COULSON FUNCTION AND TOTAL π-ELECTRON ENERGY

Ivan Gutman, Boris Furtula and Dušica Vidović

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

ABSTRACT. It is shown that the value of the variable x, at which the Coulson function $F(G,x)$ has its inflection point, is related to the respective total π-electron energy (E) and depends mainly on E and the number of non-bonding molecular orbitals.

INTRODUCTION

In another paper [1] the basic properties of the Coulson function

$$F(x) = F(G, x) = n - \frac{ix\,\phi'(G, ix)}{\phi(G, ix)} \ . \tag{1}$$

were determined. This function, introduced by Charles Coulson long time ago [2], plays a significant role in the theory of the total π-electron energy of conjugated molecules (for details see [1, 3, 4]).

On the right–hand side of Eq. (1), $\phi(G, x)$ denotes the characteristic polynomial of the molecular graph G, n is the number of vertices of G and $i = \sqrt{-1}$. For more details see [1, 3, 4].

In the work [1] the spectral decomposition of the Coulson function was shown to be

$$F(G, x) = \sum_{j=1}^{n} \frac{\lambda_j^2}{x^2 + \lambda_j^2} \tag{2}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ stand for the eigenvalues of G. To be more precise: the right–hand side of Eq. (2) is equal to the real part of the Coulson function [1], a feature that is immaterial for the present considerations.

It was demonstrated [1] that for all graphs G (either representing conjugated π-electron systems or saturated hydrocarbons or no molecular species at all), $F(G, x)$ is an even bell–shaped function. From this fact follows that $F(G, x)$ has a unique maximum at $x = 0$ and two symmetrically arranged inflection points, at $x = x_{infl} > 0$ and $x = -x_{infl}$.

In order to calculate x_{infl} one has to find the second derivative of $F(G, x)$ and to equate it with zero. Utilizing formula (2), we immediately arrive at

$$F'(G, x) = -2 \sum_{j=1}^{n} \frac{\lambda_j^2\, x}{(x^2 + \lambda_j^2)^2}$$

$$F''(G, x) = 2 \sum_{j=1}^{n} \frac{\lambda_j^2\, (3\, x^2 - \lambda_j^2)}{(x^2 + \lambda_j^2)^3}$$

implying that x_{infl} is the unique solution of the equation

$$3\, x^2 \sum_{j=1}^{n} \frac{\lambda_j^2}{(x^2 + \lambda_j^2)^3} = \sum_{j=1}^{n} \frac{\lambda_j^4}{(x^2 + \lambda_j^2)^3} \tag{3}$$

for $x > 0$.

According to (2) the Coulson function is equal to the sum of n increments, each being associated with a particular eigenvalue of the respective molecular graph G. These increments have the form

$$f(\lambda_j, x) = \frac{\lambda_j^2}{x^2 + \lambda_j^2}$$

for $j = 1, 2, \ldots, n$. If $\lambda_j = 0$ then $f(\lambda_j, x) \equiv 0$ and, consequently, the summation in (2) can be restricted to the $n - n_0$ non-zero eigenvalues of G. As before [1], n_0 denotes the multiplicity of the number zero in the spectrum of the graph G.

Recall that in Hückel molecular orbital theory, n_0 is just the number of non-bonding molecular orbitals [3, 5]. Several methods for evaluating n_0 (without actually computing the graph eigenvalues) were put forward [6, 7].

