

Topological Organic Chemistry. 2. Graph Theory, Matrix Determinants and Eigenvalues, and Topological Indices of Alkanes¹

HARRY P. SCHULTZ*

Department of Chemistry, University of Miami, Coral Gables, Florida 33124

EMILY B. SCHULTZ† and TOR P. SCHULTZ‡

School of Forest Resources, Forest Products Laboratory, and Department of Chemistry, Mississippi State University, Mississippi State, Mississippi 39762

Received June 22, 1989

Determinants of the adjacency, distance, and adjacency-plus-distance matrices that describe the structure of an alkane are calculated. The determinant of the adjacency-plus-distance matrix appears to be useful as a unique molecular topological index for each different alkane. The corresponding principal eigenvalues developed for each of the matrices also appear to be unique, single-sum topological indices.

INTRODUCTION

Graph theory offers the means to numerically characterize chemical structures. Wiener,² Hosoya,³ and Randić⁴ have devised various ways of attaining this end. Balaban,⁵ Randić,⁶ and Hansen and Jurs⁷ have reviewed methods for computing molecular topological indices. This paper continues the study of the utility of structure-descriptive matrices of alkanes as sources of single-sum numbers descriptive of alkane structures.

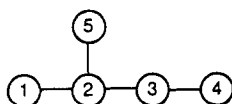
The adjacency and distance matrices that describe the structures of molecular graphs are square matrices. Every square matrix has associated with it a scalar, its determinant, which serves as its absolute single-sum value and which is probably the most fundamental mathematical property of the matrix. Some investigators⁸ have derived the eigenvalues of the adjacency matrices of alkanes, values not unique to the structures. Trinajstić⁹ has described the determinants of adjacency matrices. The determinants of the adjacency (A) and distance (D) matrices, and of their matrix (A + D) sums, are reported in this paper for a series of representative alkanes.

Concomitant necessary utilization of the corresponding matrix eigenvalues presented the opportunity to also assess the significance and utility of the principal eigenvalues of the D and A + D matrices as unique, single-sum molecular topological indices.

COMPUTATIONS

Molecular graphs, hydrogen-suppressed and with the interatomic carbon-carbon bond distances set at unity, were derived from the alkanes listed in Table I. For convenience in relating graphs to molecular structures, the IUPAC system of nomenclature and numbering was utilized, with the α carbon atoms of the side chains deriving priorities from their attachments to the main-sequence carbon atoms and with β carbon atoms, if any, deriving priorities from their respective α carbon atoms. By use of 2-methylbutane (7) as an example, the molecular graph, adjacency and distance matrices, matrix sum of the two matrices, eigenvalues, and determinants are illustrated.

Structural (Molecular) Graph.



*School of Forest Resources.

†Forest Products Laboratory and Department of Chemistry.

Matrices.

$$\begin{matrix} & \text{A} & & & & & \text{D} & & & & \text{A + D} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} & + & \begin{bmatrix} 0 & 1 & 2 & 3 & 2 \\ 1 & 0 & 1 & 2 & 1 \\ 2 & 1 & 0 & 1 & 2 \\ 3 & 2 & 1 & 0 & 3 \\ 2 & 1 & 2 & 3 & 0 \end{bmatrix} & = & \begin{bmatrix} 0 & 2 & 2 & 3 & 2 \\ 2 & 0 & 2 & 2 & 2 \\ 2 & 2 & 0 & 2 & 2 \\ 3 & 2 & 2 & 0 & 3 \\ 2 & 2 & 2 & 3 & 0 \end{bmatrix} \end{matrix}$$

Eigenvalues.

(A)	(D)	(A + D)
-	$\lambda_1 = 7.45929$	$\lambda_1 = 8.85345$
	$\lambda_2 = -3.86269$	$\lambda_2 = -3.38547$
	$\lambda_3 = -2.00000$	$\lambda_3 = -2.00000$
	$\lambda_4 = -1.08462$	$\lambda_4 = -2.00000$
	$\lambda_5 = -0.51198$	$\lambda_5 = -1.46798$

Principal Eigenvalues.

(A)	(D)	(A + D)
-	$\lambda_1 = 7.45929$	$\lambda_1 = 8.85345$

Determinants.

$$|A| = 0 \quad |D| = 32 \quad |A + D| = 176$$

The matrix determinants were computed on a Data General MV/7800XP minicomputer using the MINITAB software package.¹⁰ The alternation of the |D| and |A + D| determinant signs from plus to minus as the odd-numbered carbon atom families of isomers were succeeded by even-numbered carbon atom isomers (vide infra) was confirmed by manually calculating the determinants for matrices up through $n = 5$ by using the minor-cofactor¹¹ and the Gauss-Jordan triangularization¹² procedures, both cumbersome matrix manipulations. Because the program's matrix eigenvalues from which each determinant was calculated were obtained to six significant figures—four decimal places—the last digit for the very large determinants of D and A + D were not firm for compounds 40-42. Note that the determinant |D| values of compounds 40-42, all of which should be the same, differ by one unit (114 687-114 688).

