

GRAPH-CENTRE, SELF-RETURNING WALKS, AND CRITICAL
PRESSURE OF ALKANES

Maria Barysz^a, Danail Bonchev^b, and Ovanes Mekenyan^b

^a Institute of Chemistry, Silesian University,
Szkolna 9, 40-006 Katowice, Poland

^b Department of Physical Chemistry, Higher School
of Chemical Technology, Burgas 8010, Bulgaria

(Received: May 1986)

ABSTRACT

A graph-theoretical approach to the problem of hierarchical ordering of isomeric structures has been developed. The key to this approach is the graph centre and its self-returning walks. C₇-C₉-alkane isomers and their critical pressure have been used as illustrative example.

INTRODUCTION

Progress in experimental chemistry and physics has resulted in the accumulation of a large amount of data on molecular properties. Rationalization of these data is of considerable theoretical interest and may facilitate calculation and prediction of molecular properties. For purposes of finding structure-properties correlations, the connectedness (the atom - atom connectivity) or, more general, the topology of a molecule is one of the appro-

aches used frequently¹⁻¹². The topology of each molecular graph is converted into a descriptor which may be a matrix, a polynomial, a sequence of numbers, or a numerical index. Numerical indices developed for this purpose are called topological indices¹³.

Experience indicates that a single topological quantity is usually not sufficient for the description of molecular properties, because it emphasizes a single structural feature. Therefore, when comparing molecular data and topological descriptors, additional criteria are necessary^{5,9,10}. In this paper we propose a procedure for ordering alkanes, based on the self-returning walks of the graphs' central vertices. Next, critical quantities are analysed.

Some Definitions¹⁴

A connected acyclic graph (a tree) T is considered. Let v be a vertex in T and u_i be vertices neighbouring v . The various edges incident with v will then be denoted by $E_i = (v, u_i)$. All edges which belong to paths from v , including E_i as a first edge, form the subgraph $B_i = B(v, E_i)$. We call B_i a branch defined by E_i and v ; E_i is included in B_i . The number of edges in B_i is $n_e(B_i)$ so that

$$n_e(T) = \sum_i n_e(B_i) \quad (1)$$

The number

$$w(v) = \max n_e(B_i) \tag{2}$$

is called the tree weight at v and any branch B_i with

$$n_e(B_i) = w(v) \tag{3}$$

is a weighted branch.

The mass centre m_o is the vertex of minimal weight:

$$w(m_o) = \min w(v)$$

A finite tree T has a single mass centre when

$$w(T) \leq \frac{1}{2} v_e(T) \tag{4}$$

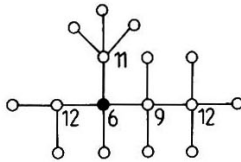
and two neighbouring mass centres when

$$w(T) \leq \frac{1}{2} v_e(T) + 1$$

The latter case occurs only when T has an odd number of edges, hence, an even number of vertices.

The weights at the nonterminal vertices of a tree are presented as an example.

The weight at each terminal vertex is 14, the number of graph edges.



The mass centre is not the sole central point specified for the graph. Another central point¹⁴ is also known which we call the (classical) graph centre. Let c be a graph ver-

Let $r(c)$ denote the maximum distance $d(c,x)$ from c to other graph vertices x :

$$r(c) = \max d(c,x) \quad (5)$$

We call c_0 the (classical) graph centre if $r(c)$ has a minimum for this vertex :

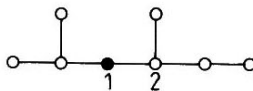
$$r(c_0) = \min r(c) \quad (6)$$

The classical graph centre definition often results in two (for trees) or more (for cyclic graphs) nonequivalent vertices. Consequently, a generalized graph centre concept has recently been proposed^{15,16} and it is specified by a hierarchical series of conditions, which preserve (6) as the first condition. When two or more vertices satisfy the condition (6), the vertex with maximum distance number $d(c)$ is specified as a graph centre :

$$d(c) = \sum_x d(c,x) \quad (7)$$

$$d(c_0) = \min d(c) \quad (8)$$

When two tree vertices have the same minimum distance number, the tree centre (condition (3)) is the one with lesser occurrence of its largest distance, or if the latter are equal, of the next largest distance, etc. The example below illustrates the approach.



