

Correlations between topological features and physicochemical properties of molecules

D V S JAIN*, SUKHBIR SINGH and VIJAY GOMBAR†

Department of Chemistry, †Department of Pharmaceutical Sciences, Panjab University, Chandigarh 160 014, India

Abstract. The article reviews in brief, the *de novo* group additivity approach and, at length, the different topological approaches to obtain predictive and internally consistent correlations between various properties and structural features of molecules. The stress has mainly been on the molecular connectivity method. A new rational scheme for nomenclature of connectivity indices of different orders and types is introduced. The concept of the perturbation connectivity parameter developed by us recently has been applied to obtain correlations for molar refraction, boiling point, molar volume, heat of atomisation, heat of combustion, heat of vaporisation, magnetic susceptibility and critical constants of alkanes, alcohols and alkylbenzenes. A comparative study of various approaches reveals that the present perturbation topological approach has an edge over other similar methods.

Keywords. Connectivity indices; topological parameters; topology-property correlations; alcohols; alkylbenzenes; physicochemical properties.

1. Introduction

Molecular topology describes the manner of connectedness of atoms in a molecule. This is one of the basic factors responsible for differences in the values of physicochemical properties of isomeric molecules. The problem of establishing quantitative relationships between topology and properties has always been an important research area because such correlations can readily estimate and predict magnitude of properties particularly when experimental determinations are difficult. A great amount of effort has, therefore, been expended in developing procedures for obtaining such correlations. The earliest work in this direction was based upon the *de novo* approach using the additivity model. Towards the middle of the century, however, the problem was looked at from a different angle where molecules are first coded, according to their topology, in terms of numerical indices which are then treated as independent variables in regression analysis. Though the two models have inherent advantages and limitations, the efforts are on to parametrize both the approaches to generate internally consistent correlations possessing predictive powers within the experimental accuracy.

We are primarily interested in the development of the topological parameters. The present article, therefore, reviews, in brief, the work accomplished using the additivity model followed by a treatment, in detail, regarding the basis, terminology and chronological development of various topological parameters with special emphasis on connectivity indices. Recently, the present authors (Gombar *et al* 1983) developed a

* To whom all correspondence should be addressed.

connectivity index, called perturbation connectivity index, which incorporates the effects of atomic environments at topological level. Its correlations with properties like molar refraction, boiling point, molar volume, heat of atomisation, heat of combustion, heat of vaporisation, magnetic susceptibility and critical constants of alkanes, alcohols and alkylbenzenes are reported and compared with earlier results.

2. Additivity model

The principle of additivity of atomic contributions was used for correlation studies on molar refraction (Bruhl 1891; Eisenlohr 1910), parachor (Sugden 1924), molar volume (Glasstone 1946) and magnetic susceptibility (Souders *et al* 1949). The predictions were found to be closer to the experimental values when the additivity principle was applied to the bond contributions instead of atomic contributions (von Steiger 1921; Smyth 1925; Fajans and Knorr 1926; Benson and Buss 1958; Pauling 1960; Benson 1976) especially when explicit corrections for bond-bond interactions, influence of nearest neighbouring atoms, steric interactions etc were incorporated (Zahn 1934; Tatevskii 1950, 1951; Allen 1959; Tatevskii and Papulov 1960; Somayajulu and Zwolinski 1966, 1972, 1974). The bond additivity scheme was further refined by taking different groups of atoms as the independent contributing units (Parks and Huffman 1932; Klages 1949; Franklin 1949; Laidler 1956). The contributions of various groups involving hetero atoms, multiple bonds, cyclic fragments etc have been calculated for a variety of physiochemical properties (Platt 1947; Bremner and Thomas 1948; van Krevelen and Chermis 1951; Tatevskii *et al* 1961; Vetere 1976; Benson 1976; Hoshino *et al* 1978; Macknick and Prausnitz 1979; Ogilvie and Abu-Elgheit 1980).