At this point of the study it became apparent that the principal eigenvalues were themselves unique and single sum and that their sequences were monotonic. Table I lists the principal eigenvalues for the D and A + D matrices. Both sets of principal eigenvalues appear capable of functioning as single-sum molecular topological indices.

Table I. D and A + D Determinant-Molecular and Principal Eigenvalue-Molecular Topological Indices and Boiling Points for Representative C₁-C₁₅ Alkanes

no.	compound	D	A + D	obsr bp, °C	calc bp, ^a °C (A + D)	λ_1		calc bp, ^b °C (λ_1 , A + D)
						D	A + D	
1	methane	0	0	-161		0	0	
2	ethane	-1	-4	-88.6	-108.8	1	2	-119.3
3	propane	4	16	-42	-47.4	2.73205	4	-48.6
4	butane	-12	-60	-0.5	-2.3	5.16228	6.53113	1.4
5	2-methylpropane	-12	-48	-11.7	-9.4	4.64575	6	-7.3
6	pentane	32	224	36.1	35.8	8.28822	9.72970	42.0
7	2-methylbutane	32	176	27.9	29.2	7.45929	8.85345	32.4
8	2,2-dimethylpropane	32	128	9.5	20.2	6.60555	8	22.1
9	hexane	-80	-817	68.7	68.9	12.1093	13.6116	76.3
10	2-methylpentane	-80	-656	60.3	63.5	11.0588	12.5154	67.7
11	3-methylpentane	-80	-644	63.2	63.1	10.7424	12.1719	64.9
12	2,3-dimethylbutane	-80	-512	58	57.3	10	11.4031	58.2
13	2,2-dimethylbutane	-80	-464	49.7	54.8	9.67021	11.0769	55.3
14	heptane	192	2840	98.4	97.9	16.6254	18.1783	105.8
15	2-methylhexane	192	2372	90	93.8	15.4048	16.9205	98.5
16	3-ethylpentane	192	2352	93.5	93.6	14.2969	15.7492	91.2
17	3-methylhexane	192	2348	92	93.6	14.8636	16.3497	95.0
18	2,4-dimethylpentane	192	1920	80.5	89.0	14.1760	15.6472	90.5
19	2,3-dimethylpentane	192	1872	89.8	88.4	13.6346	15.0731	86.7
20	2,2-dimethylpentane	192	1728	79.2	86.6	13.6353	15.0970	86.9
21	3,3-dimethylpentane	192	1680	86	86.0	13.0698	14.4942	82.7
22	2,2,3-trimethylbutane	192	1344	80.9	80.8	12.3945	13.8023	77.7
23	octane	-448	-9486	125.7	123.8	21.8363	23.4316	131.7
24	4-methylheptane	-448	-8160	117.7	120.7	19.5420	21.0738	120.9
25	3-methylheptane	-448	-8127	119	120.6	19.7628	21.3026	122.0
26	2,5-dimethylhexane	-448	-6864	109.1	117.0	19.1115	20.6428	118.8
27	3-ethyl-2-methylpentane	-448	-6832	115.7	116.9	17.4187	18.8774	109.6
28	2,4-dimethylhexane	-448	-6816	109.5	116.9	18.3964	19.8985	115
29	3,4-dimethylhexane	-448	-6700	117.7	116.5	17.6759	19.1471	111.1
30	2,2-dimethylhexane	-448	-6220	106.9	115.0	18.4133	19.9327	115.2
31	3,3-dimethylhexane	-448	-6124	112.0	114.6	17.4426	18.9193	109.9
32	3-ethyl-3-methylpentane	-448	-6076	119	114.5	16.6705	18.1096	105.4
33	2,2,3,3-tetramethylbutane	-448	-3520	106.3	102.6	14.9373	16.3459	95.0
34	nonane	1024	31007	150.8	147.7	27.7422	29.3729	154.7
35	decane	-2304	-100095	174.1	170.0	34.3429	36.0037	175.5
36	undecane	5120	320048	196.8	191.1	41.6384	43.3249	194.4
37	dodecane	-11264	-1014203	216.3	211.1	49.6287	51.3375	211.7
38	tridecane	24576	3188230	235.4	230.0	58.3139	60.0420	227.7
39	tetradecane	-53248	-9954933	253.7	248.2	67.6939	69.4390	242.5
40	pentadecane	114688	30910288	270.6	265.6	77.7687	79.5288	256.3
41	2,3-dimethyl-6-ethyl-5-isopropyloctane ^c	114688	17734496		257.1	51.9240	53.5129	(215.9)
42	2,6-dimethyl-5-ethyl-3-isopropyloctane ^c	114687	16622076		256.1	52.1361	53.7266	(216.3)