Here, vertices 1 and 2 are regarded as equivalent according to the first two criteria, i.e.: $r(1) = r(2) = 3$, $d(1) = d(2) = 13$; the third criterion, however, specifies vertex 1 as single graph centre because the maximum distance ($d_{\max} = 3$) occurs once for the vertex 1 and twice for the vertex 2.

A random walk in a graph is a sequence of edges which can be traversed starting from any vertex and ending in any vertex. Repeated use of the same edge or edges is permitted.

A self-returning walk is a random walk starting and ending at the same vertex.

Self-returning walks may be computed⁴ by the use of the diagonal elements of the first N powers of the graph adjacency matrix, \underline{A} . Each individual diagonal element $(\underline{A}^k)_{ii}$ of the matrix \underline{A}^k can be interpreted as the total number of the self-returning walks of length k proceeding from vertex i . The length of a walk is the total number of the traversed edges. For trees, a simple method for calculating $(\underline{A}^k)_{ii}$ was proposed¹⁷ recently.

A sequence of integers, $(\underline{A}^1)_{ii}, (\underline{A}^2)_{ii}, \dots, (\underline{A}^N)_{ii}$ (with $(\underline{A}^k)_{ii}$ the i th diagonal entry of the k th adjacency matrix power) assigned to each atom in a molecule is called the atomic code⁴. For trees $(\underline{A}^k = \text{odd numbers})_{ii} = 0$ and therefore the atomic code contains only even powers.

THE GRAPH-CENTRE SELF-RETURNING WALKS AND THE
ORDERING OF ISOMERIC MOLECULES

Different approaches to the ordering of isomeric molecules have been proposed^{4,5,10,18-23} including here the use of some topological indices, as well as different molecular or atomic codes.

In our study we represent the molecule by the atomic code of the graph centre as specified in the foregoing. Combining thus the centric properties of isomeric molecules with the detailed description of their shape by means of the self-returning walks we hope to reflect more adequately the topological nature of some molecular properties.

We have examined the atomic codes of both the mass-centre and the generalized graph-centre. For uniqueness of the presentation each molecule must be presented by a single atomic code. The trees, however, often have two central vertices (a bi-centre). This is not an obstacle when the two vertices are equivalent (i.e. when they belong to the same orbit of the graph's automorphism group). The generalized graph centre concept excludes the possibility of arriving at two nonequivalent central vertices. This is, however, not the case with the mass centre. This occurs, for instance, in six out of 18 cases of C_8 -alkanes. In such cases the mass centre with larger $(\underline{A})_m^2$ was chosen following the intuitive idea

that the mass centre should be the root of more branches.
 In case the two mass centres have $(\underline{A}^2)_{m_0} = (\underline{A}^2)_{m'_0}$, then we proceed with the requirement of larger $(\underline{A}^4)_{m_0}$, etc. as, for example, in case of 2-methyl heptane :



Here, $(\underline{A}^2)_{11} = (\underline{A}^2)_{22} = 2$ but $(\underline{A}^4)_{11} = 7 > 6 = (\underline{A}^4)_{22}$.
 Having determined unambiguously the central vertex, we represent the molecular graph by the atomic code of this vertex. Then, the ordering of isomeric structures is straightforward by using numerical priority when two codes are compared digit by digit. Taking into account the diversity in properties of the secondary, tertiary, etc. atoms we classify them in different groups, or regard the isomers with such central atoms as non-comparable. One thus arrives to the condition that the isomeric structure with centre i should precede the one with centre j if

$$\begin{aligned}
 (\underline{A}^2)_{ii} &= (\underline{A}^2)_{jj} \\
 (\underline{A}^4)_{ii} &\leq (\underline{A}^4)_{jj} \\
 &\cdot \\
 &\cdot \\
 &\cdot \\
 (\underline{A}^N)_{ii} &\leq (\underline{A}^N)_{jj}
 \end{aligned}$$

It should be mentioned that the Nth power considered in our study is the largest one generating for the respective n-alkane a matrix still considering entry equal to one.

