match no.1 pp. 125-134 1975

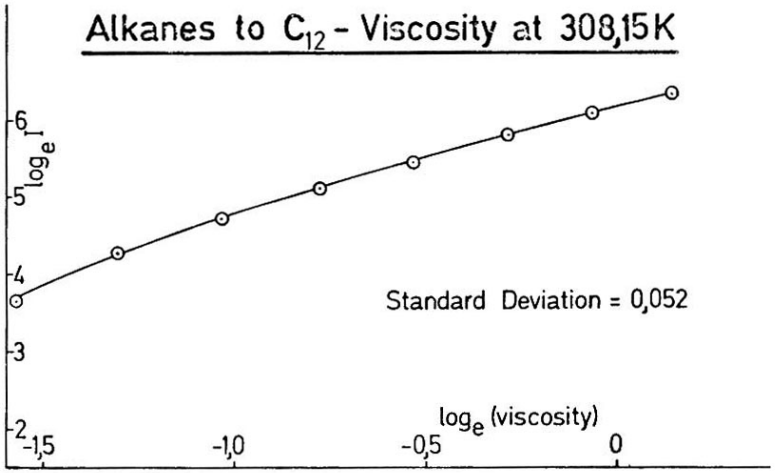
# THE VALUE OF TOPOLOGICAL INDICES IN CHEMISTRY

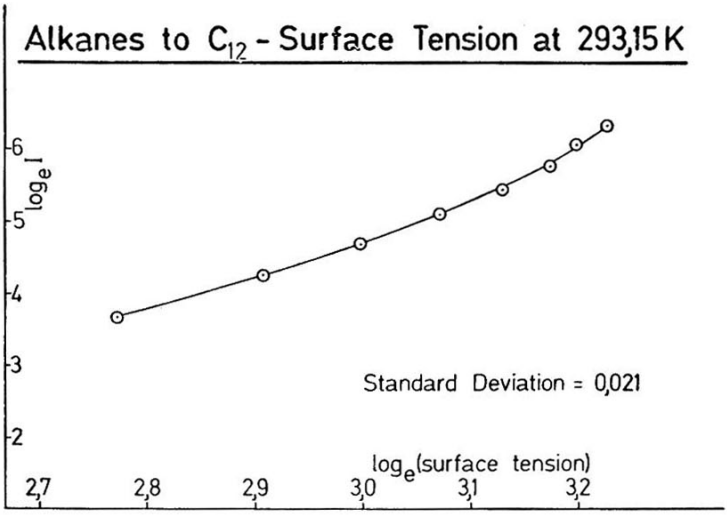
D. H. Rouvray · Chemistry Department, University of the Witwatersrand, Johannesburg, South Africa

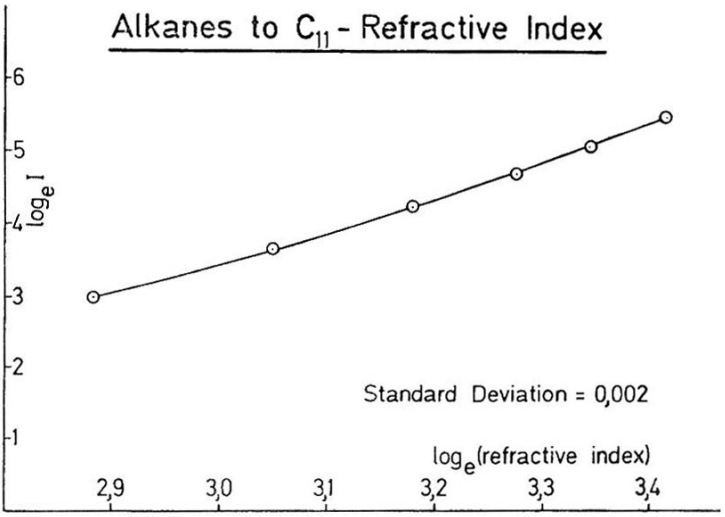
In recent years it has become increasingly apparent that many, if not all, physical and chemical properties of fluids are dependent upon and often largely determined by the topological structure of their component chemical species. In view of this development, it would thus not be unreasonable to describe chemistry today as an essentially topological subject[1]. In this paper, we briefly review some of the principal areas of activity which have led to this new situation. We start by tracing these areas back to their origins.