^aCalculated by using: bp, °C = -256.96 + 190.94(log |A + D|)^{1/2}. ^bCalculated by using: bp, °C = -190.010 + 234.848(log λ_1 , A + D).
^cSuggested by Hansen and Jurs.⁷

SUMMARY

The A determinants were not unique; indeed most, not unexpectedly, had values of zero, as discussed in detail by Trinajstić.⁹ The few exceptions were displayed by compounds 4, 23, 25, 29, 32, and 37 with |A| = +1 and by |A| = -1 for compounds 2, 9, 11, 35, and 39. Therefore, no further attention was given to the A determinants and eigenvalues.

The D determinants for each family of isomers of the same carbon atom content possessed the same values, alternating from negative values for all isomers of even carbon atom content to positive values for all isomers of odd carbon atom content. Most importantly, the A + D determinant of each alkane in Table I was a unique, single-number scalar, part of an essentially monotonic series, negative for substances of even carbon atom content and positive for those of odd numbers of carbon atoms in the molecule. The A + D determinants were very large for alkanes of higher carbon atom content. The normal alkane parent of each family of isomers always possessed the highest absolute A + D determinant value, with the more highly substituted members of a given isomer family having smaller values, lowest for what appears intuitively to be the most highly substituted isomer of a family. But note the slight differences between the A + D determinants for 16

and 17 and for 26-28. Perhaps the derived topological index may be viewed as a description of molecular complexity. If defined in this fashion, 3-methylhexane is slightly more complex than, for example, 3-ethylpentane. The data demonstrate the discrimination attained by using the A + D determinants as molecular topological index numbers of alkanes.

It is also interesting to observe that a plot of the square root of the log of the absolute values for the A + D determinants versus the boiling points of the alkanes enumerated in Table I gave a straight line expressed by

$$\text{bp, } ^\circ\text{C} = -256.96 + 190.94(\log |A + D|)^{1/2}$$

with an r^2 value of 0.994 (Figure 1). The calculated boiling points derived from the above equation are also listed for all substances, excluding methane, in Table I, which also compares the principal eigenvalues, emphasizing the surprising result that the principal eigenvalues for both D and A + D matrices were unique and single-sum and that each series was monotonic; moreover, the figures obtained for even the higher molecular weight molecules were of reasonable dimension and always of a positive value, unlike the corresponding D and A + D determinants. It is pleasing to observe that there was apparently sufficient dissimilitude inherent in the D matrix

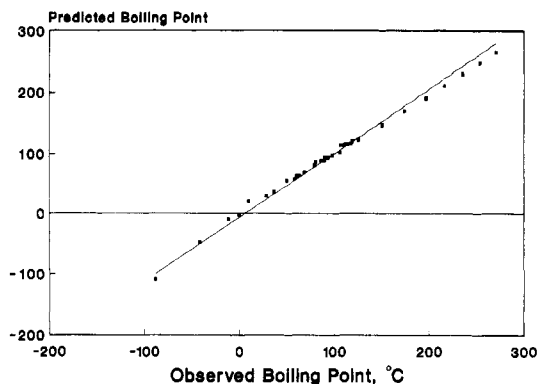


Figure 1. Observed versus predicted boiling points derived from the square roots of the log (A + D) determinants of the alkanes listed in Table I, methane excluded.

alone to provide a unique principal eigenvalue as a molecular topological index—but only barely so in the instance of alkanes 19 and 20, whose principal D eigenvalues differed by only 0.0007 unit but whose A + D eigenvalues varied by more than 0.02 unit.

The D principal eigenvalues versus boiling points of the alkanes, methane excluded, shown in Table I, resulted in a linear relationship expressed by

$$\text{bp, } ^\circ\text{C} = -142.044 + 204.401 \log (\text{main eigenvalue D})$$

with an r^2 value of 0.969. The same comparison between the A + D principal eigenvalues and boiling points, methane excluded, gave

$$\text{bp, } ^\circ\text{C} = -190.010 + 234.848 \log (\text{main eigenvalue A + D})$$

with an r^2 value of 0.987. Table I also lists the calculated alkane boiling points, methane excluded, obtained from application of this latter equation.

