communications in mathematical
match
and in computer chemistry

no. 44, October 2001
MATCDY (44) 279 - 286 (2001)
ISSN 0340 - 6253

A THEOREM ON GRAPH VALENCE SHELLS

István Lukovits
Department of Surface Science and Corrosion, Chemical Research Center,
Budapest, Hungary

Abstract. Recently Randić [1] has proposed a set of new graph invariants: the graph valence shells. A valence shell $s(u)_k$ is the sum of valences (degrees) of all vertices placed at distance $k - 1$ from vertex u. The graph valence shell s_k, is the sum of valence shells $s(u)_k$. Consider an acyclic graph T. It has been proved that $s_k = p_{k-1} + p_k$, where p_k denotes the number of paths of length k in T.

INTRODUCTION

Recently Randić [1] has proposed a set of new graph invariants: the graph valence shells. A valence shell $s(u)_k$ is the sum of valences (degrees) of the vertices placed at a constant distance $k - 1$ from a given vertex u. Randić has not given an explicit formula for $s(u)_k$, but the following expression is in full agreement with Randić's definition:

$$s(u)_k = \sum_i \delta(k - 1 - d_{u,i})\, v_i \tag{1}$$

where $d_{u,i}$ denotes the distance between vertices u and i, and v_i denotes the valence of a vertex i, δ is a function which is equal to 1 if the argument is zero, and is zero in all other cases. δ picks out a vertex placed at distance $d_{u,i}$ from u. The summation has to be performed for all N vertices in T. Graph valence shell s_k can be obtained by summing expression (1) for all vertices u, and dividing the result by two:

$$s_k = \sum_u s(u)_k/2 \tag{2}$$

The calculation of s_k can be illustrated by using a simple example (FIGURE 1). As an illustration obtain s_5: The valence shell of vertex 1 is 2 (valence of vertex 5). Contribution of vertex 2 is 1 (valence of vertex 6), contribution of vertex 5 is 1 + 1 = 2 (sum of valences of vertices 1 and 7), contribution of vertex 6 is 6 (sum of valences of vertices 2, 8 and 9). The contribution of vertex 7 is 2 (valence of vertex 5) and the contributions of vertices 8 and 9 are 1, each. Therefore $s_5 = (2 + 1 + 2 + 5 + 2 + 2)/2 = 7$. Graph invariants based on 'concentric shells' (or layers) have already been discussed [2-6] and graph valence shells happen to be a subclass of this more general approach.

Graph valence shells s_2, s_3 and s_4 have been shown by Randić to form an appropriate set of indexes that can be used in structure-property relationships. More specifically Randić found that combinations of p_2, p_3, p_4 and s_2, s_3, s_4, respectively, resulted in multiple regression equations with identical statistical parameters. (The boiling points of octanes were considered as the dependent variable.) This result could only have been obtained if p_2, p_3, p_4 and s_2, s_3, s_4 are linear transforms of each other. Actually Randić has observed that

$$p_1 + p_2 = s_2 \tag{3}$$

$$p_2 + p_3 = s_3 \tag{4}$$

$$p_3 + p_4 = s_4 \tag{5}$$

The aim of this paper is to prove that these relationships are valid, and that they are instances of a more general theorem:

Theorem: Let T be a tree and let p_k denote the number of distances of length k, while s_k denotes the k-th graph valence shell in T. Then

$$p_{k-1} + p_k = s_k \tag{6}$$

PROOF OF THE THEOREM

Let's first show that

$$p_1 + p_2 = s_2 \tag{7}$$

because the argumentation will reveal the relationship between s_2 and the Zagreb indices [7]. Vertices contributing to s_2 are the first neighbours of vertex u (eq. 2). Each vertex u has v_u first neighbours. For each first neighbour the contribution to s_2 is equal to v_u, therefore the total contribution of u is v_u^2, and we obtain:

$$s_2 = \sum_u v_u^2/2 \tag{8}$$

Division by 2 has been introduced in accordance with Randić's definition [1]. Eq. (8) indicates the close relationship between s_2 and one of the Zagreb indices [7]. Eq. (8) may be rewritten:

$$s_2 = \sum_u v_u^2/2 = \sum_u v_u(v_u - 1)/2 + \sum_u v_u/2 \tag{9}$$

The second term on the right hand side of eq. (9) is equal to p_1 (the number of edges in T) because of the 'handshake' lemma. The first term is equal to the number of times a path of length two (l_2) can be placed on vertex u. Therefore the first term of eq. (9) is equal to p_2 and $p_1 + p_2 = s_2$.