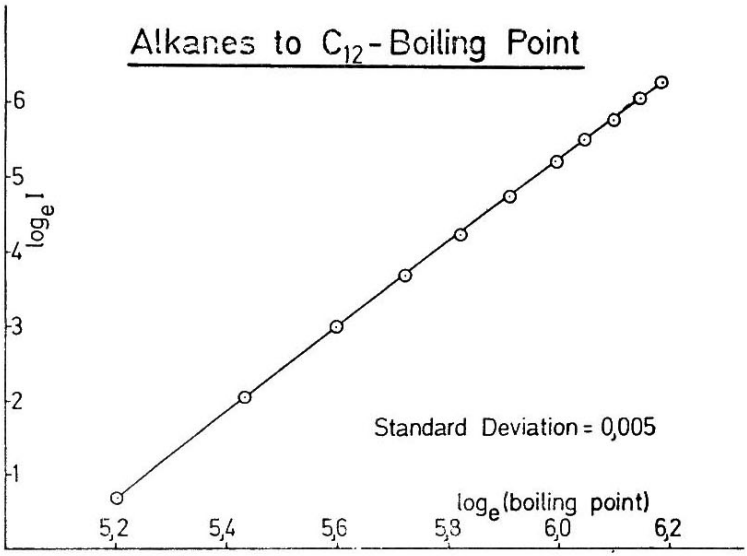
The earliest work along these lines, was that involving studies of the additive properties of compounds, initiated during the first half of the nineteenth century. As an example, we cite the work of Kopp[2], who demonstrated that the thermodynamic parameters of compounds such as boiling point and specific heat were roughly additive in the sense that an approximately constant difference in the parameter was observed for compounds differing by a given chemical grouping, such as a methylene unit. He was also the first person to express his results in the form of algebraic equations. During the succeeding decades, Kopp's researches were extended by a variety of workers. One of these, Markownikov, showed that the melting point of a branched hydrocarbon is higher than that for the corresponding straight chain species[3].

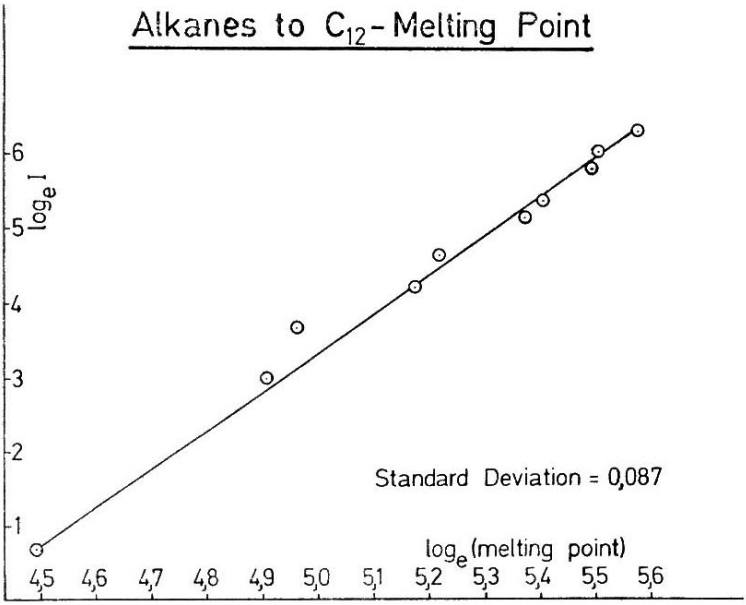
In the present century, Hückel developed his now famous procedures for obtaining appproximate answers to the Schrödinger equation for hydrocarbon species[4]. Extensions to his work, by investigators such as Coulson[5], led to the introduction of several new quantum chemical parameters characteristic of hydrocarbon species. These included the π-electronic charge density, the bond order, and the polarizability. In the early 1950's it was demonstrated by Ruedenberg[6] that all such parameters were dependent upon the molecular topology. Shortly thereafter, the correspondence between the Hückel matrix and what is now referred to as the topological matrix was pointed out by