Hydrophobic constants (log P) of organic compounds have been shown to be very useful for the correlation of biological and medical activities with chemical structures, as discussed by Fujita et al.¹³ and Hansch and Lien.¹⁴ It is of interest and utility to observe that the plot of the square roots of the principal eigenvalues of the A + D matrices (methane omitted) versus the "new log P values (octanol-water)" for seven lower alkanes (1–6 and 8) reported by Leo et al.¹⁵ gave an excellent linear fit (Figure 2) expressed by

$$\log P = 0.50665 + 0.92435(\lambda_1(\text{A + D}))^{1/2}$$

with an r^2 value of 0.999.

It is interesting to compare the relative intricacy of the alkanes based upon determinants and shown by principal eigenvalues. For a given family of isomers of the same carbon atom content, both the determinants and the principal eigenvalues decreased in magnitude as the molecules became more branched, with the straight-chain parent isomer always having the highest index number. But such pairs as 16 and

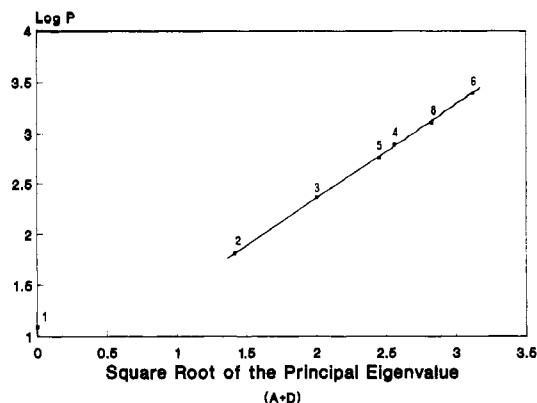


Figure 2. Square roots of the A + D principal eigenvalues versus hydrophobicity (log P). The numbers identify specific compounds in Table I, methane excluded.

17, 19 and 20, 24 and 25, and 28 and 29, to cite obvious comparisons, have different orders of intricacy when compared by the two systems of arriving at a molecular topological index. A further comparison with the earlier published list¹ of the same alkanes (with, however, a different numbering order consistent with the data presented in that table) whose topological indices were calculated from the $v(\text{A + D})$ matrix values, again showed the same pairs of compounds differing slightly in intricacy comparisons.

CONCLUSIONS

The determinants and principal eigenvalues of the structure-descriptive matrices of alkanes are unique and can serve as single-sum molecular topological indices. An appropriate computer program is necessary to acquire the requisite determinants and eigenvalues.

REFERENCES AND NOTES

- (1) Paper 1 of this series: Schultz, H. P. *J. Chem. Inf. Comput. Sci.* **1989**, 29, 227–228.
- (2) Wiener, H. *J. Am. Chem. Soc.* **1947**, 69, 17–20.
- (3) Hosoya, H. *Bull. Chem. Soc. Jpn.* **1971**, 44, 2332–2339.
- (4) Randić, M. *J. Am. Chem. Soc.* **1975**, 97, 6609–6615.
- (5) Balaban, A. T. *J. Chem. Inf. Comput. Sci.* **1985**, 25, 334–343.
- (6) Randić, M. *J. Chem. Inf. Comput. Sci.* **1988**, 28, 142–147.
- (7) Hansen, P. J.; Jurs, P. C. *J. Chem. Educ.* **1988**, 65, 574–580.
- (8) Herndon, W. *Tetrahedron Lett.* **1974**, 8, 671–672.
- (9) Trinajstić, N. *Chemical Graph Theory*; CRC: Boca Raton, FL, 1983; Vol. 1, p 37.
- (10) Minitab, Inc., 3081 Enterprise Dr., State College, PA 16801; 1986. MINITAB is a registered trademark of Minitab, Inc.
- (11) Campbell, H. G. *An Introduction to Matrices, Vectors, and Linear Programming*, 3rd ed.; Prentice-Hall: Englewood Cliffs, NJ, 1977; p 86.
- (12) Painter, R. J.; Yantis, R. P. *Elementary Matrix Algebra with Linear Programming*; Prindle, Weber and Schmidt: Boston, MA, 1977; p 139.
- (13) Fujita, T.; Iwasa, J.; Hansch, C. *J. Am. Chem. Soc.* **1964**, 86, 5175–5180.
- (14) Hansch, C.; Lien, C. *J. Med. Chem.* **1971**, 14, 653–670.
- (15) Leo, A.; Jow, P. Y. C.; Silipo, C.; Hansch, C. *J. Med. Chem.* **1975**, 18, 865–868.