In order to derive $p_2 + p_3 = s_3$, observe that all second neighbours of u have to be considered now. The second and third columns of Table 1 indicate how many times paths of length two (l_2) and of length three (l_3) can be placed upon the substructures k_{mn} [8,9]. Note that k_{mn} will denote the substructure and the number this structure appears in T. The forth column of Table 1 can be obtained by inspection of the respective substructures k_{mn}. As an exercise consider substructure k_{13}: the endpoint will contribute (two times) to $s(u)_3$, and its contribution is 1 for each neighbour (free valence), therefore the total contribution of k_{13} is equal to 2. As a second example consider k_{22}: Each vertex will contribute 2 to $s(u)_3$, the sum of these contributions is equal to 4. In general the contribution of k_{mn} to $s(u)_3$ is equal to (m-1)n + (n-1)m.

TABLE 1. Substructures k_{mn} and their contributions to $s(u)_3$.
Endpoints are denoted by black dots.

Substructure	Number of times l_2 can be placed upon k_{mn}	Number of times l_3 can be placed upon k_{mn}	Contribution to $s(u)_3$
●—● k_{11}	0	0	0
●—○— k_{12}	1	0	1
●—○< k_{13}	2	0	2
●—○⋲ k_{14}	3	0	3
—○—○— k_{22}	2	1	4
—○—○< k_{23}	3	2	7
—○—○⋲ k_{24}	4	3	10
>○—○< k_{33}	4	4	12
k_{mn}	(m-1)+(n-1)	(m-1)x(n-1)	(m-1)n+(n-1)m

The sum of substructures is equal to the number of edges [8,9]:

$$\sum_{mn} k_{mn} = p_1 \tag{10}$$

The number of paths of length two (p_2) is (Table 1):

$$\sum_{mn} k_{mn}(m + n - 2) = 2p_2 \tag{11}$$

The factor of two on the right hand side appears because each vertex was considered two times in eq. (11). The number of paths of length three (p_3) is (Table 1):

$$\sum_{mn} k_{mn}(m - 1)(n - 1) = p_3 \tag{12}$$

Finally for $s_3$we obtain (Table 1):

$$\sum_{mn} k_{mn}(2mn - m - n) = 2s_3 \tag{13}$$

Combining eqs. (11), and (12) we obtain:

$$2p_2 + 2p_3 = \sum_{mn} k_{mn}[(m + n - 2) + 2(m - 1)(n - 1)] = \sum_{mn} k_{mn}(2mn - m - n) = 2s_3 \tag{14}$$

and from this $p_2 + p_3 = s_3$ follows immediately.

Next prove that $p_{k-1} + p_k = s_k$. In order to obtain e.g. $p_3 + p_4 = s_4$, leave the free valences in TABLE 1 unaltered, and replace the edge with symbol – –, where the square denotes a vertex of any degree. Replace symbols l_2 and l_3 by l_3 and l_4, respectively, and modify the definition of substructure k_{mn}: k_{mn} will now denote the substructure (and the number of such substructures), which was obtained replacing the edge in k_{mn} by – –. Then an analogous argumentation given for the $p_2 + p_3 = s_3$ case can be applied. Similarly if in turn is replaced by – (etc.), the same reasoning can be applied. Therefore we obtain that $p_{k-1} + p_k = s_k$.

The theorem remains valid for p_0 and p_1. Since p_0 (the number of paths of length zero) is equal to zero, and $p_1 = N - 1$ (the number of edges in T), $p_0 + p_1 = s_1$, because of the 'handshake' lemma.

DISCUSSION

Path-counts can easily be determined by inspection of the matrix D composed of the distances $d_{i,j}$. The number of times any number k appears in the upper right hand side of D is equal to p_k. Therefore all values of s_k can be obtained in a more simple manner, than by using eq. (1).

Eq. 6 can be expressed in matrix form:

$$\mathbf{Bp} = \mathbf{s} \tag{15}$$

where **p** is vector containing path counts p_k ($k = 1,\ldots,m$), and m denotes the value of the maximal distance in T. Similarly **s** is a vector consisting values of the individual graph valence shells s_k ($k = 1,\ldots,m$). The structure of matrix **B** is the following in case N = 5:

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Through inversion of **B** we obtain:

$$\mathbf{B}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 1 & 0 \\ 1 & -1 & 1 & -1 & 1 \end{pmatrix}$$

And therefore:

$$\mathbf{p} = \mathbf{B}^{-1}\mathbf{s} \tag{16}$$

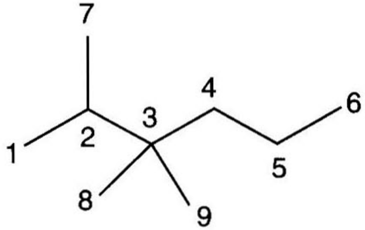

FIGURE 1. Hydrogen suppressed graph of 2,3,3-trimethyl-hexane.