If $\lambda_j \neq 0$, then the increment $f(\lambda_j, x)$ is an even bell–shaped function, and its inflection points are at $x = |\lambda_j|/\sqrt{3}$ and $x = -|\lambda_j|/\sqrt{3}$, see [1]. Then, in view of (2), x_{infl} will be greater than the minimal value of $|\lambda_j|/\sqrt{3}$, provided $\lambda_j \neq 0$, and smaller than the maximal value of $|\lambda_j|/\sqrt{3}$. In view of this, the average value of $|\lambda_j|/\sqrt{3}$ (over all non-zero eigenvalues)

appears to be a reasonable approximation for x_{infl}. If so, then

$$x_{infl} \approx \frac{1}{n - n_0} \sum_{j=1}^{n} |\lambda_j| \ . \tag{4}$$

The sum occurring on the right–hand side of (4) is immediately recognized as the *energy of the graph* G [4, 8], a graph invariant much studied in both theoretical chemistry (see, for instance, [4,9–14] and the references cited therein), and mathematics (see, for instance, [15–21] and the references cited therein). Note that if G is a molecular graph of an alternant hydrocarbon, then its "energy" is precisely equal to the total π-electron energy, as computed within the Hückel molecular orbital model and expressed in units of the carbon–carbon resonance integral β [5]. Note also that the same quantity, under the name *eigenvalue sum* was recently used for describing physico–chemical properties of alkanes [22]. In fact, the "eigenvalue sum" is simply equal to one half of the "energy".

In what follows we denote the quantity $|\lambda_1| + |\lambda_2| + \cdots + |\lambda_n|$ by E and refer to it as either "total π-electron energy" or "graph energy". Then formula (4) is rewritten as

$$x_{infl} \approx \frac{E}{n - n_0} \ . \tag{5}$$

The quantity E is in the following manner related with the Coulson function [2]:

$$E = \frac{1}{\pi} \int_{-\infty}^{+\infty} F(x) \, dx \ .$$

Thus, in order to compute E one has to integrate the Coulson function over the entire interval $(-\infty, +\infty)$. We now show that some information about E can be obtained by examining the Coulson function in just a single point – at its inflection point. This result is obtained by means of the approximation (5).

TESTING THE APPROXIMATION (5)

In order to test the formula (5) we have first computed x_{infl} (as the solution of Eq. (3)) for all chemical trees with $n \leq 10$. Recall that chemical trees are trees in which the maximal vertex degree does not exceed 4; these are molecular graphs of alkanes. A typical result, for $n = 10$, is shown in Figure 1.
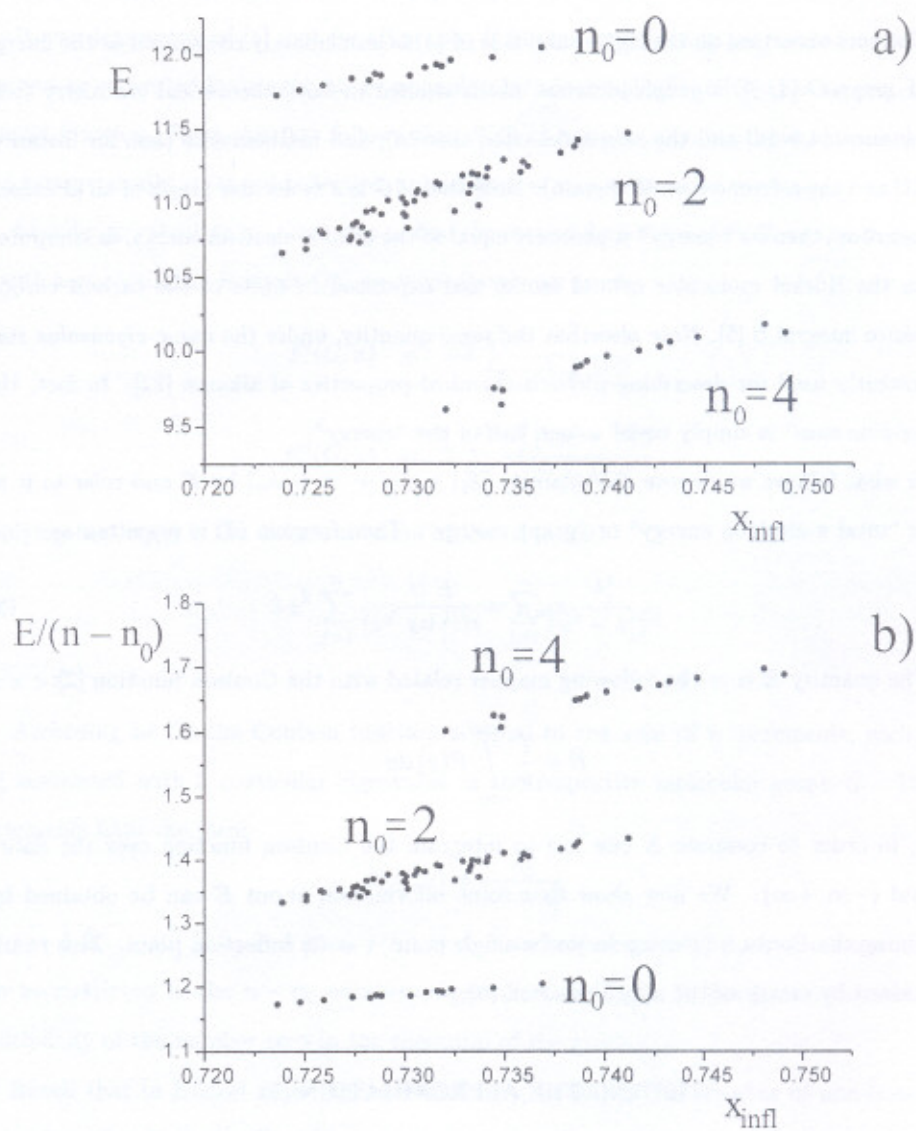
Figure 1. Relation between the energy (E) of the 10-vertex chemical trees and the value of the variable x at which the corresponding Coulson function has an inflection point (x_{infl}); n_0 is the multiplicity of the number zero in the graph spectrum; for details see text.