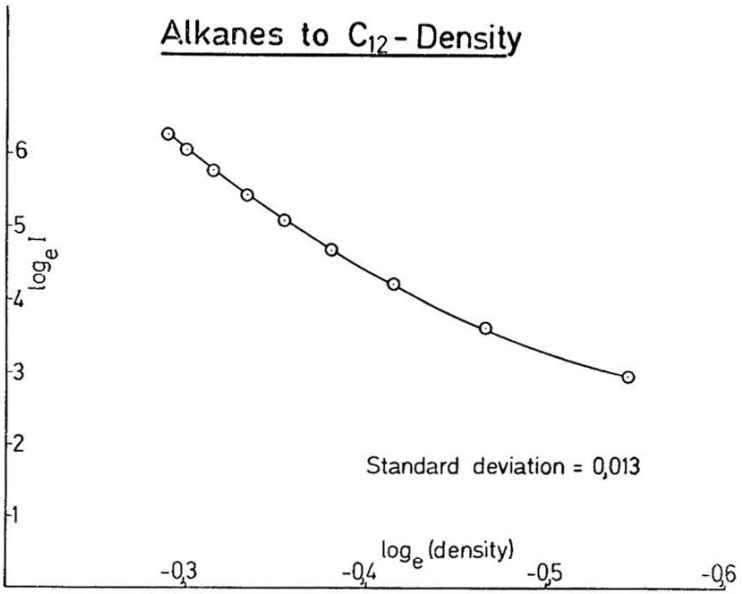


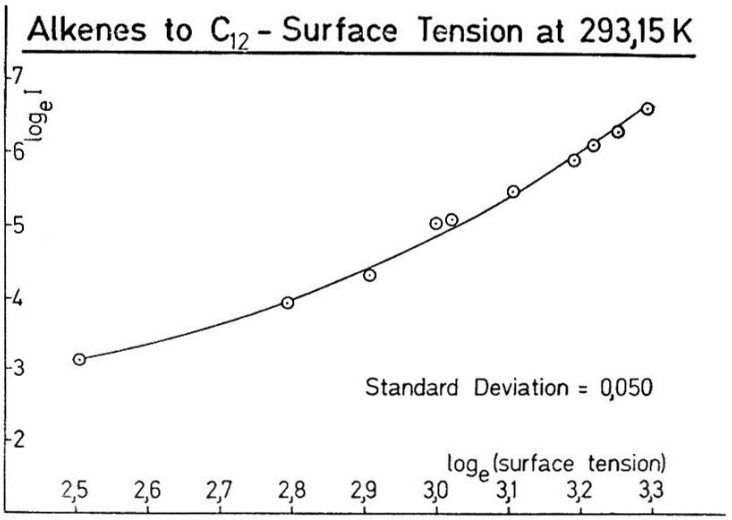

Alkanes to $C_{12}$ - Surface Tension at 293,15 K
$\log_e I$
6
5
4
3
2
Standard Deviation = 0,021
$\log_e$(surface tension)
2,7
2,8
2,9
3,0
3,1
3,2

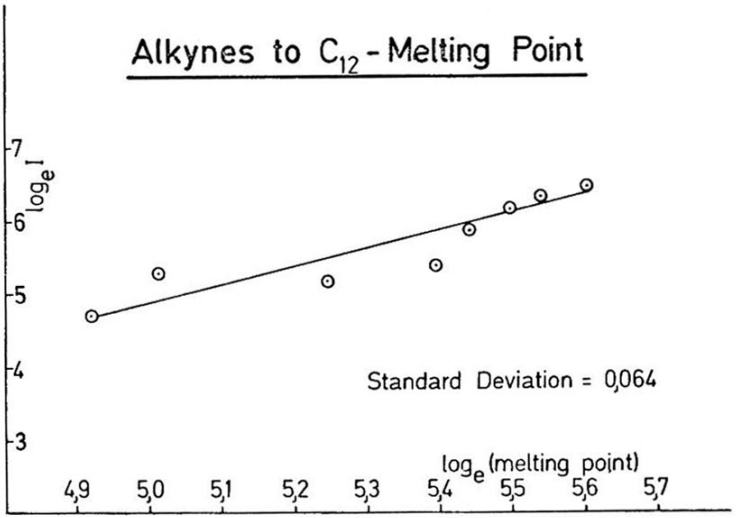

Alkanes to $C_{11}$ - Refractive Index
$\log_e I$
6
5
4
3
2
1
Standard Deviation = 0,002
$\log_e$(refractive index)
2,9
3,0
3,1
3,2
3,3
3,4

