

**A FINAL SOLUTION OF THE PROBLEM CONCERNING THE (N,M,K)-DEPENDENCE
OF THE TOTAL π -ELECTRON ENERGY OF CONJUGATED SYSTEMS ?**

J. Cioslowski

*Department of Chemistry and Supercomputer Computations Research Institute
Florida State University, Tallahassee, Florida 32306-3006, USA*

(Received: June 1989)

Abstract : An approximate, empirical parameter-free formula that expresses the total π -electron energy of alternant ground state hydrocarbons as the function of the number of carbon atoms (N), the number of carbon-carbon bonds (M) and the algebraic structure count (K) is derived from the principle of corresponding spectral densities. The calculated energies match the exact energies within 0.13 % of the mean and 0.76 % of the maximal error. The present approach offers a solution of the problem concerning the (N,M,K)-dependence of E_{π} that is satisfactory from both theoretical (no empirical parameters) and practical (high accuracy) point of view.

INTRODUCTION

Among the concepts of chemical topology, the adjacency matrix, \underline{A} , plays a central role [1]. Let us introduce a formal definition of \underline{A} :

Definition: Let the square matrix \underline{A} has the elements A_{ij} . If these elements conform to the conditions:

1. A_{ij} is equal to 0 or 1 for any pair (i,j),
2. $A_{ij} = A_{ji}$ for any pair (i,j),
3. $A_{ii} = 0$ for any i,
4. $1 \leq \sum_j A_{ij} \leq 3$ for any i,

then \underline{A} is called the adjacency matrix.

In the present paper we deal with the adjacency matrices associated with alternant non-radical hydrocarbons, i.e. matrices that bear two additional properties:

1. All the eigenvalues of \underline{A} have non-zero values.
2. \underline{A} can be factorized as follows:

$$\underline{A} = \begin{bmatrix} \underline{0} & \underline{B} \\ \underline{B}^+ & \underline{0} \end{bmatrix} \quad (1)$$

The problem to be solved is derivation of an approximate formula that relates the trace of the absolute value of \underline{A} :

$$E_{\pi} = \text{Tr } |\underline{A}| \quad (2)$$

to the dimension of the matrix \underline{A} (N), its second moment:

$$2M = \text{Tr } \underline{A}^2 \quad (3)$$

and its determinant:

$$(-1)^{N/2} K^2 = \det \underline{A} \quad (4)$$

The above quantities can be also interpreted [1] as the number of carbon atoms (N), the number of carbon-carbon bonds (M) and the algebraic structure count (K) of the molecule described by A . The total π -electron energy of the molecule calculated within Hückel theory is identical with E_{π} [1] and can be used for assessing the stability of the molecule in question. Therefore, estimation of E_{π} from N , M and K is of great importance for it allows to estimate the energy without explicit diagonalization of A . It also constitutes an interesting mathematical problem.

HISTORICAL OVERVIEW

The problem of an approximate dependence of E_{π} upon N , M and K was addressed as early as in 1949 [2]. The formula proposed by Carter read:

$$E_{\pi} \approx A N + B \ln K + C, \quad (5)$$

where A , B and C were empirical parameters. Since then, more than twenty different formulae for E_{π} have been proposed, all of them containing some empirical parameters obtained by a numerical fitting [3]. Let us review some of them:

Eq.(2) has an advantage of being size consistent, but provides rather poor numerical approximation for E_{π} . A significant reduction of the discrepancy between the exact and approximate E_{π} is achieved in the equation [4]:

$$E_{\pi} \approx A N + B M + C [N (\ln K)^5]^{1/6} + D \quad (6)$$

with four empirical parameters. Equations:

$$E_{\pi} \approx A (Mq)^{1/2} + B K (q/M)^{q/2} + C \quad [5], \quad (7)$$

where:

$$q = (N/2) - 1; \quad (8)$$

and

$$E_{\pi} \approx A N + B M + C K D^{M-N} \quad [6] \quad (9)$$

offer a similar accuracy. Eventually, the formula [7]:

$$E_{\pi} \approx A (2MN)^{1/2} + B N K^{2/N} \quad (10)$$

has only two empirical parameters, but works as satisfactorily as the above formulae with three or four parameters.