In Figure 1a are plotted the energies of the 75 distinct 10-vertex alkane graphs versus the corresponding x_{infl}-values. The data points are, in an obvious manner, separated into three clusters. By direct checking we established that all data points, belonging to the same cluster, pertain to molecular graphs with equal number of zero eigenvalue (n_0). In the case of 10-vertex chemical trees n_0 is equal to either 0 or 2 or 4, hence the three clusters.

The same clustering of data points is found also when instead of E we use $E/(n - n_0)$. This is shown in Figure 1b. Here, of course, the position of the clusters is inverse to what we have in Figure 1a.

Fully analogous results were observed also for chemical trees with $n = 6, 7, 8$, and 9 vertices, and there is no doubt that regularities of the same kind hold also for $n > 10$.

The above described relation between E and x_{infl} is by no means restricted to chemical trees. This is seen from the example shown in Figure 2, where we have plotted E and $E/(n - n_0)$ versus x_{infl} for all 10-vertex trees. There exist 106 such trees and their n_0-values are equal to either 0 or 2 or 4 or 6 or 8. There is a single 10-vertex tree with $n_0 = 8$, namely the star. In both Figure 2a and Figure 2b we see that the data points form four well-separated clusters, with a single point (corresponding to the star), not belonging to any of the clusters. Each cluster of data points correspond to trees with the same value of n_0.

Fully analogous results were observed also for trees with $n = 6, 7, 8$, and 9 vertices, and we believe that regularities of the same kind hold also for $n > 10$.

From both Figures 1 and 2 it is seen that for any fixed value of n_0, there is a relatively good (but far from perfect) correlation between $E/(n - n_0)$ and x_{infl}.

From Eq. (5) one would expect that this correlation is linear and, furthermore, that the respective line goes through the origin. Already from Figures 1 and 2 we see that such expectations are only partially fulfilled. The main violation from Eq. (5) is the fact that the data points lie not on a single, but on several correlation lines.

