CROATICA CHEMICA ACTA CCACAA 49 (1) 115—121 (1977)

CCA-982

YU ISSN 0011—1643
539.19:516
Original Scientific Paper

Note on Branching

D. M. Cvetković and I. Gutman***

*Faculty of Electrical Engineering, University of Belgrade, P.O.B. 816, 11001 Beograd, Yugoslavia, Technological University Eindhoven, Department of Mathematics. P.O.B. 513, Eindhoven, The Netherlands
and **Ruđer Bošković Institute, P.O.B. 1016, 41001 Zagreb, Croatia, Yugoslavia

Received September 22, 1976

It has been demonstrated that the spectrum of the molecular graph contains information about the extent of branching of the molecular skeleton. In particular, the largest eigenvalue, x_1, in the spectrum has been shown to be closely related to the total number of walks in the graph (eqs. (11) and (15)). Thus, a justification of the recent empirical finding that x_1 is a measure of branching[9,16] has been obtained.

Many physical and chemical properties (viscosity, surface tension, refractive index, boiling point, density, thermochemical constants *etc.*) of saturated hydrocarbons depend on the extent of branching of the carbon skeleton of the molecule[1]. Similar regularities exist also in homologous series of other compounds, *e. g.* alcohols, fatty acids, alkens *etc.* The recent interest in topological ideas in chemistry[2–4] stimulated the study of branching[5–9]. A number of topological indices have been proposed[1,5,8,9] in order to express branching in a quantitative manner.

In this respect, the fact that the spectrum[4,10] of a graph contains information about branching may be of great importance. Namely, this would allow the application of the mathematical apparatus of graph spectral theory[10] in theoretical study of saturated organic compounds[11].

The first indication of the branching dependence of graph spectral quantities came from Collatz and Sinogowitz[13]. Hosoya[5,14] demonstrated that for acyclic molecules his topological index Z_G is in a simple way related to the characteristic polynomial[15] $P(G, x)$ of the molecular graph G, that is $Z_G = i^N P(G, i)$, where $i = \sqrt{-1}$.

Relations between the graph spectrum and branching were noticed also by Lovász and Pelikán[16]. They concluded that the maximal eigenvalue x_1 of a tree is a measure of branching[17] and that trees can be well ordered according to x_1. This finding was later confirmed by extensive numerical work performed on molecular graphs[9]. In addition recent numerical investigation on cubic graphs[18] (these are regular graphs of degree tree; all cubic graphs have $x_1 = 3$), have shown that the values of the second, third *etc.* eigenvalue of a graph have good ordering properties. However, the mathematical background of these regularities has not yet been understood.

In the present work we would like to offer a result which shows the reasons why and how x_1 is related to the branching of a graph.

We shall use the following notation[4]. The molecular graph G has N vertices $v_1, v_2, \ldots, v_N$ and M edges $e_1, e_2, \ldots, e_M$. Its adjacency matrix $\mathbf{A}$ is defined as

$$\mathbf{A}_{pq} = \begin{cases} 1 & \text{if the vertices } v_p \text{ and} \\ & v_q \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

The degree D_p of the vertex v_p is the number of edges which are incident to v_p. Evidently

$$D_p = \sum_{q=1}^{N} A_{pq} \tag{2}$$

Let D_{max} and D_{min} be the maximal and minimal vertex degrees, respectively, in the graph. In the molecular graphs of saturated hydrocarbons, $D_{max} \leqslant 4$, $D_{min} \geqslant 1$. A graph is called regular of degree D if $D_{max} = D_{min} = D$.

A walk[19] of the length n in a graph is a sequence of n + 1 vertices and n edges, $v_1, e_1, v_2, e_2, \ldots, e_n, v_{n+1}$, such that the edge e_j is incident to the vertices v_j and v_{j+1} ($j = 1, 2, \ldots, n$). This walk is said to connect the vertices v_1 and v_{n+1}. A walk with the properties $v_j \neq v_k$ and $e_j \neq e_k$ for all $j \neq k$ is called a path[19]. Let $W_{pq}(n)$ and $P_{pq}(n)$ be the number of distinct walks and paths, respectively, of the length n, connecting the vertices v_p and v_q.

