Theor Chim Acta (1988) 74: 55-62

Theoretica Chimica Acta

Topological hamiltonian spectra of polycyclic benzenoid hydrocarbons

J. Cioslowski* **and O. E. Polansky**

Max-Planck-Institut für Strahlenchemie, Stiftstraße 34-36, D-4330 Mülheim a.d. Ruhr, Federal Republic of Germany

(Received July 21, revised October 7/Accepted October 15, 1987)

A general theory which points out the relations between Hückel π-electron energy, the number of Kekulé structures and the HOMO-LUMO separation is presented. Some normalized topological invariants are derived from the concept of the spectral density function. A reasonably simple (three parameters) model spectral density function leads to universal relations between topological invariants that, although valid for any alternant molecule, were tested numerically for polycyclic benzenoid hydrocarbons. Some general conclusions concerning a distribution of the adjacency matrix eigenvalues are drawn.

Key words: Conjugate molecules — Polycyclic benzenoid hydrocarbons, total π-electron energy — Spectral density function

Introduction

After more than 10 years in the realm of chemical sciences, chemical topology has established a position as an approach enabling us to investigate the molecular properties of conjugated compounds. Building a bridge between the abstract concepts of graph theory and common experimental observations, this branch of chemistry is a proven tool of practical importance. However, as far as π-electron properties are concerned, it seems unavoidable that the focus of applications of

---

* *Present address*: Department of Chemistry, University of North Carolina at Chapel Hill, NC 27514, USA

chemical topology is on the chemistry and physics of polycyclic aromatic hydrocarbons. For these molecules, which have essentially identical carbon-carbon bonds and have only one kind of atom involved in the π-electron framework, we have a strong basis for belief in the validity of the Hückel theory, which is one of the building blocks of chemical topology.

Taking this into account, we decided to suggest in the title of the present paper that the approach we present is especially useful in the case of the benzenoid hydrocarbons. However, it should be emphasized that utilization of a spectral density function, which describes the distribution of eigenvalues of a particular Hamiltonian, enables one to construct some approximate relations between graph invariants and the electronic properties of any π-conjugated molecule; moreover, these relations can probably be extended beyond the Hückel Hamiltonian.

In the first section of this paper we demonstrate the existence of some general relations between topological invariants and both the total π-electron energy and the HOMO-LUMO separation. The results are used to construct a model approach pointing out these relations for benzenoid hydrocarbons, which are finally discussed in the third section.

General relations between normalized topological invariants

A great part of the work toward elucidation of the stability of conjugated alternant molecules on the ground of chemical topology has been devoted to the derivation and numerical verification of various approximate formulae that express the total π-electron energy (E_π) as a function of the number of carbon atoms (N), carbon-carbon bonds (M) and Kekulé structures (K) [1-8]. Much less attention has been directed to analogous expressions for the HOMO-LUMO separation (X_{HL}) [8-12], which is another measure of the molecular reactivity [13]. Unfortunately, a large number of the $E_\pi(N, M.K)$ formulae are incorrect because they do not conform to some "first principle" constraints.

In order to get more insight into this matter one should consider the following reasoning. Let A be the adjacency matrix corresponding to some particular conjugated molecule Q, and $x_1, \ldots, x_N$ be the set of its eigenvalues. The spectral density function is defined as:

$$F(Q; x) = \sum_{i=1}^{N} \delta(x - x_i). \tag{1}$$

The function F fulfils the relations [7, 8]:

$$\int_{-\infty}^{+\infty} F(Q; x)\, dx = N; \tag{2}$$

$$\int_{-\infty}^{+\infty} F(Q; x)|x|\, dx = E_\pi; \tag{3}$$

$$\int_{-\infty}^{+\infty} F(Q; x) x^2 \, dx = 2M; \tag{4}$$

$$\int_{-\infty}^{+\infty} F(Q; x) \, ln \, |x| \, dx = 2 \, ln \, K. \tag{5}$$

The last relation is valid only (in general) for benzenoid hydrocarbons. Moreover:

$$X_{HL} = 2a; \qquad a = \max_{x} x\colon \ \underset{0 \le t < x}{\forall} \ F(Q; t) = 0. \tag{6}$$