In most cases the correlation between $E/(n - n_0)$ and x_{infl} is linear. One example is shown in Figure 3, in which the data points for 9-vertex trees with $n_0 = 3$ are depicted. We see that with a single exception all data points lie near a straight line; the outlier (the "single exception") is indicated by an arrow. More details on this and other linear correlations are found in Tables 1 and 2, where also the outliers are specified.
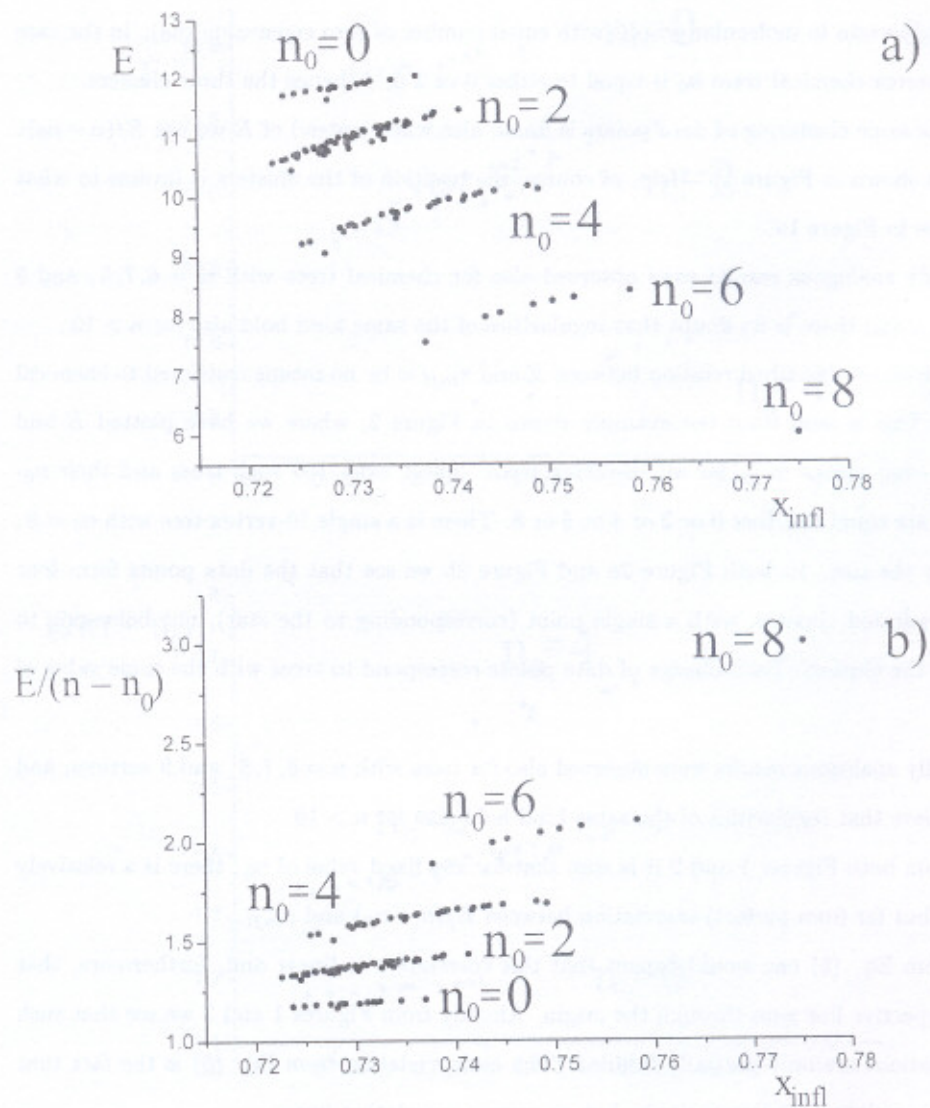
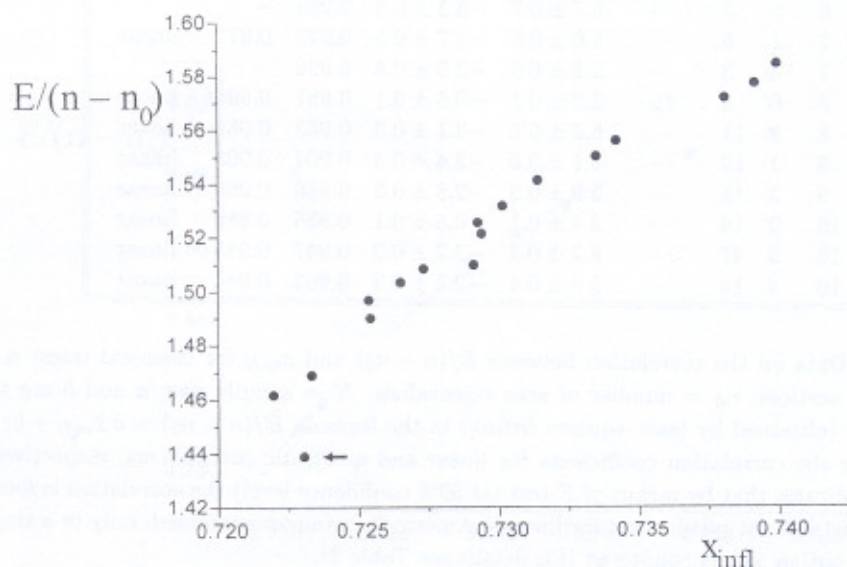**Figure 2.** Same data as in Figure 1, for the 10-vertex trees.