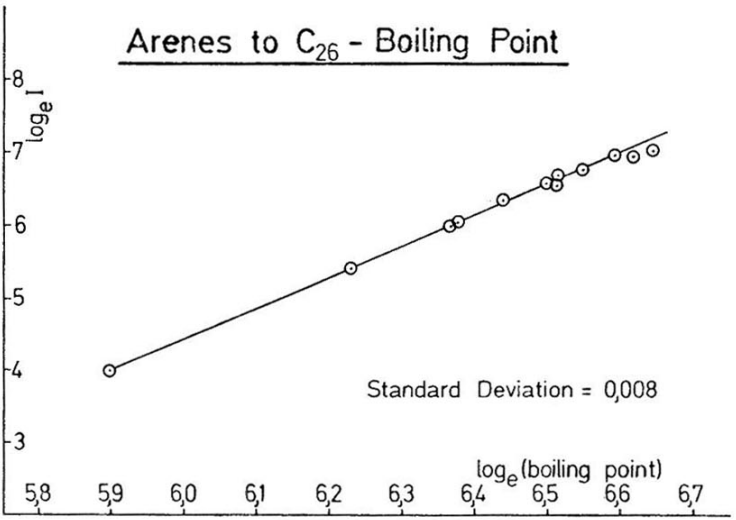

Alkanes to $C_{12}$ - Boiling Point
$\log_e I$
6
5
4
3
2
1
Standard Deviation = 0,005
$\log_e$ (boiling point)
5,2
5,4
5,6
5,8
6,0
6,2


Alkanes to $C_{12}$ - Melting Point
$\log_e I$
6
5
4
3
2
1
Standard Deviation = 0,087
$\log_e$ (melting point)
4,5
4,6
4,7
4,8
4,9
5,0
5,1
5,2
5,3
5,4
5,5
5,6


Alkanes to $C_{12}$ - Density
$\log_e I$
6
5
4
3
2
1
Standard deviation = 0,013
$\log_e$ (density)
-0,3
-0,4
-0,5
-0,6


Alkenes to $C_{12}$ - Surface Tension at 293,15 K
$\log_e I$
7
6
5
4
3
2
Standard Deviation = 0,050
$\log_e$ (surface tension)
2,5
2,6
2,7
2,8
2,9
3,0
3,1
3,2
3,3


Alkynes to $C_{12}$ - Melting Point
$\log_e I$
7
6
5
4
3
Standard Deviation = 0,064
$\log_e$ (melting point)
4,9
5,0
5,1
5,2
5,3
5,4
5,5
5,6
5,7


Arenes to $C_{26}$ - Boiling Point
$\log_e I$
8
7
6
5
4
3
Standard Deviation = 0,008
$\log_e$ (boiling point)
5,8
5,9
6,0
6,1
6,2
6,3
6,4
6,5
6,6
6,7

Günthard and Primas[7]. The Pairing Theorem is also realised to be a graph-theoretical result in essence[8]. Many original contributions to this general area have been made within the last five years by Gutman and Trinajstić[9].

The evolution of topological indices in their own right has come about during the last decade. The results obtained in this area up to 1973 have been reviewed by the present author[10]. The fundamental idea in this approach is that a topological index of the structure of a component species should reflect the bulk properties of the fluid formed from the species in question. The condition that all such species belong to a given chemical series, ensures that interactions between species will all be of the same general type. By making this condition, intermolecular forces may thereafter be neglected, at least to a first approximation.

To date fairly good correlations have been obtained with topological indices and the boiling points of members of several homologous series. Examples of such correlations are to be found in the works of Hosoya[11] and Randić[12]. For members of these series beyond about $C_{15}$, however, the correlations have frequently been rather poor. In an attempt to overcome these setbacks, the present author developed a new index, designed to be valid for a much wider range.

The new index proposed is based on the distance matrix of a chemical species. The index is derived by simply summing all the entries in this matrix, and so the new index is given by the summation

$$I = \sum_{i,j} d_{ij}$$

Although in such an index the structure within the matrix is lost, it was thought that this index might provide a better representation of a species than some of the other indices previously proposed, because it differentiates more clearly between similar structures such as 2-methyl hexane and 3-methyl hexane.