3. Topological approach

3.1 Molecules as graphs

In the topological approach a molecule is represented by a hydrogen-suppressed graph (HSG) in which each vertex represents a non-hydrogen atom and each edge a bond (Sylvester 1878). A HSG (figure 1) consists of various subgraphs involving different number and arrangement of edges. The number of edges determines the order m and the arrangement decides the type t of the subgraph. Mathematically a HSG is usually represented by a binary adjacency matrix of order $n \times n$ (Sylvester 1878) in which unit and zero elements, respectively, represent bonded and non-bonded situations in a molecule.

3.2 Development of topology based parameters

Various indices have been defined in order to numerically encode molecules according to their topology. For n -alkanes the number of vertices (carbon atoms) carries enough topological quantification and significantly correlates with a number of properties (Aranda 1936; Komoshilov 1939; 1940; Colingoert *et al* 1941; Huggins 1941, 1954; Kurtz and Lipkin 1941; Francis 1941; Mibashan 1945; Smithenberg and Mulder 1948; Kurtz and Sankin 1954; Li *et al* 1956). In view of the correlations which could be applicable to branched alkanes as well, Wiener (1947, 1948) proposed that the difference ΔX between any property X of isomeric alkanes could be predicted in terms

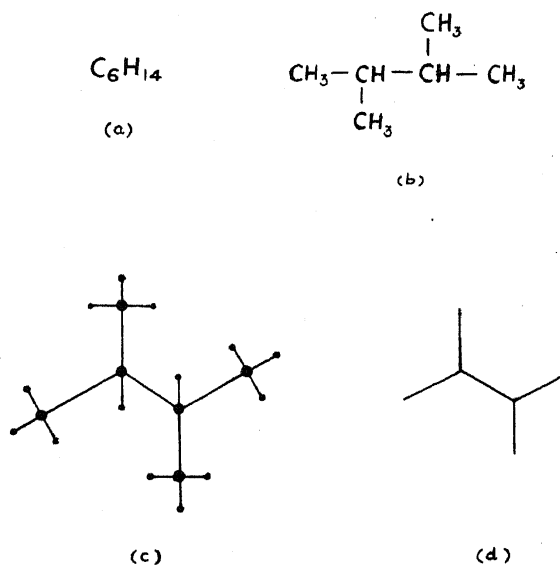


Figure 1. 2,3-dimethylbutane a. molecular formula, b. structural formula, c. molecular graph, d. hydrogen-suppressed graph.

of two parameters, *viz* w (path number) and p (polarity number), according to equation (1),

$$\Delta X = \frac{k}{n^2} \Delta w + b \Delta p, \quad (1)$$

where n is the number of vertices in the HSG of the alkanes concerned and k and b are regression coefficients. The polarity number p is the number of pairs of carbon atoms which are separated by three carbon-carbon bonds, and w is calculated according to (2) where d_{ij} is the general element of the distance matrix corresponding to the HSG.

$$w = \sum_{i < j}^n d_{ij}, \quad (2)$$

Gordon and Scantlebury (1964) defined a branching index of second order subgraphs B_2 which increases with degree of branching, by (3)

$$B_2 = N_2 / (n - 1), \quad (3)$$

where N_2 is the number of second order sub-graphs and n is the number of vertices. Using various subgraphs as the contributing units Smolenski (1964) blended the additivity and the topological approaches and observed a good correlation with standard heats of formation in terms of nine types of subgraphs. An integer topological index Z was proposed by Hosoya (1971). Z is related to the number of ways in which k edges are so chosen from the HSG that no two of them are connected. Gutman *et al* (1975) defined a topological index for acyclic polymers and concluded that their thermodynamic stability decreased and reactivity, in general, increased with branching. Randic (1975) distinguished between ten possible types of edges in the HSG of a hydrocarbon by associating each edge between vertices i and j with an index (edge index) equal to $(v_i v_j)^{1/2}$; v_i , being the number of edges converging at vertex i . He defined

a topological index (branching index) to be the sum of all the edge indices existing in a given HSG.