The total number of paths, $P = \sum_{n} \sum_{p,q} P_{pq}(n)$ in an acyclic molecular graph has been shown to be a quantity which can be excellently correlated with a number of branching-dependent molecular properties[1,7]. Hence, P is a good measure of what one intuitively associates with the notion of branching. Similarly, it should be evident that in a general case, the larger is the branching of a graph, the larger is the value of any $W_{pq}(n)$. Namely, every branch introduced in the graph will result in numerous new possible walks. Therefore, the total number $W(n)$ of walks of the length n in the graph

$$W(n) = \sum_{p=1}^{N} \sum_{q=1}^{N} W_{pq}(n) \tag{3}$$

should also possess convenient properties for being a measure of branching[21]. In the text below we will see that $W(n)$ is closely related to the largest eigenvalue of the graph.

The eigenvalues x_j and the eigenvectors $\mathbf{C}_j = (C_{j1}, C_{j2}, \ldots, C_{jN})$ of the graph G fulfil the relation[4] $\mathbf{C}_j\mathbf{A} = x_j\mathbf{C}_j$ ($j = 1, 2, \ldots, N$). Since $\mathbf{A}$ is a symmetric matrix, its eigenvalues and eigenvectors are real. These eigenvalues form the spectrum of the graph[4,10], and will be labelled in non-increasing order: $x_1 \geqslant x_2 \geqslant \ldots \geqslant x_N$. The ortogonal matrix $\mathbf{C} = (\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_N)^T$ has, hence, the property

$$\mathbf{C}\,\mathbf{A}\,\mathbf{C}^T = \text{diag}\,(x_1, x_2, \ldots, x_N) \tag{4}$$

from which it follows,

$$\mathbf{C}\,\mathbf{A}^n\,\mathbf{C}^T = \text{diag}\,((x_1)^n, (x_2)^n, \ldots, (x_N)^n) \tag{5}$$

$$W_{pq}(n) = (\mathbf{A}^n)_{pq} \tag{6}$$

which combined with eq. (5) yields

$$W_{pq}(n) = \sum_{j=1}^{N} (x_j)^n C_{jp} C_{jq} \tag{7}$$

Finally, from eq. (3) we get

$$W(n) = \sum_{j=1}^{N} (x_j)^n (\Omega_j)^2 \tag{8}$$

where

$$\Omega_j = \sum_{p=1}^{N} C_{jp} \tag{8'}$$

Formula (8) has also been found independently by F. Harary and A. J. Schwenk[23].

It is known that x_1 fulfils the inequality

$$D_{min} \leqslant x_1 \leqslant D_{max} \tag{9}$$

and moreover, if G is a connected graph, the eigenvector $\mathbf{C}_1$ is positive, that is $C_{1j} > 0$ for all $j = 1, 2, \ldots, N$. This is the graph theoretical interpretation[13] of the well known Frobenius theorem[24], which applies to non-negative matrices. In chemistry eq. (9) was first used by Coulson[25]. This inequality can be interpreted as an indication that graph spectra contain information about branching.

Let us first consider regular graphs of degree D. Because of (9), $x_1 = D$. Besides, the eigenvectors of a regular graph have the following interesting property:

$$\Omega_j = \begin{cases} \sqrt{N} & \text{if } j = 1 \\ 0 & \text{if } j > 1 \end{cases} \tag{10}$$

In order to prove the above equation, note that since $x_1 = D$, $\mathbf{C}_1 = (N)^{-1/2}$ $(1, 1, \ldots, 1)$. Therefore, according to eq. (8'),

$$\Omega_1 = \sum_{j=1}^{N} (N)^{-1/2} = \sqrt{N}$$

All other Ω_j 's are equal to zero because of the mutual ortogonality of the graph eigenvectors,

$$\mathbf{C}_1 \mathbf{C}_j^T = 0 \qquad \text{for } j \neq 1$$

Substituting eq. (10) back into (8), one gets a simple expression

$$W(n) = N D^n = N (x_1)^n \tag{11}$$

Another elementary combinatorial derivation of eq. (11) is the following. In regular graphs a walk of length n can start from any of the N vertices and can be continued at each vertex in D different ways.