Figure 3. Relation between the energy (E) and x_{infl} for the 9-vertex trees with three zero eigenvalues ($n_0 = 3$). The outlier, indicated by arrow, correspond to the tree T_B depicted in Figure 5. More data on this correlation are found in Table 2.

An inspection of Figure 2 reveals that not all correlations are linear. A convincing example for this is provided in Figure 4, depicting the data points for the 9-vertex trees with $n_0 = 5$. For further details see Tables 1 and 2 as well as the discussion following them.

In order to learn more on the nature and quality of the correlation between $E/(n - n_0)$ and x_{infl} we have examined all sets of trees and chemical trees with fixed values of the parameters n and n_0, for $n \leq 10$. For obvious reasons, sets containing less than three elements were not analyzed. In all other sets we have drawn the best straight line through the data points (by means of the method of least squares). If the number of elements was greater than 3, then we have also drawn the best quadratic parabola through the data points (an example is shown in Figure 4), which should account for curvilinearity. Whether the correlation between $E/(n - n_0)$ and x_{infl} is linear or curvilinear was then determined by means of F-test.

The results obtained are collected in Tables 1 and 2.

n	n_0	N	outlier	a	b	R	R_q	remark
6	2	3	—	6.7 ± 0.7	-3.3 ± 0.5	0.994	—	
7	1	6	—	5.6 ± 0.6	-2.7 ± 0.5	0.975	0.977	linear
7	3	3	—	5.3 ± 0.5	-2.2 ± 0.4	0.996	—	
8	0	5	T_A	2.5 ± 0.1	-0.6 ± 0.1	0.997	0.9996	linear
8	2	11	—	6.3 ± 0.4	-3.1 ± 0.3	0.983	0.985	linear
9	1	19	—	5.1 ± 0.6	-2.4 ± 0.4	0.904	0.905	linear
9	3	15	—	5.9 ± 0.3	-2.8 ± 0.2	0.986	0.990	linear
10	0	14	—	2.4 ± 0.1	-0.6 ± 0.1	0.986	0.986	linear
10	2	47	—	6.2 ± 0.3	-3.2 ± 0.2	0.947	0.948	linear
10	4	14	—	5.6 ± 0.4	-2.5 ± 0.3	0.963	0.984	linear

Table 1. Data on the correlation between $E/(n - n_0)$ and x_{infl} for chemical trees: $n =$ number of vertices; $n_0 =$ number of zero eigenvalues; $N =$ sample size; a and b are the coefficients (obtained by least–squares fitting) in the formula $E/(n - n_0) \approx a\,x_{infl} + b$; R and R_q are the correlation coefficients for linear and quadratic correlations, respectively; "linear" indicates that by means of F-test (at 99% confidence level) the correlation is found to be linear (and not quadratic curvilinear). Among the samples examined, only in a single sample an outlier was encountered (for details see Table 2).