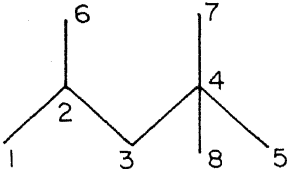




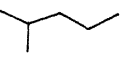
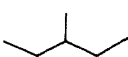
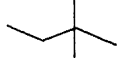
Some topological parameters *viz* information content (IC), complementary information content (CIC), structural information content (SIC) and relative nonstructural information content (RNSIC) have been defined (Basak *et al* 1979, 1982a, b; Basak and Magnuson 1983; Ray *et al* 1981, 1982, 1983; Raychaudhury *et al* 1980; Gombar and Wadhwa 1982) by applying information theory to the total molecular graph. These have, however, been mostly employed in correlation studies on biological properties. The discussion on these parameters is, therefore, not included in this paper. On the same grounds the DARC/PELCO method (Dubois *et al* 1967, 1973, 1976; Mercier and Dubois 1979; Mercier *et al* 1981; Sobel *et al* 1981) is not discussed here.

3.3 Connectivity indices

Kier *et al* (1975) defined a unitless topological quantifier called connectivity index χ identical to Randic's branching index for a series of alcohols and esters by treating oxygen atoms equivalent to carbon atoms and double bonds equivalent to single bonds. They obtained useful correlations with many physicochemical and non-specific biological properties (Kier *et al* 1975a, b; Hall *et al* 1975; Murray *et al* 1975, 1976). Since a single parameter cannot completely quantify the topology of a molecule, the approach of connectivity index was extended to higher order terms (Kier *et al* 1976). For more effective quantification of other salient topological features like types of atoms, multiplicity of bonds, cyclic fragments etc, valence connectivity indices, χ^v , of different types were defined (Kier and Hall 1976a, b, 1977; Kier 1980). A connectivity index dependent upon bond length, χ^r , was defined by Gupta and Singh (1978) and its correlations investigated with different biological properties. Jain and Gombar (1979) furthered the concept of path-connectivity indices to binary mixtures of *n*-alkanes. They assumed connectivity index of a binary mixture to be equal to molefraction average of the values of the components. Using first and third order connectivity indices of mixtures they obtained a highly predictive relationship for estimating excess volumes of binary mixtures of *n*-alkanes at different temperatures.

3.3a Nomenclature of connectivity indices Some of the higher order subgraphs, as defined by Kier and Hall (1976a), for 1,1,4-trimethylpentane, along with their counts and notations are given in table 1. It can be seen that according to notation given by Kier and Hall (1976a) the sub-graphs with different arrangement of vertices contribute towards same connectivity index. The ambiguity in the definitions is further accentuated when still higher order terms are considered. We have developed a uniform scheme (Gombar-Surjit-Jain (GSJ) scheme) for characterising unique subgraphs. The GSJ notations for connectivity indices corresponding to some of the subgraphs 1,1,4-trimethylpentane are also given in table 1. The superscript on χ equals the number of edges involved and the number of subscripts refers to the number of branchings in the contributing subgraph. The base of a subscript gives the branching position and its power represents the length of the branch. A terminal vertex is always assigned zero position and other vertices are numbered sequentially so that the branched position gets minimum number. Some of the sixth order connectivity indices, thus defined according to the GSJ scheme, are given in table 2. In the term ${}^8\chi_{2^2,2^2}$, read as eight-chi-two-square-two-square, therefore, the contributing subgraph (figure 2) involves 8 edges such that two branchings each of length two are present at position 2.

Table 1. Nomenclature, occurrence count and representative subgraphs in HSG of 1,1,4-trimethylpentane.

				
Order	Representative subgraphs		Resulting connectivity index	
	Topology	Occurrence	Kier's-notation	GSJ-notation
2		10	${}^2\chi$	${}^2\chi$
3		5	${}^3\chi_c$	${}^3\chi_1$
4		1	${}^4\chi_c$	${}^4\chi_{11}$
4		4	${}^4\chi_{pc}$	${}^4\chi_1$
5		7	${}^5\chi_{pc}$	${}^5\chi_1$
5		0	${}^5\chi_{pc}$	${}^5\chi_2$
5		1	${}^5\chi_{pc}$	${}^5\chi_{11}$

3.3b Calculation of connectivity indices For calculation of connectivity index of order m and type t ($=$ P, C, PC or CH) Kier and Hall (1977) extended Randic's (1975) idea of edge indices. Each vertex in a HGS is associated with a numerical value, called δ -value. The contribution C_j of a subgraph of order m and type t is calculated from (4).