ORDERING OF C₇-C₉ ALKANES AND THEIR
CRITICAL PRESSURES

The mass centre atomic codes of all isomeric compounds having 7,8, and 9 carbon atoms are presented in Table 1. As seen, the code has a good discriminating power providing different numerical sequences for all 62 compounds examined except one pair of C₉-isomers (compounds 44/45). Clearly, the number of such degenerate codes may increase at higher alkanes. Further, when the generalized graph centre is considered instead of the mass centre, two more degeneracies appear (compounds 12/13 and 32/30).

Table 1. The Isomeric Alkanes C₇ to C₉, Their Mass Centre Atomic Codes and Critical Pressures²⁴

<u>N = 7</u>	Mass centre atomic code	Critical pressure (atm)
1. n heptane	(2,6,20)	27.0
2. 2M hexane	(2,7,26)	26.98
3. 2,4EM pentane	(2,8,32)	27.01

Table 1
(continued)

4. 3M hexane	(3,11,42)	27.77
5. 3E pentane	(3,12,48)	28.53
6. 2,3MM pentane	(3,12,50)	28.70
7. 2,2MM pentane	(4,17,74)	27.37
8. 3,3MM pentane	(4,18,82)	29.07
9. 2,2,3MMM butane	(4,18,84)	29.15
<u>N = 8</u>		
10. n octane	(2,6,20,69)	24.54
11. 2M heptane	(2,7,26,98)	24.52
12. 2,5MM hexane	(2,7,27,106)	24.54
13. 3M heptane	(3,11,42,163)	25.13
14. 4M heptane	(3,11,43,171)	25.09
15. 2,4MM hexane	(3,11,43,173)	25.23
16. 3E hexane	(3,12,49,201)	25.74
17. 2,3MM hexane	(3,12,51,219)	25.94
18. 3,4MM hexane	(3,12,51,221)	26.57
19. 2M,3E pentane	(3,13,57,251)	26.65
20. 2,3,4MM pentane	(3,13,59,269)	26.94
21. 2,2MM hexane	(4,17,74,326)	24.96
22. 2,2,4MMM pentane	(4,17,75,338)	25.34
23. 3,3MM hexane	(4,18,83,385)	26.19
24. 2,2,3MMM pentane	(4,18,85,409)	26.94
25. 3M,3E pentane	(4,19,91,436)	27.71
26. 2,3,3MM pentane	(4,19,93,458)	27.83
27. 2,2,3,3MMM butane	(4,19,97,508)	28.3
<u>N = 9</u>		
28. n nonane	(2,6,20,70)	22.58
29. 2M octane	(2,6,21,78)	22.6

Table 1
(continued)