The regular structure of $\mathbf{B}^{-1}$ will not change with other values of N, and we can write:

$$(\mathbf{B}^{-1})_{i,j} = \begin{cases} 0 \text{ if } i < j \\ (-1)^{i+j} \text{ if } i \geq j \end{cases}$$

and therefore expressing path counts in term of graph valence shells:

$$p_1 = s_1 \tag{17}$$

$$p_2 = -s_1 + s_2 \tag{18}$$

$$p_3 = s_1 - s_2 + s_3 \tag{19}$$

$$p_4 = -s_1 + s_2 - s_3 + s_4 \tag{20}$$

$$p_5 = s_1 - s_2 + s_3 - s_4 + s_5 \tag{21}$$

etc.

Several examples of equivalences between topological indices have been reported in the literature [10,11]: it was found that the well-known Wiener index [12] and the Schultz index [13] are related and the resistance distance [14] and the "quasi Wiener index" [15] are also related. Basak et al. [16] after inspecting more than 200 topological indices concluded that the indices could be grouped into various (more or less distinct) classes. Klein and Gutman [17] found that distance based indices are all mutually related and can be expressed in terms of distance distribution moments. Finally we have shown that s_2 is equivalent with one of the Zagreb indices [7].

As a matter of fact it is questionable whether a graph-theoretical index exists, which is *not* related to any other (graph) invariant at all. Skvortsova et al. [18] found several invariants they considered basic. A similar result was obtained by Estrada and Rodriguez [19], who found that the edge-connectivity index does not (practically) depend on other considered types of topological indices.

Path-counts p_{k-1}, p_k and the graph valence shells s_k are linear transforms, therefore, and in contradiction to Randić's conclusion, they *do not* represent independent features. On the other hand, if such (qualitatively different) indices like path counts and graph valence shells are equivalent, then - using Kirby's wording [20] - what does a topology index index?

REFERENCES

[1] M. Randić *J. Chem. Inf. Comput. Sci.*, **2000**, *41*, 000-000.

[2] M. V. Diudea, D. Horvath, D. Bonchev, *Croat. Chem. Acta*, **1995**, *68*, 131-148.

[3] V. A. Skorobogatov, A. A. Dobrynin, *MATCH*, **1988**, *23*, 105-151.

[4] A. T. Balaban, M. V. Diudea, *J. Chem. Inf. Comput. Sci.*, **1993**, *33*, 421-428.

[5] M. V. Diudea, O. M. Minailiuc, A. T. Balaban, *J. Comput. Chem.* **1991**, *12*, 527-535.

[6] M. V. Diudea, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1064-1071.

[7] S. Nikolić, N. Trinajstić, I. M. Tolić, *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 920-926, and references therein.

[8] M. I. Skvortsova, I. I. Baskin, O. L. Slovokhotova, V. A. Palyulin, N. S. Zefirov, *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 630-634.

[9] L. H. Hall, R. S. Dailey, L. B. Kier, *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 598-603.

[10] D. J. Klein, Z. Mihalić, D. Plavšić, N. Trinajstić, *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 304-305.

[11] I. Gutman, B. Mohar, *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 982-985.

[12] H. Wiener, *J. Am. Chem. Soc.* **1947**, *69*, 17-20.

[13] H. P. Schultz, *J. Chem. Inf. Comput. Chem.* **1989**, *29*, 228-228.

[14] D. J. Klein, M. Randić, *J. Math. Chem.* **1993**, *12*, 81-95.

[15] N. Trinajstić; D. Babić, S. Nikolić, D. Plavšić, D. Amić, Z. Mihalić, *J. Chem. Inf. Comput. Sci.*, **1994**, *34*, 368-376.

[16] S. C. Basak, A. T. Balaban, G. D. Grunwald, B. D. Gute, *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 891-898.

[17] D. J. Klein, I. Gutman, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 534-536.

[18] M. I. Skvortsova, I. I. Baskin, L. A. Skvortsov, V. AQ. Palyulin, N. B. Zefirov, I. V. Stankevich, *Theochem – Journal of Molecular Structure* **1999**, *466*, 211-217.

[19] E. Estrada, L. Rodriguez, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 1037-1041.

[20] E. C.Kirby, *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1030-1035.