From Eqs. (2)-(6) we learn that the knowledge of F is sufficient to calculate N, M, K, E_π and X_{HL}. However, when we seek to derive an approximate formula for E_π (or for X_{HL}) as a function of N, M and K, we impose certain approximations on the form of the spectral density function. Being more precise, we hope that the knowledge of only N, M and K determines F to a large extent, and then we assume that F depends only on these quantities. Only under this strong assumption we can expect to express E_π (and X_{HL}) solely in terms of N, M and K.

Two entirely different kinds of errors in calculated values of $E_\pi(X_{HL})$ may arise from this approximate formula. First, there is some error due to the fact that F also depends on other structural features of A, so that molecules having the same values of N, M and K have (we hope only slightly) different E_π and X_{HL}. Second, there *may* be some error because our formula *may*, on the average, incorrectly account for the dependence between $E_\pi(X_{HL})$ and N, M, K.

With the above assumption we put:

$$F(Q; x) \equiv F(N, M, K; x). \tag{7}$$

Introducing an auxiliary function

$$G(N, M, K; x) = \sqrt{\frac{2M}{N^3}} \, F\left(N, M, K; \sqrt{\frac{2M}{N}} \, x\right) \tag{8}$$

we can rewrite Eqs. (2)-(6) as:

$$\int_{-\infty}^{+\infty} G(N, M, K; x) \, dx = 1; \tag{9}$$

$$\int_{-\infty}^{+\infty} G(N, M, K; x) |x| \, dx = \frac{E_\pi}{\sqrt{2MN}}; \tag{10}$$

$$\int_{-\infty}^{+\infty} G(N, M, K; x) x^2 \, dx = 1; \tag{11}$$

$$\int_{-\infty}^{+\infty} G(N, M, K; x) \, ln \, x \, dx = ln \frac{K^{2/N}}{\sqrt{2M/N}}. \tag{12}$$

It is not difficult to realize that in order to fulfil Eqs. (9), (11) and (12) for any combination of N, M and K, the function G has to depend only on the argument

x and the parameter

$$k=\frac{K^{2/N}}{\sqrt{2M/N}} \tag{13}$$

which we call the normalized structure count (NSC) [8].

From Eq. (10) we arrive immediately at the conclusion that the McClelland quotient (MCQ) [7, 8],

$$e=\frac{E_\pi}{\sqrt{2MN}}, \tag{14}$$

can be expressed solely as a function of k. From Eq. (6) we also see that the normalized HOMO-LUMO separation (NHLS) [8],

$$h=\frac{1}{2}\frac{X_{HL}}{\sqrt{2M/N}}, \tag{15}$$

also depends only on k.

The above results can be compiled in the following.

Theorem. *Any approximate formulae expressing* E_π *and* X_{HL} *as a function of only* N, M *and* K *must have the form*:

$$E_\pi=\sqrt{2MN}\,R\left[\frac{K^{2/N}}{\sqrt{2M/N}}\right] \tag{16}$$

and

$$X_{HL}=2\sqrt{2M/N}\,S\left[\frac{K^{2/N}}{\sqrt{2M/N}}\right] \tag{17}$$

where R and S are some functions with the property

$$R(1)=1; \qquad S(1)=1. \tag{18}$$

The relations (18) are easily derived from the fact [8, 14] that, when $k=1$, both e and h approach their extreme values of one, which correspond to the function F having the form:

$$F(N, M, K; x)=\frac{N}{2}\left[\delta\left(x-\sqrt{\frac{2M}{N}}\right)+\delta\left(x+\sqrt{\frac{2M}{N}}\right)\right]. \tag{19}$$

We end this section with some important remarks. The first one is that (unfortunately) the majority of the known topological formulae for E_π do not conform to Eq. (16). This calls their reliability into question. Even worse, many of them also do not fulfil an obvious size consistency condition:

$$E_\pi(uN, uM, K^u)=uE_\pi(N, M, K); \qquad u=2, 3, \ldots \tag{20}$$

The above theorem provides us with a general form of E_π which is indeed size consistent.