In practice, we have found good to excellent correl-

ations between our index and the bulk parameters of liquids. In our work we considered the members of the alkane, alkene, and alkyne homologous series up to the $C_{40}$ members, and arenes up to the $C_{26}$ members. Using these molecules, we looked for correlations of our index against boiling point, melting point, surface tension, viscosity, density, and refractive index. A relationship of the general form

$$P = k_1(I)^{k_2}$$

was assumed, where P is the bulk physicochemical parameter, $k_1$ and $k_2$ are constants to be determined for a given series, and I is our index. In all cases ln P was plotted against ln I to test the relationship.

Our investigations have revealed that in all cases a correlation of the type assumed does exist. For the first dozen or so members of all series investigated a fairly good approximation to a straight line was obtained. The best correlations were obtained for boiling points, and the worst for melting points. Standard deviations were usually of the order of 0,01 or less in this range. After the first dozen or so members, the plots assumed a distinct curvature, and standard deviations became correspondingly poorer. All plots were found to be convex with respect to the ln I axis, with the exception of the viscosity plots which were always concave.

The shapes of the curves may be accounted for on the following qualitative basis. We suppose that the molecular topology is of overriding influence in determining physicochemical parameters within a given series, as the intermolecular forces may all be assumed of the same type. This supposition explains in general terms the correlations we have found. But, because molecules are not rigid, we can expect an increasing amount of folding and bending in the liquid phase as the carbon content increases. As a result of this, a relatively large molecule folded into the form of a hairpin might behave as a molecule of only one half its true length. Such folding would amount to an effective loss of topological structure, which should become more pronounced with increasing carbon content. This would lead

to a progressive lowering of the effective value of our index, and thus to the convex curves which we obtained.

In the case of the viscosity plots, a loss of topological structure should result in a reduced resistance to flow, and therefore to an increased viscosity. On this basis a concave curve is thus to be expected. Work is now underway to make these qualitative explanations quantitative by study of simulated liquid structures. The relatively poor correlations with melting points may be due to the fact that lattice forces are superimposed upon the usual intermolecular forces, and these to some extent modify the force field in which the molecules find themselves. Lattice forces being relatively independent of topological structure, it is only to be expected that poor correlations were found for this parameter.

In conclusion, we may say that, because our correlations were generally good, that the importance of topology as a determinant of physicochemical behaviour has been established. In fact, in some cases the correlations were so good that it would be feasible to make predictions of unknown parameters from our index. Furthermore, the curvature found for all our plots might be utilized to provide information on the intermolecular forces in liquids and the relative arrangement of the molecules in this phase. Up to the present, the only type of parameters to have been investigated have been those of a physical or chemical nature. In future, it may prove possible to correlate biological parameters against a suitable index. Work is at present underway in Johannesburg to correlate the degree of carcinogenicity in arene species in this way. It would thus appear that this fruitful new line of inquiry is far from exhausted.

References

1. R.W. Jotham, Chem. Soc. Reviews 2, 457 (1973).
2. H. Kopp, Ann. Chem. 41, 79, 169 (1842).
3. V.V. Markovnikov, Ann. Chem. 182, 324 (1876).
4. E. Hückel, Z. Phys. 70, 204, 279 (1931); idem. ibid. 72, 310 (1931).
5. C.A. Coulson, Proc. Roy. Soc. London A169, 413 (1939).
6. K. Ruedenberg, J. Chem. Phys. 22, 1878 (1954).
7. H.H. Günthard and H. Primas, Helv. Chim. Acta 39, 1645 (1956).
8. D. Cvetković, Publ. Fac. Electrotechnique, Ser. Mat. Fiz. (Univ. Belgrade) 354-356, 1 (1971).
9. I. Gutman and N. Trinajstić, Topics in Curr. Chem. 42, 49 (1973).
10. D.H. Rouvray, Amer. Sci. 61, 729 (1973).
11. H. Hosoya, Bull. Chem. Soc. Japan 44, 2332 (1971).
12. M. Randić, J. Amer. Chem. Soc., to appear.