Cioslowski [10] for which  $h \leq 8$ ), the correlation between  $K$  and  $K^{2/n}$  has a correlation coefficient as high as 0.98 - 0.99. This means that if the true  $E$  values are well correlated with the r.h.s. of eq. (1), then they will be well correlated also with the r.h.s. of eq. (3) and vice versa. Therefore both (1) and (3) may be of similar numerical quality and from the point of view of data fitting they need not contradict to each other.

REFERENCES

- [1] For a detailed exposition of the theory of total  $\pi$ -electron energy see: I. Gutman and O.E. Polansky, *Mathematical Concepts in Organic Chemistry*, Springer-Verlag, Berlin 1986, pp. 135-154.
- [2] P.C. Carter, *Trans. Faraday Soc.* 45, 597 (1949).
- [3] I. Gutman, N. Trinajstić and C.F. Wilcox, *Tetrahedron* 31, 143 (1975).
- [4] I. Gutman and N. Trinajstić, *Croat. Chem. Acta* 48, 297 (1976).
- [5] I. Gutman, *J. Chem. Phys.* 66, 1652 (1977).
- [6] G.G. Hall, *Publ. Math. Inst. Appl.* 17, 70 (1981).
- [7] I. Gutman and S. Petrović, *Chem. Phys. Letters* 97, 292 (1983).
- [8] I. Gutman, Lj. Nedeljković and A.V. Teodorović, *Bull. Soc. Chim. Beograd* 48, 495 (1983).
- [9] I. Gutman, A.V. Teodorović and Lj. Nedeljković, *Theoret. Chim. Acta* 65, 23 (1984).
- [10] J. Cioslowski, *Match* 20, 95 (1986).
- [11] J. Cioslowski and O.E. Polansky, to be published.
- [12] G.G. Hall, *Internat. J. Math. Educ. Sci. Technol.* 4, 233 (1973).