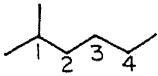
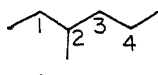
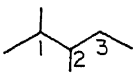
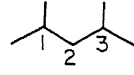
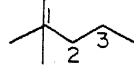
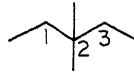
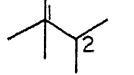
$$C_j = \prod_{i=1}^{m+1} (\delta_i)^{-1/2}. \quad (4)$$

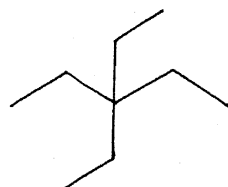
The total of C_j values calculated for all the subgraphs, mg_t , of order m and type t gives the required connectivity index, ${}^m\chi_t$, equation (5)

$${}^m\chi_t = \sum_{j=1}^{{}^mg_t} C_j. \quad (5)$$

The simplest version of connectivity indices is represented by ${}^m\chi_t^s$. In this case heteroatoms are treated equivalent to carbon atoms and double bond equivalent to single bond. The δ -value associated to each vertex in this case is δ^s which is equal to the number of vertices connected to it. In order to treat hetero-

Table 2. Gombur-Surjit-Jain (GSJ) representations of the sixth order connectivity indices and contributing subgraphs.

Subgraph	GSJ notation of connectivity index
	${}^6\chi_1$
	${}^6\chi_2$
	${}^6\chi_{12}$
	${}^6\chi_{13}$
	${}^6\chi_{11}$
	${}^6\chi_{22}$
	${}^6\chi_{112}$

**Figure 2.** A representative subgraph contributing towards ${}^8\chi_{2222}$.

atoms and bond multiplicities specifically which is essential to distinguish molecules $\text{CH}_3(\text{CH}_2)_2\text{CH}_3$, $\text{CH}_3(\text{CH}_2)_2\text{OH}$, $\text{CH}_3(\text{CH}_2)_2\text{NH}_2$, $\text{CH}_3\text{CH}_2\text{CHO}$, $\text{CH}_3\text{CH}_2\text{OCH}_3$, $\text{CH}_3\text{CH}=\text{CHCH}_3$ etc., at topological level, Kier and Hall (1976b) proposed the use of δ^v for δ -values. For the first row heteroatoms δ^v is defined by equation (6) and for higher members in a group by equation (7)*. In (6) and (7) Z_i and Z_i^v , respectively, denote total and valence shell electrons in atoms represented by vertex i and h is the number of hydrogens suppressed at vertex i in the HSG.

$$\delta_i^v = Z_i^v - h, \quad (6)$$

$$\delta_i^v = \frac{Z_i^v - h}{Z_i - Z_i^v}, \quad (7)$$

* For halogen atoms a set of empirical δ -values was proposed earlier but in recent compilations (Kier 1980; Kier and Hall 1983) the values derived from general equations (6) and (7) have been quoted for them. For subgraphs involving a fluorine atom the contribution C (cf eq (4)) is empirically assumed to be negative.

Recently Kier and Hall (1983) have proposed a general definition, equation (7a) of valence δ -values

$$\delta_i^v = \frac{Z_i^v - h}{Z_i - Z_i^v - 1} \quad (7a)$$

The resulting connectivity indices are called valence connectivity indices and are represented by ${}^m\chi_i^v$.

3.3c Computation of connectivity indices In the literature an unpublished computer program CFUNC (Hall)* is often cited for computer calculation of connectivity indices upto sixth order. We (Singh *et al* 1984) have developed a program, KNKTVT, in standard FORTRAN which can compute path-connectivity terms upto the highest order possible in a HSG and all the branched-connectivity indices upto seventh order.