30.	2,6MM heptane	(2,6,22,86)	22.7
31.	3M octane	(2,7,27,107)	23.1
32.	2,5MM heptane	(2,7,28,115)	23.2
33.	3,5MM heptane	(2,8,34,146)	23.7
34.	2,2 MM heptane	(2,8,35,156)	22.9
35.	2,2,5MMM hexane	(2,8,36,164)	23.0
36.	2,2,4MMM hexane	(2,9,42,197)	23.5
37.	2,2,4,4MMMM pentane	(2,10,50,250)	24.5
38.	4M octane	(3,11,43,172)	23.1
39.	2,4MM heptane	(3,11,44,182)	23.1
40.	3E heptane	(3,12,49,202)	23.7
41.	4E heptane	(3,12,50,210)	23.6
42.	2M,4E hexane	(3,12,50,212)	23.7
43.	2,3MM heptane	(3,12,51,220)	23.7
44.	2,3,5MMM hexane	(3,12,52,230)	23.7
45.	3,4MM heptane	(3,12,52,230)	24.3
46.	2M,3E hexane	(3,13,58,260)	24.2
47.	3M,4E hexane	(3,13,58,262)	24.8
48.	2,3,4MMM hexane	(3,13,60,280)	24.9
49.	2,2,3MMM hexane	(3,13,62,302)	24.6
50.	2,4MM,3E pentane	(3,14,66,312)	24.9
51.	2,2MM,3E pentane	(3,14,68,336)	25.2
52.	2,2,3,4MMMM pentane	(3,14,70,354)	25.75
53.	3,3MM heptane	(4,18,83,386)	24.0
54.	4,4MM heptane	(4,18,84,396)	24.0
55.	2,4,4MMM hexane	(4,18,84,398)	24.0
56.	3M,3E hexane	(4,19,92,447)	25.2
57.	2,3,3MMM hexane	(4,19,94,469)	25.2
58.	3,3,4MMM hexane	(4,19,94,471)	25.9

Table 1
(continued)

59. 3,3EE pentane	(4,20,100,500)	26.4
60. 2,3MM,3E pentane	(4,20,102,522)	26.5
61. 2,3,3,4MMMM pentane	(4,20,104,544)	26.8
62. 2,2,3,3MMMM pentane	(4,20,106,572)	27.05

Inspection of Table 1 shows a close parallelism between the ordering of the isomeric alkanes according to their mass centre atomic codes and critical pressures^{5,23}. Three groups of isomers (A,B, and C, respectively) are formed for each of the C₇, C₈, and C₉ alkanes; i.e., isomers having secondary, tertiary, and quaternary atoms as mass centres.

For C₇-isomers only one minor irregularity in the ordering of the critical pressure occurs with compounds 1 and 2. For isomers with eight carbon atoms the ordering of the compounds with quaternary central atom displays full parallelism with their critical pressures while for those with tertiary and secondary central atoms only one minor disagreement occurs with $\Delta P_{cr} = 0.04$ and 0.02 atm, respectively. Even for the largest class of 35 C₉-isomers the agreement is complete for group C, three discrepancies occur for isomers of group B, and a single but strong displacement is found for 3,5MM heptane which should be placed three positions down the group A scale.

It is of interest to check which ordering provides better agreement with the critical pressures: that with atomic codes based on the mass centre or on the generalized graph centre. Actually, the two graph centres do not coincide for only about 1/6 of the compounds examined (compounds 7; 11, 13, 15, 21, 22, 24; 34, 40, 43, 53). Interestingly, the six discrepancies for C₈-isomers occur only for the compounds having two non-equivalent mass centres and they would disappear if the opposite choice of a single mass centre has been made ($(\underline{A}^2)_{m_0} < (\underline{A}^2)_{m_0'}$, and if $(\underline{A}^2)_{m_0} = (\underline{A}^2)_{m_0'}$), $(\underline{A}^4)_{m_0} < (\underline{A}^4)_{m_0'}$). The resulting ordering of the 60 compounds examined is given below (the isomers of groups A, B and C are divided by semicolons):

C₇: 1, 2, 3, 7; 4, 5, 6; 8, 9

C₈: 1, 2, 3, 4, 6, 12, 13; 5, 7, 8, 9, 10, 11, 15; 14, 16, 17, 18

C₉: 28, 29, 30, 34, 31, 32, 40, 43, 33, 35, 53, 36, 37; 38, 39, 41, 42, 44, 45, 46, 47, 48, 49, 50, 51, 52; 54, 55, 56, 57, 58, 59, 60, 61, 62

Comparing these sequences with the critical pressure values from Table 1, one concludes that the isomers with a quaternary central atom preserve their ordering, those with a tertiary central atom slightly improve it while for the enlarged groups A some more discrepancies appear. Further, the overlap between the values of the critical pressure of these three groups of isomeric molecules is now diminished.