Eq. (11) can be rewritten also in the form

$$T_n \equiv (W(n)/N)^{1/n} = x_1 \; (= D) \tag{12}$$

For convenience we have defined a topological index T_n

$$T_n = \sqrt[n]{\frac{1}{N} \sum_{j=1}^{N} (x_j)^n (\Omega_j)^2} \tag{13}$$

which can be completely calculated from the eigenvalues and eigenvectors of the graph. Thus, in regular graphs T_n is independent of n.

If G is not regular, eqs. (10)—(12) are no more valid and T_n depends on n. However, from eq. (13),

$$T = \lim_{n \to \infty} T_n = x_1 \tag{14}$$

where we have used the relation

$$\lim_{n \to \infty} \sqrt[n]{\sum_{j=1}^{J} F_j (f_j)^n} = \max \{f_1, f_2, \ldots, f_J\}$$

with F_j (j = 1, 2, . . ., J) being any real numbers such that

$$\sum_{j=1}^{J} F_j (f_j)^n > 0 \text{ for sufficiently large n.}$$

Formula (11), together with the remark made immediately after it, suggest that T_n can be understood as a mean value of vertex degrees with respect to walks of length n. This mean value for long walks (that is for $n \to \infty$), which is equal to x_1, is called the dynamical mean of the vertex degrees[10]. On the other hand, any definition of branching must be (explicitely or implicitely) related to vertex degrees. Intuitively it is clear that a measure of branching must be based on some kind of a mean value of vertex degrees[26]. For example, the measure of branching defined by[9] $Q = \Sigma D_j^2$ is based on the quadratic mean

$$\overline{D} = \sqrt{\frac{1}{N} \sum_{j=1}^{N} D_j^2}$$

and we have $Q = N \overline{D}^2$. In other words, Q is based on the mean value of vertex degrees with respect to walks of length two.

Since x_1 can be interpreted as a certain kind of mean vertex degree, it is not surprising to find correlations between x_1 and any properly defined measure of branching.

If we apply elementary combinatorial reasoning, used after formula (11), to non-regular graphs we immediately get

$$W(n) \approx N (x_1)^n \tag{15}$$

where we have replaced the constant vertex degree D in regular graphs by the dynamical mean x_1 of vertex degrees in the general case.

Of course, we have a better approximation for W (n) if we just take the leading term of (8), namely

$$W(n) \approx (\Omega_1)^2 (x_1)^n \tag{16}$$

but (15) is also sufficiently good at least for molecular graphs. This claim is supported by the following, more or less empirical facts.

In non-regular graphs $C_{1j} \neq (N)^{-1/2}$, and Randić has shown[27] that the largest C_{1j}'s belong usually to the vertices of maximal degree, while the smallest C_{1j}'s belong to terminal vertices and their neighbours. However, also in non-regular graphs $\Omega_1 \approx \sqrt{N}$ and $\Omega_1 \gg \Omega_j$ for $j \neq 1$. Our experience suggests that in medium size molecular graphs eq. (15) is a good approximation for all n except for n = 1, 2 and 3.

In fact, eqs. (8), (11), (15) and (16) show the way in which x_1 is related to the total number of walks in a graph. Thus we have a relation between a spectral (x_1) and a combinatorial ($W(n)$) property of a graph. The earlier found fact[9,16] that the largest eigenvalue of the molecular graph is a measure of branching is now completely understood.