n	n_0	N	outlier	a	b	R	R_q	remark
6	2	3	—	6.7 ± 0.7	-3.3 ± 0.5	0.994	—	
7	1	6	—	5.6 ± 0.6	-2.7 ± 0.5	0.975	0.977	linear
7	3	4	—	7.1 ± 0.9	-3.5 ± 0.6	0.985	0.9999	curvilinear
8	0	5	T_A	2.5 ± 0.1	-0.6 ± 0.1	0.997	0.9996	linear
8	2	12	—	6.7 ± 0.5	-3.5 ± 0.3	0.977	0.983	linear
8	4	5	—	8.3 ± 1.1	-4.3 ± 0.8	0.974	0.99988	curvilinear
9	1	20	—	5.5 ± 0.6	-2.6 ± 0.4	0.910	0.913	linear
9	3	20	T_B	6.6 ± 0.3	-3.3 ± 0.2	0.986	0.994	linear
9	5	6	—	8.7 ± 1.2	-4.6 ± 0.9	0.965	0.9995	curvilinear
10	0	15	T_C	2.4 ± 0.1	-0.6 ± 0.1	0.986	0.986	linear
10	2	52	T_D	6.2 ± 0.3	-3.1 ± 0.2	0.956	0.957	linear
10	4	31	T_E	6.7 ± 0.3	-3.3 ± 0.2	0.975	0.992	curvilinear
10	6	7	—	10.2 ± 1.3	-5.6 ± 1.0	0.960	0.9989	curvilinear

Table 2. Same data as in Table 1 for trees. In five cases one data point was an outlier; then this point was not taken into account and the size of the sample employed was by one smaller that N. The trees corresponding to the outliers are depicted in Figure 5. By "curvilinear" is indicated that the quadratic correlation is significantly better than the linear one (as decided by means of F-test at 99% confidence level); "linear" indicates that the opposite applies.
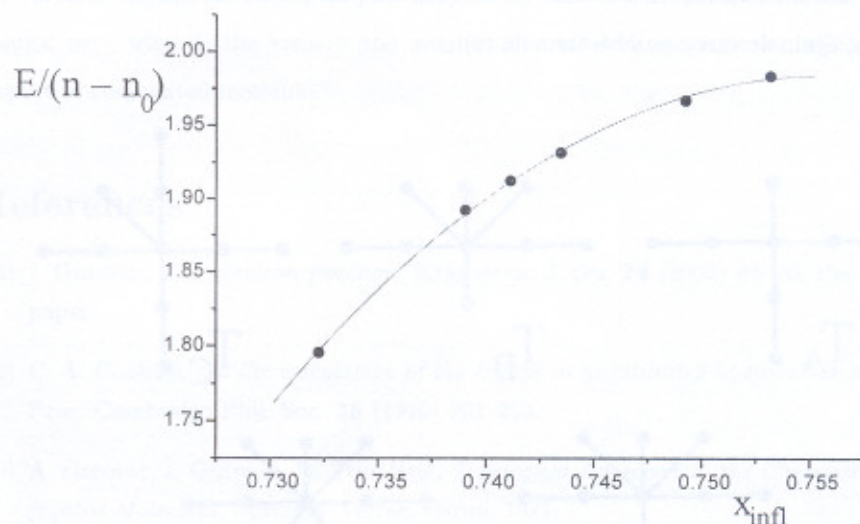
Figure 4. An example of curvilinear relation between the energy (E) and x_{infl}: the 9-vertex trees with five zero eigenvalues $(n_0 = 5)$.

In five cases, specified in Tables 1 and 2, a single data point was an obvious outlier. (For one such case see Fig. 3.) These outliers were excluded from the statistical analysis. The trees corresponding to the outliers are depicted in Figure 5.