The calculation of connectivity indices for a given set of δ -values requires counting of subgraphs of desired order and type. Our program, however, requires only the connection table as the input besides the appropriate δ -values. The program generates a logical matrix of order $n \times n$ whose elements are initiated .TRUE. or .FALSE. depending upon the input connection table. For calculation of a connectivity index of any order and type all the appropriate .TRUE. paths are then picked and their contributions computed.

4. Perturbation connectivity index

The δ^v values do distinguish between heteroatoms and bond multiplicities. However, a look at the δ^v values for methyl propionate (figure 3) reveals that in spite of different atomic environments the two oxygens are assigned equal values. Similar is the case with two terminal methyl carbons as well. We have incorporated the effects of atomic environments by defining perturbation δ -value, δ^p , which is related to δ -value as:

$$\Gamma \delta^p = \delta. \quad (8)$$

Here Γ is a square matrix whose diagonal elements are unity and off-diagonal elements are zero or $-\gamma_{ij}$ depending upon whether element $i-j$ corresponds to non-bonded or bonded situation in a molecule. The perturbation parameter γ_{ij} determines the environmental effect of the nearest atom j bonded to atom i . The elements of the column vector δ can be δ_i^s , δ_i^v or Z_i^v . The selection of δ^s , δ^v or Z^v for δ -values yields, respectively, the connectivity indices ${}^m\chi^p$, ${}^m\chi^{p(v)}$ or ${}^m\chi^{p(e)}$.

To calculate δ_i^p from (8) it is necessary to know the numerical value of γ_{ij} . As there is no *a priori* method for such assignment, the perturbation parameters have to be treated

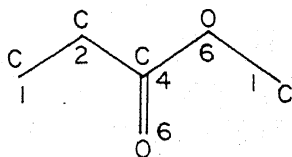


Figure 3. δ^v values assigned to vertices of methyl propionate.

* Program listings can be obtained from: Professor L. G. Hall, Eastern Nazarene College, Massachusetts.

empirically. It can be seen that the Kier's approach (Kier and Hall 1976b) becomes a special case of our general approach when all γ_{ij} are zero. Cammarata (1979) also used a similar approach to quantify molecular topology. However he defined only two parameters M' and A' as:

$$M' = \sum_i^n \delta_i^P, \quad (9)$$

$$A' = \sum_i^n (1/\delta_i^P), \quad (10)$$

where δ_i^P were calculated from (7) using $\gamma_{ij} = \gamma_{ji} = \gamma = 0.1$ and employing Z_i^v for δ -values.

4.1 Correlations using perturbation connectivity indices

A summary of quantitative correlations of a number of physicochemical and biological properties with simple and valence connectivity indices was reported by Kier and Hall (1977). Many more relationships have been obtained since then. The properties correlated so far are collected in table 3. In the following sections we discuss the correlations obtained between the newly defined perturbation connectivity indices ${}^m\chi^{P(e)}$ and some of the physicochemical properties of alkanes, alcohols and alkyl-benzenes. The indices ${}^m\chi^{P(e)}$ used in the present correlations have been calculated assuming $\gamma_{ij} = \gamma_{ji} = 0.1$ for all vertices i and j . The correlations thus obtained have been compared with those reported in the literature (Kier and Hall 1976a). For all the properties reported here we have obtained the correlations using both valence and perturbation connectivity indices. It was felt necessary first of all to assess the relative merits of the two because the sample set was either different from that reported by Kier and Hall or the systematic procedure was not followed by them for regression analysis.