TABLE

Compound	x_1
n-hexane	1.802
2-methyl-pentane	1.902
3-methyl-pentane	1.932
2,3-dimethyl-butane	2.000
2,2-dimethyl-butane	2.074
n-heptane	1.848
2-methyl-hexane	1.932
3-methyl-hexane	1.970
2,4-dimethyl-pentane	2.000
3-ethyl-pentane	2.000
2,3-dimethyl-pentane	2.053
2,2-dimethyl-pentane	2.101
3,3-dimethyl-pentane	2.136
2,2,3-trimethyl-butane	2.175

Table presents a chemical example, namely the x_1 values of all hexanes and heptanes[28]. The ordering of these molecules according to increasing x_1 follows the intuitive notion of branching. The behaviour of higher alkanes is similar[28]. Numerical investigations have shown that simple correlations exist between x_1 and the measured chemical and physical properties of alkanes and other saturated series. These correlations are, however, of similar accuracy as those with the previously proposed[1,5,8,9] topological indices and will not be presented here. Further examples of the applicability of x_1 in chemistry can be found in Ref. 9.

As a final comment we wish to note that the graphs of 3-ethyl-pentane and 2,4-dimethyl-pentane have equal x_1 values (= 2.000). Such cases of coincidence occur more frequently in graphs with larger number of vertices. Moreover, numerous pairs (triplets *etc.*) of isospectral nonisomorphic trees exist if N is large enough[29]. The same holds for molecular graphs[30]. Therefore it must be concluded that x_1 (or even the whole graph spectrum) cannot in all cases differentiate between the branching of two molecular skeletons. Nevertheless, if two molecules have equal x_1, their branching should be expected to be similar.

Acknowledgement. During the preparation of this paper Professor Milan Randić (Ames, Iowa) gave numerous useful comments and suggestions. His help is gratefully acknowledged.