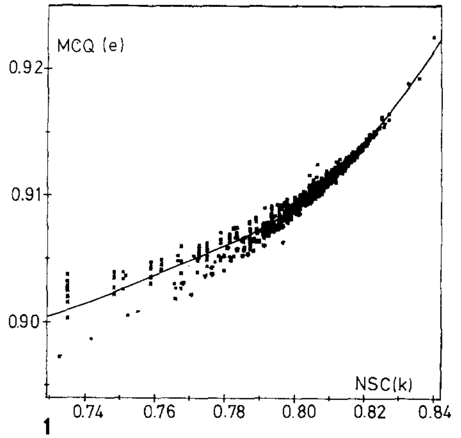
The next point is that our considerations remain valid for any conjugated systems if we replace the number of Kekulé structures (K) by the algebraic structure count [2]. Finally we point out that what we are offered from Eqs. (16) and (17) is *the best* we can expect if we attempt to calculate E_π and X_{HL} with information that is confined to the knowledge of N, M and K. One should therefore realize that the results presented in the next section cannot be further improved (except by some more sophisticated numerical fitting). This should clean out the plethora of reported formulae, of which some tens are currently known [1].

A model spectral density function for benzenoid hydrocarbons

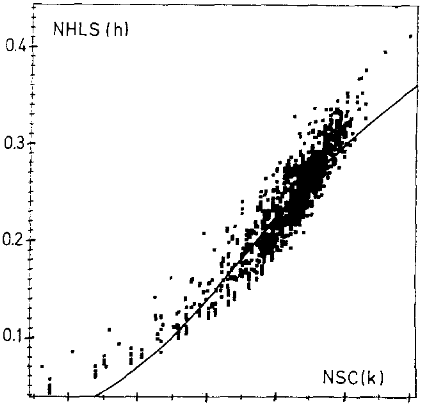
The general results from the previous section, which have been derived before using very particular assumptions about the function F ("the uniform distribution approach") [7, 8], suggest a route for more detailed considerations on the dependencies between NSC, MCQ and NHLS. Polycyclic benzenoid hydrocarbons, BHs, are an especially good field for such an analysis because:

1. They form a class of molecules having similar topological characteristics.
2. Most of the known topological formulae were tested on BHs.
3. There is a standard set of BHs known to statistically represent the entire class of molecules [15].

Before making any assumptions about the approximate form of F, we first have to check whether there are any reasonable correlations between k, e and h. For this purpose we plotted the respective quantities calculated for 1030 singlet ground state planar BHs. These plots, which we present in Figs. 1-3, show that, despite some scatter in the points, the normalized topological invariants are indeed highly correlated.