4.1a Molar refraction This is one of the most commonly investigated molecular property for assessing the potential of any newly defined topology quantifier because fairly accurate experimental data on molar refraction R_m are available in the literature for a number of compounds. Kier and Hall (1976a) have investigated at great length the correlations between connectivity indices of various order and types and molar refraction of alcohols (31, 0.153, ${}^1\chi^v$, ${}^3\chi_c^v$, ${}^3\chi_p^v$, ${}^3\chi_c^v$)*, ethers (9, 0.291, ${}^1\chi^v$, ${}^3\chi_c^v$), amines (22, 1.280, ${}^1\chi^v$, ${}^3\chi_p^v$, ${}^3\chi_p^v$), alkyl halides (17, 0.108, ${}^1\chi$, ${}^1\chi^v$, ${}^3\chi_c^v$, ${}^3\chi_c^v$, ${}^3\chi_p^v$), mixed aliphatic compounds (64, 0.364, ${}^1\chi$, ${}^1\chi^v$, ${}^3\chi_c^v$, ${}^3\chi_c^v$, δ_N^v), alkanes (46, 0.027, ${}^1\chi$, ${}^2\chi$, ${}^3\chi_p^v$, ${}^4\chi_c^v$, ${}^4\chi_{pc}^v$), mono olefinic alkenes (39, 0.147, ${}^0\chi$, ${}^1\chi$, ${}^1\chi^v$, ${}^2\chi$, ${}^3\chi_p^v$, ${}^4\chi_c^v$) and alkyl substituted benzenes (70, 0.163, ${}^1\chi$, ${}^1\chi^v$, ${}^2\chi$, ${}^2\chi^v$). We have investigated a sample of 55 alkanes and obtained a three-variable correlation, equation (11),

$$R_m = -8.171 {}^0\chi^{P(e)} + 37.711 {}^1\chi^{P(e)} + 8.506 {}^2\chi^{P(e)} + 10.429$$

$$n = 55; r = 0.9999; s = 0.027; F > 9 \times 10^5, \quad (11)$$

which produces statistics comparable to that obtained by Kier and Hall (1976a) using five indices for a relatively smaller sample of 46 alkanes. Equation (11) yields maximum

* The arguments in parantheses are sample size ($n = 31$), standard error ($s = 0.153$) and various regression variables.

Table 3. Physicochemical properties correlated with connectivity indices.

Property	Class of Compounds	<i>n</i>	<i>r</i>	Reference
Aqueous solubility	Alcohols	51	0.978	Hall <i>et al</i> (1975)
	Alcohols	48	0.988	Kier and Hall (1976b)
	Alkanes	18	0.958	Hall <i>et al</i> (1975)
	Alkyl benzenes	13	0.9890	Kier and Hall (1976a)
	Alcohols	51	0.965	Cammarata (1979)
	Esters	38	0.990	Kier and Hall (1977)
	Ethers	22	0.989	Kier and Hall (1976a)
Boiling point	Alcohols	63	0.970	Hall <i>et al</i> (1975)
	Alcohols	48	0.993	Kier and Hall (1976b)
	Alcohols	28	0.9996	Kier and Hall (1976a)
	Alkanes	51	0.9969	Kier and Hall (1976a)
	Amines (primary)	12	0.9991	Kier and Hall (1976a)
	Amines (secondary)	13	0.999	Kier and Hall (1976a)
	Amines (tertiary)	4	0.997	Kier and Hall (1976a)
	Amines (primary)	21	0.999	Kier and Hall (1976b)
	Alkyl halides	24	0.992	Kier and Hall (1976b)
	Ethers	11	0.9882	Kier and Hall (1976a)
	Polyols	9	0.995	Kier and Hall (1976b)
	Carboxylic acids	22	0.996	Gombar (1981)
	Alkanes and alcohols	69	0.978	Kier <i>et al</i> (1975a)
Cavity surface area	Aliphatic liquid Hydrocarbons	82	0.9889	Kier <i>et al</i> (1976)
Density	Hydrocarbons	13	0.9104	Kier and Hall (1976a)
	Aliphatic alcohols	42	0.9305	Kier <i>et al</i> (1976)
	Aliphatic acids	20	0.9831	Kier <i>et al</i> (1976)
	Aliphatic ethers	13	0.9104	Kier and Hall (1976a)
	Hydrocarbons	27	0.9999	Kier and Hall (1976a)
Diamagnetic susceptibility	Alcohols	15	0.9993	Kier and Hall (1976a)
Excess volume	Binary mixtures of <i>n</i> -alkanes			Jain and Gombar (1979)
Heat of atomisation	Alkanes	44	0.9999	Kier and Hall (1976a)
	Alcohols	19	0.9999	Kier and Hall (1976a)
	Ethers	12	0.9999	Kier and Hall (1976a)
	Monoolefines	37	0.9999	Kier and Hall (1976a)
	Thiols	14	0.9999	Kier and Hall (1976a)
Heat of formation	Alkanes	67	0.9996	Kier and Hall (1976a)
	Alcohols	20	0.9948	Kier and Hall (1976a)
Heat of vaporisation	Alkanes	44	0.9999	Kier and Hall (1976a)
	Alcohols	20	0.9982	Kier and Hall (1976a)
Molar refraction	Alcohols	31	0.9998	Kier and Hall (1976a)
	Alkanes	46	0.999	Kier and Hall (1976a)
	Amines	22	0.9881	Kier and Hall (1976a)
	Alkyl benzenes	70	0.9995	Kier and Hall (1976a)
	Ethers	9	0.9989	Kier and Hall (1976a)
	Halo compounds	17	0.999	Kier and Hall (1976a)
	Mixed	65	0.990	Kier and Hall (1976a)
	Monoolefinic alkenes	39	0.9994	Kier and Hall (1976a)
	Phenyl substituent	25	0.990	Kier and Hall (1976b)
	<i>n</i> -alkanes at 20°C	14	0.9999	Jain and Gombar (1979)
Molar volume	<i>n</i> -alkanes at 50°C	13	0.9999	Jain and Gombar (1979)
	<i>n</i> -alkanes at 100°C	15	0.9999	Jain and Gombar (1979)
Molecular polarizability	Mixed	36	0.990	Kier <i>et al</i> (1975a)