DISCUSSION AND CONCLUDING REMARKS

In this work we examined only acyclic (molecular) graphs and, strictly speaking, all our conclusions hold only for them. We, however, believe that the same or similar results hold also for cycle–containing systems. Anyway, the study of the inflection point of the Coulson function of cycle–containing (molecular) graphs is a task that awaits to be accomplished in the future.

The calculated values for the coefficients a (given in Tables 1 and 2) indicate the there is an approximate (linear or curvilinear) proportionality between $E/(n - n_0)$ and x_{infl}, but that the slopes significantly and in a manner difficult–to–predict depend on the parameters

n and n_0. Furthermore, in all cases studied the a-values significantly differ from $a \approx 1$, a value that would be expected on the basis of Eq. (5).

The calculated values for the coefficients b are, in all cases studied, significantly different from zero, again at variance with formula (5).
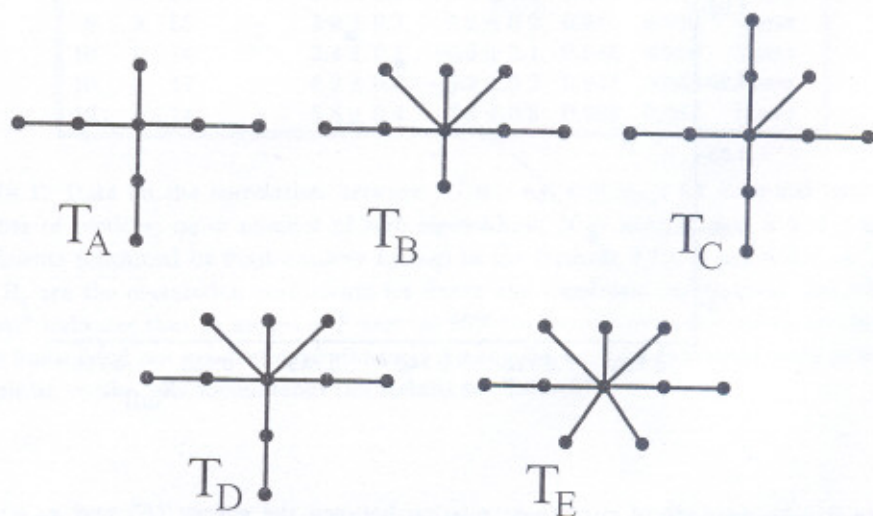

Figure 5. The trees whose (E, x_{infl}) data points significantly deviate from the lines formed by other members of the same cluster; for example see Figure 3; for details see Table 2.

Based on the above presented evidence we conclude that the approximation (5) is valid only to a limited extent. It correctly predicts that $E/(n - n_0)$ is proportional to x_{infl}, although the proportionality found is not always linear. It, however, fails to foresee that the correlation between $E/(n - n_0)$ and x_{infl} is dependent on both n and n_0. In other words, for each particular choice of the parameters n and n_0 one has a distinct correlation line. Some of these correlations are linear, some curvilinear. Some of these correlations are rather good (with correlation coefficients close to unity, $R > 0.98$), some much weaker (but, in all cases, $R > 0.9$).

On the other hand, we have established that there certainly exists some relation between x_{infl} and the graph energy. Thus, a single point of the Coulson function (namely the value of the variable x at which the Coulson function has its inflection point) carries a lot of information on the energy of the respective graph, or – what is the same – on the total

π-electron energy of the respective conjugated molecule. Another parameter that influences the value of x_{infl} is the multiplicity of the number zero in the spectrum of the respective graphs, or – what is the same – the number of non-bonding molecular orbitals of the respective conjugated molecule.

References

[1] I. Gutman, *The Coulson function*, Kragujevac J. Sci. **24** (2002) 65–70, the preceding paper.