REFERENCES

1. H. Wiener, *J. Amer. Chem. Soc.* **69** (1947) 17; *J. Phys. Chem.* **52** (1948) 425; J. R. Platt, *J. Chem. Phys.* **15** (1947) 419; *J. Phys. Chem.* **56** (1952) 328; E. A. Smolenski, *Zhur. Fiz. Khim.* (U.S.S.R.) **38** (1964) 1288; M. Gordon and G. R. Scantelbury, *Trans. Faraday Soc.* **60** (1964) 604.
2. V. Prelog, *Chem. Britain* **4** (1968) 382; V. I. Sokolov, *Uspekhi Khim.* **42** (1973) 1035.
3. For review on the application of graphs in chemistry see: D. H. Rouvray, *R. I. C. Rev,* **4** (1971) 173; H. Hosoya, *Kagaku no Ryoiki* **26** (1972) 21; R. W. Jotham, *Chem. Soc. Revs.* **2** (1973) 457.
4. For review on the application of graph spectra in chemistry of conjugated compounds see: I. Gutman and N. Trinajstić, *Topics. Curr. Chem.* **42** (1973) 49; *Croat. Chem. Acta* **47** (1975) 507.
5. H. Hosoya, *Bull. Chem. Soc. Japan* **44** (1971) 2332; H. Hosoya, K. Kawasaki, and K. Mizutani, *ibid.* **45** (1972) 3415.
6. M. Gordon and J. W. Kennedy, *J. C. S. Faraday* II, **69** (1973) 484.
7. D. H. Rouvray, *Amer. Sci.* **61** (1973) 729; *MATCH (Mathematical Chemistry)* **1** (1975) 125.
8. M. Randić, *J. Amer. Chem. Soc.* **97** (1975) 6609.
9. I. Gutman, B. Ruščić, N. Trinajstić, and C. F. Wilcox, *J. Chem. Phys.* **62** (1975) 3399.
10. For review on graph spectral theory see: R. J. Wilson, in D. J. A. Welsh and D. R. Woodall (Eds.), *Combinatorics,* IMA, Southend-on-Sea, 1973, pp. 295—321; D. Cvetković, *Univ. Beograd Publ. Elektrotehn. Fak. Ser. Mat. Fiz.* **354—356** (1971) 1; see also Refs. 13 and 28.
11. Note that presently graph spectra are used mainly in discussing of properties of conjugated molecules[4]. Applications of topological ideas to inorganic chemistry were also attempted[12].
12. H. H. Schmidtke, *J. Chem. Phys.* **48** (1968) 970; S. F. A. Kettle and V. Tomlinson, *Theoret. Chim. Acta* **14** (1969) 175; H. Bock and W. Ensslin, *Angew. Chem. Internat. Ed.* **10** (1971) 404; R. W. Rudolph and W. R. Petzer, *Inorg. Chem.* **11** (1972) 1974; I. R. Epstein, *ibid.* **12** (1973) 709.
13. L. Collatz and U. Sinogowitz, *Abh. Math. Sem. Univ. Hamburg* **21** (1957) 63.
14. H. Hosoya, *Theoret. Chim. Acta* **25** (1972) 215.
15. Graph spectral theory investigates both the graph eigenvalues and the characteristic polynomial, since there is a one-one correspondence between them. Therefore all quantities which can be expressed by means of $P(G, x)$ belong to spectral properties of a graph G.
16. L. Lovász and J. Pelikán, *Period. Math. Hung.* **3** (1973) 175.
17. In Ref. 16 is proved that among acyclic graphs, the chain has a minimal and the star a maximal x_1 value (see also Ref. 13). This is a necessary condition which must fulfil every measure of branching[9].
18. F. C. Bussemaker, S. Čobeljić, D. M. Cvetković, and J. J. Seidel, *Technological Univ. Eindhoven, T. H.-Report* 76-WSK-01.
19. This terminology follows that of Ref. 20, p. 13.
20. F. Harary, *Graph Theory,* Addison-Wesley, Reading 1969.
21. The analogy between the number $P(n)$ of paths of the length n and $W(n)$ is, however, not complete. While there exists a number n_0 such that for all $n > n_0$, $P(n) = 0$, the values of $W(n)$ rapidly increase with increasing n. Of course, such a property of $W(n)$'s follows from eq. (15).
22. Theorem 13.1 in Ref. 20, p. 151.
23. F. Harary and A. J. Schwenk, *Pacific J. Math.,* in press.
24. See for example in: M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities,* Allyn and Bacon, Boston 1964, Chapter 3, Theorem. 3.1.1.
25. C. A. Coulson, *Proc. Cambridge Phil. Soc.* **46** (1949) 202.
26. For different definitions of means see: D. S. Mitrinović, *Analytic Inequalities,* Springer-Verlag, Berlin—Heidelberg—New York 1970.
27. M. Randić, *J. Chem. Inf. Comput. Sci.* **15** (1975) 105.
28. Graph spectral data are taken from: E. Nosal, *M. Sc. Thesis,* University of Calgary, Alberty 1970.

29. A. J. Schwenk, in F. Harary (Ed.), *New Directions in the Theory of Graphs*, Academic Press, New York 1973, pp. 275—307.
30. T. Živković, N. Trinajstić, and M. Randić, *Mol. Phys.* **30** (1975) 517; M. Randić, N. Trinajstić, and T. Živković, *J. C. S. Faraday* II, **72** (1970) 244.

SAŽETAK

Jedna napomena o grananju

D. M. Cvetković i I. Gutman

Dokazano je da spektar molekularnog grafa sadržava informaciju o veličini grananja molekularnog skeleta. Pokazano je da je najveća vlastita vrijednost x_1 u spektru molekularnog grafa u bliskoj vezi s ukupnim brojem puteva u tom grafu (jednadžbe (11) i (15)). Na taj je način obrazloženo jedno svojstvo x_1 koje je nedavno nađeno empirijski[9,16], naime da x_1 pretstavlja mjeru grananja.

ELEKTROTEHNIČKI FAKULTET
UNIVERZITETA U BEOGRADU
11000 BEOGRAD
i
INSTITUT »RUĐER BOŠKOVIĆ«
41001 ZAGREB

Prispjelo 22. rujna 1976.