Table 3. (Continued)

Property	Class of Compounds	n	r	Reference
Octanol-water partition coefficient	Alcohols	49	0.997	Murray <i>et al</i> (1975)
	Alcohols	14	0.999	Boyd <i>et al</i> (1982)
	Amines	12	0.998	Boyd <i>et al</i> (1982)
	Amines	28	0.979	Murray <i>et al</i> (1975)
	Carboxylic acids	9	0.996	Murray <i>et al</i> (1975)
	Hydrocarbons	45	0.975	Murray <i>et al</i> (1975)
	Esters	24	0.999	Murray <i>et al</i> (1975)
	Ethers	12	0.976	Murray <i>et al</i> (1975)
	Ketones	16	0.993	Murray <i>et al</i> (1975)
	Phenols	16	0.988	Boyd <i>et al</i> (1982)
	Mixed	66	0.970	Boyd <i>et al</i> (1982)
	Mixed	138	0.986	Murray <i>et al</i> (1975)
Positronium annihilation cross-section area	Benzene and derivatives	17	0.997	Gombar <i>et al</i> (1980)
	Alkanes	47	0.999	Gombar <i>et al</i> (1980)
	Compounds containing halogens	31	0.998	Gombar <i>et al</i> (1980)
	Compounds containing oxygen	46	0.998	Gombar <i>et al</i> (1980)
	Unsaturated compounds	33	0.984	Gombar <i>et al</i> (1980)
Retention indices in squalene	Mixed	182	0.996	Gombar <i>et al</i> (1980)
	Alcohols		0.994	Kier and Hall (1979)
	Esters		0.991	Kier and Hall (1979)
	Ethers		0.968	Kier and Hall (1979)
	Ketones		0.996	Kier and Hall (1979)
Van der Waals constant-b	Mixed	20	0.981	Kier and Hall (1976b)

absolute deviation of $0.1 \text{ cm}^3 \text{ mol}^{-1}$ and predicts an absolute deviation of more than $0.05 \text{ cm}^3 \text{ mol}^{-1}$ only for three of the 55 alkanes. It may be noted that the accuracy in the experimental data ranges from 0.007 to $0.05 \text{ cm}^3 \text{ mol}^{-1}$. For a series of 31 alcohols a 4-variable ($^1\chi^v$, $^3\chi^v$, $^3\chi_P$, $^3\chi_c$) equation has been reported (Kier and Hall 1976a) which gives a standard error of $0.153 \text{ cm}^3 \text{ mol}^{-1}$. Using the zero order perturbation connectivity index, $^0\chi^{P(e)}$, we have found a correlation which produces comparable statistics for a sample of 32 alcohols. It was observed that a 5-variable equation, equation (12), yielded a standard error of $0.02 \text{ cm}^3 \text{ mol}^{-1}$ which is of the order of the experimental error.