[2] C. A. Coulson, *On the calculation of the energy in unsaturated hydrocarbon molecules*, Proc. Cambridge Phil. Soc. **36** (1940) 201–203.

[3] A. Graovac, I. Gutman, N. Trinajstić, *Topological Approach to the Chemistry of Conjugated Molecules*, Springer–Verlag, Berlin, 1977.

[4] I. Gutman, O. E. Polansky, *Mathematical Concepts in Organic Chemistry*, Springer–Verlag, Berlin, 1986.

[5] C. A. Coulson, B. O'Leary, R. B. Mallion, *Hückel Theory for Organic Chemists*, Academic Press, London, 1978.

[6] D. M. Cvetković, I. Gutman, *The algebraic multiplicity of the number zero in the spectrum of a bipartite graph*, Mat. Vesnik (Beograd) **9** (1972) 141–150.

[7] D. Cvetković, I. Gutman, N. Trinajstić, *Graphical studies on the relations between the structure and reactivity of conjugated systems: The role of non-bonding molecular orbitals*, J. Mol. Struct. **28** (1975) 289–303.

[8] I. Gutman, *The energy of a graph*, Ber. Math.-Statist. Sekt. Forschungszentrum Graz **103** (1978) 1–22.

[9] I. Gutman, T. Soldatović, D. Vidović, *The energy of a graph and its size dependence. A Monte Carlo approach*, Chem. Phys. Lett. **297** (1998) 428–432.

[10] J. H. Koolen, V. Moulton, I. Gutman, *Improving the McClelland inequality for total π-electron energy*, Chem. Phys. Lett. **320** (2000) 213–216.

[11] I. Gutman, A. Nikolić, Ž. Tomović, *A concealed property of total π-electron energy*, Chem. Phys. Lett. **349** (2001) 95–98.

[12] I. Gutman, Y. Hou, *Bipartite unicyclic graphs with greatest energy*, Commun. Math. Chem. (MATCH) **43** (2001) 17–28.

[13] I. Gutman, Y. Hou, *Hyperenergetic line graphs*, Commun. Math. Chem. (MATCH) **43** (2001) 29–39.

[14] I. Gutman, D. Vidović, *Quest for molecular graphs with maximal energy: A computer experiment*, J. Chem. Inf. Comput. Sci. **41** (2001) 1002–1005.

[15] H. B. Walikar, H. S. Ramane, P. R. Hampiholi, *On the energy of a graph*, in: R. Balakrishnan, H. M. Mulder, A. Vijayakumar (Eds.), *Graph Connections*, Allied Publishers, New Delhi, 1999, pp. 120–123.

[16] J. Koolen, V. Moulton, *Maximal energy graphs*, Adv. Appl. Math. **26** (2001) 47–52.

[17] I. Gutman, *The energy of a graph: Old and new results*, in: A. Betten, A. Kohnert, R. Laue, A. Wassermann (Eds.), *Algebraic Combinatorics and Applications*, Springer-Verlag, Berlin, 2001, pp. 196–211.

[18] H. B. Walikar, H. S. Ramane, *Energy of some cluster graphs*, Kragujevac J. Sci. **23** (2001) 51–62.

[19] H. B. Walikar, H. S. Ramane, *Energy of some bipartite cluster graphs*, Kragujevac J. Sci. **23** (2001) 63–74.

[20] A. Graovac, I. Gutman, P. E. John, D. Vidović, I. Vlah, *On statistics of graph energy*, Z. Naturforsch. **56a** (2001) 307–311.

[21] H. Fripertinger, I. Gutman, A. Kerber, A. Kohnert, D. Vidović, *The energy of a graph and its size dependence. An improved Monte Carlo approach*, Z. Naturforsch. **56a** (2001) 342–346.

[22] M. Randić, M. Vračko, M. Novič, *Eigenvalues as molecular decriptors*, in: M. V. Diudea (Ed.), *QSPR/QSAR Studies by Molecular Descriptors*, Nova, Huntington, 2001, pp. 147–211.