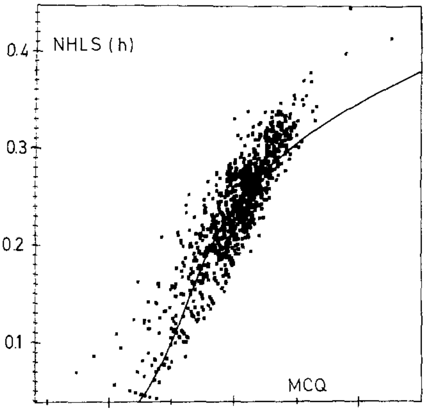
1



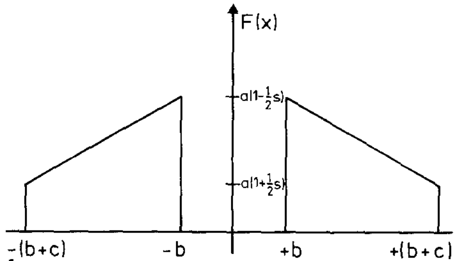
2

Fig. 1. NSC (k) vs MCQ (e) plot for benzenoid hydrocarbons

Fig. 2. NSC (k) vs NHLS (h) plot for benzenoid hydrocarbons



3



4

Fig. 3. MCQ (e) vs NHLS (h) plot for benzenoid hydrocarbons

Fig. 4. A model spectral density function

Because of this, we decided to seek some simple form of model spectral density function which could mimic in a global way the real distribution of eigenvalues. The simplest model F that we tested numerically was a saw-like function (Fig. 4):

$$F(x)=\begin{cases} a\left[1+s\left(\frac{|x|-b}{c}-0.5\right)\right] & x\in\langle b, b+c\rangle\cup\langle-(b+c), -b\rangle \\ 0 & \text{otherwise.}\end{cases} \tag{21}$$

For this particular form of the spectral density function b is equal to half the HOMO-LUMO separation, $v=b/c$ describes the relative position of the distribution and s is the slope. The parameters a, b, c and s can be found from Eqs. (2)-(5):

$$N=2ac; \tag{22}$$

$$E_\pi=2ac^2[v+\tfrac{1}{2}+\tfrac{1}{12}s)]; \tag{23}$$

$$2M=2ac^3[v^2+2v(\tfrac{1}{2}+\tfrac{1}{12}s)+(\tfrac{1}{3}+\tfrac{1}{12}s)]; \tag{24}$$

$$2\,ln\,K=2ac(1-\tfrac{1}{2}s)[(v+1)\,ln\,(v+1)-v\,ln\,v-1] + 2acs\left[\frac{(v+1)^2}{2}\,ln\,(v+1)-\frac{v^2}{2}\,ln\,v-\tfrac{1}{2}v-\tfrac{1}{2}v-\tfrac{1}{4}\right]. \tag{25}$$

The normalized topological invariants can be then found as [8]:

$$e=\left[1-\frac{\frac{1}{12}-\frac{1}{144}s^2}{v^2+2v(\frac{1}{2}+\frac{1}{12}s)+(\frac{1}{3}+\frac{1}{12}s)}\right]^{1/2}; \tag{26}$$

$$ln\,k=(1-\tfrac{1}{2}s)[(v+1)\,ln\,(v+1)-v\,ln\,v-1]+s\left[\frac{(v+1)^2}{2}\,ln\,(v+1);\right. \left.-\frac{v^2}{2}\,ln\,v-\tfrac{1}{2}v-\tfrac{1}{4}\right]-\tfrac{1}{2}\,ln\,[v^2+2v(\tfrac{1}{2}+\tfrac{1}{12}s)+(\tfrac{1}{3}+\tfrac{1}{12}s)]; \tag{27}$$

$$h=v[v^2+2v(\tfrac{1}{2}+\tfrac{1}{12}s)+(\tfrac{1}{3}+\tfrac{1}{12}s)]^{-1/2}. \tag{28}$$

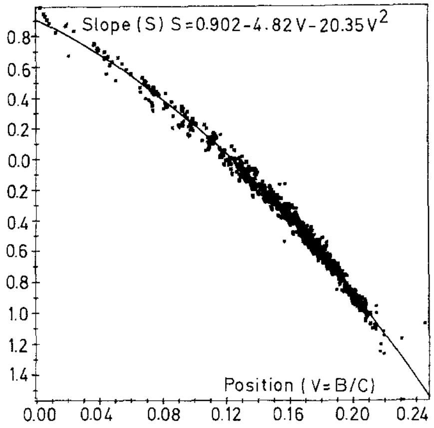


Fig. 5. Relative position (v) vs slope (s) correlation

For the slope s there are several possibilities to be investigated numerically:

1. We can, with a given s, calculate v from the knowledge of k and then obtain estimates of e and h. The slope can be:
(a) fixed at $s=0$, which corresponds to a rectangular distribution of the eigenvalues; (b) optimized for the entire set of BHs.