The existence of several different, though quite accurate, relationships between E_{π} , N , M and K is certainly confusing and surprising at the first glance. The dependence of E_{π} upon K is expressed as linear (eqs.7 and 9), logarithmic (eq.5), or even more complicated, like $(\ln K)^{5/6}$ (eq.6) or $K^{2/N}$ (eq.10). How are all these different functions capable of providing quite good estimates of E_{π} ? The answer is simple. As pointed out by Gutman et al. [3], in the case of benzenoid hydrocarbons, the dependence on N and M accounts for more than 99 % of the energy. The remaining 1 % is due to variation in K (about 0.9 %) and other structural factors (about 0.1 %). Therefore it is not unexpected that almost any function of K (especially the one that has a lot of empirical parameters) can fit reasonably well to the exact energies.

The above conclusion has twofold consequences. The first one is of practical sort. The total π -electron energy can be easily estimated from the values of N , M and K and it really does not matter what approximate formula one uses. On the other hand, one has to admit that there is no *purely theoretical* approach to the problem of the (N, M, K) dependence of E_{π} and this prompts us to present the approach that bears *no empirical parameters* and, at the same time, provides very reasonable results.

THEORY

Let us denote the molecular graph describing a given hydrocarbon by G and the corresponding adjacency matrix by $A(G)$. The knowledge of the eigenvalues of $A(G)$, $\{x_i, i=1, N\}$ provides all the necessary information to write down the spectral density function [7-9]:

$$\Gamma(G, t) = \sum_{i=1}^N \delta(t - x_i) \quad (11)$$

The spectral density has to conform to the following constraints:

$$\int_{-\infty}^{+\infty} \Gamma(G, t) dt = N, \quad (12)$$

$$\int_{-\infty}^{+\infty} \Gamma(G, t) |t| dt = E_{\pi}, \quad (13)$$

$$\int_{-\infty}^{+\infty} \Gamma(G, t) t^2 dt = 2M, \quad (14)$$

and

$$\int_{-\infty}^{+\infty} \Gamma(G, t) \ln|t| dt = \ln K^2. \quad (15)$$

In the case when only the information about N , M and K is available, one has to make an assumption that the spectral density function depends only on these topological invariants:

$$\Gamma(G, t) \approx \Gamma(N, M, K, t). \quad (16)$$

Let us use the substitution [9]:

$$\Gamma(N, M, K, t) = (N^3/2M)^{1/2} G[N, M, K, (N/2M)^{1/2}t] \quad (17)$$

that brings the equations (12)-(15) into:

$$\int_{-\infty}^{+\infty} G(N, M, K, t) dt = 1, \quad (18)$$

$$\int_{-\infty}^{+\infty} G(N, M, K, t) |t| dt = (2MN)^{-1/2} E_{\pi} = e, \quad (19)$$

$$\int_{-\infty}^{+\infty} G(N, M, K, t) t^2 dt = 1, \quad (20)$$

$$\int_{-\infty}^{+\infty} G(N, M, K, t) \ln|t| dt = \ln [K^{2/N} (2M/N)^{-1/2}] = \ln k \quad , \quad (21)$$

and yields natural definitions for the McClelland quotient (e) and the normalized structure count (k) [8]. The next step in our reasoning is obvious: in order to satisfy the equations (18)-(21) for any value of N , M and K , $G(N, M, K, t)$ must be a sole function of only two parameters, namely k and t . As an immediate conclusion we obtain the result that e is related in an approximate way to k by a universal function $F(k)$ [7-10]:

$$e \approx F(k) \quad , \quad \text{or} \quad E_{\pi} \approx (2MN)^{1/2} F[K^{2/N} (2M/N)^{-1/2}] \quad . \quad (22)$$

In the derivation of this "principle of corresponding spectral densities" [11] no assumptions are made about the nature of the molecule represented by the graph G , except for the condition that it is an alternant hydrocarbon with $K \neq 0$.