Hence, we conclude that the self-returning atomic codes of the mass centres and generalized graph centres do not differ essentially in ordering the critical pressures of C_7 to C_9 -alkanes. A more diverse^{of} molecular properties should be examined in order to make such a choice possible.

REFERENCES

1. A.T.Balaban, Ed., Chemical Applications of Graph Theory. Academic Press, London, 1976.
2. L.B.Kier and L.H.Hall, Molecular Connectivity in Chemistry and Drug Research, Medicinal Chemistry 14, Academic Press, New York, 1976.
3. D.Bonchev and N.Trinajstić, J. Chem. Phys., 67, 4517 (1977)
4. M.Randić, Chem. Phys. Lett., 55, 5217, (1978); Int. J. Quantum Chem., 5, 245, (1978); Math. Chem.(Mülheim/Ruhr) 7, 5, (1979); J.Comput. Chem. 1, 386 (1980).
5. M.Randić and C.L.Wilkins, J. Chem. Inf. Comp. Sci., 19, 31 (1979); Chem.Phys.Lett., 63, 332 (1979); J. Phys. Chem., 83, 1525 (1979); Int. J. Quantum Chem., 18, 1005 (1980).
6. D.Bonchev, Ov.Mekenyan, G.Protić and N. Trinajstić, J. Chromatography, 176, 149 (1979).
7. A.T.Balaban, A.Chiriac, I.Motoc and Z.Simon, Steric

- Fit in QSAR, Lecture Notes in Chemistry (Springer, Berlin, 1980), vol. 15.
8. O.Mekenyan, D.Bonchev and N.Trinajstić, Int. J. Quantum Chem. 13, 369 (1980).
 9. M.Randić, N.Trinajstić, Math. Chem. (Mulheim/Ruhr), 13, 271 (1982).
 10. D.Bonchev and O.Mekenyan, Chem. Phys. Lett., 98, 134 (1983); J.Chem. Soc. Faraday Trans. 2, 80, 695 (1984).
 11. N.Trinajstić, Chemical Graph Theory, vol.I, II. CRC Press, Boca Raton, 1983.
 12. D.Bonchev, Information-Theoretic Indices for Characterization of Chemical Structures, John Wiley, Chichester, 1983.
 13. A.T.Balaban, I.Motoc, D.Bonchev, and O.Mekenyan, in Steric Effects in Drug Design (M.Charton and I.Motoc, Ed.). Top. Curr. Chem. 114, Springer, Berlin, 1983, p. 21-56.
 14. O.Ore, Theory of Graphs, American Math. Society Colloquium Publications, Vol. XXXVIII, American Math. Soc., Rhode Island, 1962.
 15. D.Bonchev, A.T.Balaban, O.Mekenyan, J.Chem.Inf.Comp. Sci. 20, 106 (1980).
 16. D.Bonchev, A.T.Balaban, and M.Randić, Intern. J. Quantum Chem. 19, 61 (1981)

17. M.Barysz, N.Trinajstić, Int. J. Quantum Chem., Symposium, 18, 661 (1984)
18. G.H.Hardy, J.E.Littlewood and G.Polya, Inequalities, Cambridge University Press, London, 1934, p.45.
19. Kamarata, Publ. Math. de l'Université de Belgrade, 1, 145, 1932.
20. H.Hosoya, Bull. Chem. Soc. Japan, 44, 2332 (1971).
21. D.Bonchev, O.Mekenyan, and N.Trinajstić, J.Comput. Chem. 2, 127 (1981).
22. M.Gordon and J.W.Kennedy, J.Chem. Soc., Faraday Trans. 2, 69, 484 (1973).
23. M.Randić, G.M.Brissey, R.B.Spencer and C.L.Wilkins, Computers & Chemistry, 4, 27 (1980).
24. Selected Values of Properties of Hydrocarbons and Related Compounds, American Petroleum Institute Research Project 44, Texas, 1969.