2. Another possibility is to calculate s and v from the knowledge of k and e and then to seek for a possible correlation between s and v.

As expected, the second approach gives much better results. Moreover, there is a definite correlation between s and v as can be seen from Fig. 5. Having fitted a parabola to the points representing the set of hydrocarbons we end up with the following algorithm to calculate NHLS and MCQ if NSC is known:

1. Solve the system of Eq. (27) and

$$s = 0.902 - 4.82v - 20.35v^2 \qquad (29)$$

with respect to s and v. One should note, that necessarily $v > 0$.

2. From Eqs. (26) and (28) calculate e and h.

Using this algorithm we were able to draw the theoretical curves in Figs. 1-3. It is apparent that the overall agreement is satisfactory.

Discussion

The formalism derived above enables us to highlight some general properties of benzenoid hydrocarbons:

1. It is clear that there is a definite correlation between the total π-electron, the number of Kekulé structures and the HOMO-LUMO separation. However, these quantities have to be normalized properly before looking for any regression between them.

2. There are two distinct classes of benzenoid hydrocarbons. The first one consists of the compounds with a relatively small stability (having small values of h, k, e and v). These molecules tend to have their adjacency matrix eigenvalues grouped at rather high values of energy (slope $s > 0$). The hydrocarbons belonging to the second class are characterized by a high stability and their eigenvalues are clustered around the first one (slope $s < 0$). This means that only two different distributions of the eigenvalues are realized: either there is the first eigenvalue with a relatively high value of x, which is close to the subsequent eigenvalues, or there is the first eigenvalue with a small value of x, but then largely separated from the next ones. This observation can have practical importance for photochemistry, in which we are interested in the accessibility of higher excited electronic states whose energy depends in some (indirect) way on the distribution of the Hückel Hamiltonian eigenvalues.

3. The model can also be used for an approximate calculation of other topological invariants like moments of $\boldsymbol{A}$ or the modified topological index Z.

The theoretical curves drawn in Figs. 1-3 make it possible to perform a fast calculation of E_π and X_{HL} without performing any computations. We therefore hope that they will provide the organic chemists with a possibility of "paper and pencil" estimation of the stability of benzenoid hydrocarbons. It shouls also be pointed out that new reactivity indeces can be constructed by considering derivatives of F with respect to N, M and K. Work on this problem is in progress.

Acknowledgement. One of the authors (J.C.) expresses his thanks to the Max-Planck-Gesellschaft for providing him with a fellowship during his stay in Mülheim.

References

1. Gutman I, Nedeljković L, Teodorović AV (1983) Bull Soc Chim Beograd 48:495
2. Graovac A, Gutman I, Trinajstić N (1977) Topological approach to the chemistry of conjugated molecules. Springer, Berlin Heidelberg New York
3. Gutman I, Trinajstić N, Wilcox C F, Jr (1975) Tetrahedron 31:143
4. Hall GG (1981) Inst Math Appl 17:70
5. Gutman I, Petrović S (1983) Chem Phys Lett 97:242
6. Gutman I, Teodorović AV, Nedeljković L (1984) Theor Chim Acta 65:23
7. Cioslowski J (1986) Match 20:95; Gutman I (1986) MATCH 21:95
8. Cioslowski J (1987) Int J Quantum Chem 31:581
9. Hall GG (1977) Mol Phys 33:551
10. Graovac A, Gutman I (1980) Croat Chem Acta 53:45
11. Gutman I (1981) MATCH 11:75
12. Gutman I, Rouvray D H (1979) Chem Phys Lett 62:384
13. Fukui K (1970) Topics Current Chem, vol 15. Springer, Berlin Heidelberg New York
14. Gutman I, Cioslowski J (1987) Bull Chem Soc Beograd 42:125
15. Cioslowski J (1987) J Comput Chem 8:906