Now one should pose the question about explicit form of the universal function $F(k)$. Simple linear and quadratic approximations have been proposed [7,8] together with a more elaborate approach involving a model spectral density function [9]. Very recently a further investigation on the form of $F(k)$ has been pursued by Gutman et al. [10]. All those attempts lean on fitting of empirical parameter to the exact total π -electron energies.

Since only very little was assumed about nature of the graph G , we can use the topological characteristics of some well-defined classes of conjugated molecules to enlighten the form of $F(k)$. The goal of this study would be twofold: First, one could verify whether different algebraic forms of $F(k)$ that originate from different classes of molecules yield similar values of e , therefore confirming (or denying) the validity of the hypothesis of corresponding spectral densities. Second, one could obtain empirical parameter-free formulae for E_{π} .

Two choices of the model classes appear obvious: aromatic annulenes and polyenes with an even number of carbon atoms. For the first family of molecules, e and k is related to the number of carbon atoms (N) through the formulae:

$$e = 2^{3/2} N^{-1} [\sin(\pi/N)]^{-1} \quad (23)$$

and

$$k = 2^{(4-N)/2N} \quad , \quad (24)$$

which yields:

$$e = F_1(k) = 2^{-1/2} u(k) \{ \sin [\pi u(k)/4] \}^{-1} \quad , \quad (25)$$

where:

$$u(k) = \log_2 2k^2 \quad . \quad (26)$$

For polyenes, the analogous relations read:

$$e = [2 / N(N-1)]^{1/2} \left[\{ \sin [\pi / (2N+2)] \}^{-1} - 1 \right] \quad , \quad (27)$$

$$k = [N/(2N-2)]^{1/2} \quad (28)$$

and

$$e = F_2(k) = k^{-1} (2k^2-1) \{ [\sin (\pi v(k)/2)]^{-1} - 1 \} \quad (29)$$

with:

$$v(k) = (2k^2-1) / (4k^2-1) \quad . \quad (30)$$

In spite of very different algebraic forms of $F_1(k)$ and $F_2(k)$, they provide remarkably similar values of e for k spanning a wide range of 0.7-1.0 (Table 1). This undoubtedly supports the above hypothesis.

NUMERICAL TESTING AND DISCUSSION

In order to obtain a further insight into numerical accuracy of the formulae (25) and (29), we performed calculations of the exact and approximate total π -electron energies for all planar, singlet ground state benzenoid hydrocarbons with 1-10 ten rings. The database comprised 18,388

structures and covered values of k lying within the range of 0.713-0.891. Both models were tested together with a simple linear and quadratic approximations to $F(k)$ [7,8]. The results are presented in Table 2.

It is clear that among the models without empirical parameters, the linear polyenes model gives the best results. The formulae with empirical parameters perform slightly better in average, but they result in higher maximal error. Taking the above observation into account, we conclude that the function $F_2(k)$ is an extremely good approximation to the relationship between k and e for benzenoid hydrocarbons and therefore their total π -electron energies can be adequately approximated by the equation:

$$E_{\pi} \approx 2N [K^{2/N} - (M/N) K^{-2/N}] \{ [\sin (\pi w/2)]^{-1} - 1 \} , \quad (31)$$

with

$$w = [K^{2/N} - (M/N) K^{-2/N}] / [2 K^{2/N} - (M/N) K^{-2/N}] . \quad (32)$$

Eventually, in Table 3 we present the results of test calculations for the exact and approximate energies of benzenoid hydrocarbons. The parameter-free formula (31) works as good as the approximations (6) and (9) that have four empirical parameters! In fact, these formulae have been the best approximations known until now.

The above examples convince us to consider regard the above theory as the final solution for the problem of the (N, M, K) dependence of E_{π} . The derived approximation provides results of a considerable accuracy. The approximate energy is size-extensive and is an increasing function of the algebraic structure count, K , as expected from the VB theory.

We believe that in this paper we have been able to shed more light at the topological dependence of the total π -electron energy of conjugated systems. Our approach does not require any empirical parameters. Instead, we use the energies of linear polyenes as a template for all other systems.

Therefore the present work is not just another numerical experiment on the E_r , but has a sound mathematical foundations. It can be a starting point for more research, particularly on the resonance energy.

Table 1. Comparison of the functions $F_1(k)$ and $F_2(k)$.

k	$F_1(k)$	$F_2(k)$
1.00	1.000	1.000
0.98	0.988	0.989
0.96	0.977	0.977
0.94	0.966	0.967
0.92	0.956	0.956
0.90	0.947	0.947
0.88	0.938	0.937
0.86	0.931	0.929
0.84	0.924	0.921
0.82	0.918	0.914
0.80	0.912	0.908
0.78	0.908	0.903
0.76	0.904	0.900
0.74	0.902	0.898
0.72	0.901	0.899
0.70	0.900	0.902

Table 2. Errors of various approximations to e.

Approximation	Error		
	mean dev.	RMS	maximal
aromatic annulenes model	$2.98 \cdot 10^{-3}$	$3.11 \cdot 10^{-3}$	$5.26 \cdot 10^{-3}$
linear polyenes model	$1.21 \cdot 10^{-3}$	$1.74 \cdot 10^{-3}$	$6.85 \cdot 10^{-3}$
linear fit	$7.66 \cdot 10^{-4}$	$1.16 \cdot 10^{-3}$	$1.77 \cdot 10^{-2}$
quadratic fit	$5.90 \cdot 10^{-4}$	$8.39 \cdot 10^{-4}$	$9.37 \cdot 10^{-3}$

Table 3. Relative errors of various (N,M,K)-approximations to E_x .

Approximation	Error		
	mean dev.	RMS	maximal
aromatic annulenes model	0.33 %	0.34 %	0.58 %
linear polyenes model	0.13 %	0.19 %	0.76 %
linear fit	0.08 %	0.13 %	1.88 %
quadratic fit	0.07 %	0.09 %	0.99 %
eq.9 [6]	0.22 %	0.23 %	0.75 %
eq.6 [4]	0.08 %	0.14 %	1.17 %

REFERENCES

- [1] for a review see: I.Gutman and O.E.Polansky, "Mathematical Concepts in Organic Chemistry", Springer-Verlag, Berlin, 1986 and also A.Graovac, I.Gutman and N.Trinajstić, "Topological Approach to the Chemistry of Conjugated Molecules", Springer-Verlag, Heidelberg, 1977
- [2] P.C.Carter, Trans.Paraday Soc. 45 (1949) 597
- [3] for an extensive, though not exhaustive review see: I.Gutman, Lj.Nedeljković and A.V.Teodorović, Bull.Soc.Chim.Beograd 48 (1983) 495
- [4] I.Gutman and N.Trinajstić, Croat.Chem.Acta 48 (1976) 297
- [5] I.Gutman, A.V.Teodorović and I.Lj.Nedeljković, Theor.Chim.Acta 65 (1984) 23
- [6] G.G.Hall, Inst.Math.Appl. 17 (1981) 70
- [7] J.Ciosłowski, MATCH 20 (1986) 95
- [8] J.Ciosłowski, Intern.J.Quantum.Chem. 31 (1987) 581
- [9] J.Ciosłowski and O.E.Polansky, Theor.Chim.Acta, 74 (1988) 55
- [10] I.Gutman, S.Marković and M.Marinković, MATCH 22 (1987) 277
- [11] this principle has been called "a uniform distribution approach (UDA)" in refs. 7 and 10