$$R_m = 6.032 \ ^0\chi^{P(e)} + 8.211 \ ^1\chi^{P(e)} + 3.157 \ ^2\chi^{P(e)} + 1.115 \ ^4\chi^{P(e)} + 2.966 \ ^4\chi_1^{P(e)}$$

$$n = 32; r = 0.9999; s = 0.02; F > 7 \times 10^5. \quad (12)$$

We also studied a sample of 120 alkyl-substituted benzenes and obtained a 2-variable correlation, equation (13),

$$R_m = 10.113 \ ^0\chi^{P(e)} + 1.323 \ ^2\chi^{P(e)} - 5.994$$

$$n = 120; r = 0.9999; s = 0.17; F > 3 \times 10^6, \quad (13)$$

which produces statistics comparable to that obtained by Kier and Hall (1976a) using indices $^1\chi$, $^1\chi^v$, $^2\chi$ and $^2\chi^v$ for a smaller sample of 70 alkyl benzenes. It was observed that the stepwise regression procedure picked a 3-variable equation in χ^v , equation (14),

which was better than the 3-variable correlation in $\chi^{P(e)}$.

$$R_m = 2.168 {}^0\chi^v + 5.560 {}^1\chi^v + 0.925 {}^2\chi^v + 6.507$$

$$n = 120; r = 0.9999; s = 0.11; F > 5 \times 10^6. \quad (14)$$

4.1b Boiling point The estimation of boiling points t_b has been the subject of several studies (Wiener 1947; Kudchadkar *et al* 1968; Hosoya 1971; Randic 1975; Hall *et al* 1975). For isomeric hexanes and heptanes Kier and Hall (1976b) have compared the quality of correlations obtained between boiling point and various topological indices; Randic's edge index (Randic 1975) was shown to correlate the best. Using simple and valence connectivity indices they have reported correlations for alkanes (51, 2.53, ${}^1\chi$, ${}^4\chi_{pc}$), alcohols (62, 0.86, ${}^1\chi$, ${}^1\chi^v$, ${}^3\chi_p$, ${}^3\chi_p^v$, ${}^3\chi_c$), aliphatic ethers (11, 5.39, ${}^1\chi$, ${}^1\chi^v$), primary amines (12, 1.56, ${}^1\chi$, ${}^1\chi^v$), secondary amines (13, 2.20, ${}^1\chi$, ${}^1\chi^v$), tertiary amines (4, 2.92, ${}^0\chi^v$), aliphatic polyols (9, 3.78, ${}^1\chi$, ${}^1\chi^v$) and alkyl halides (17, 2.57, ${}^1\chi^v$, ${}^3\chi_p$, ${}^3\chi_c$). We studied a sample of 235 alkanes, including 96 *n*-alkanes, ranging between 5 and 100 carbon atoms. We obtained a highly significant nonlinear correlation, equation (15).

$$t_b = 916.978 - \frac{12406.583}{11.666 + {}^1\chi} + 4.634 {}^4\chi_1,$$

$$n = 235; r = 0.9999; s = 2.789; F > 2 \times 10^7. \quad (15)$$

A correlation obtained using ${}^1\chi^{P(e)}$ and ${}^4\chi_1^{P(e)}$, however, was slightly inferior ($s = 4.439^\circ\text{C}$). For the sub-sample of 96 *n*-alkanes we found that ${}^1\chi$ or ${}^1\chi^{P(e)}$ resulted in equation (16) with identical statistical quality.

$$t_b = 916.978 - \frac{4944.446}{4.806 + {}^1\chi^{P(e)} \text{ (or } {}^1\chi)}$$

$$n = 96; r = 0.9999; s = 1.185; F > 8 \times 10^7. \quad (16)$$

It was observed that the maximum absolute deviation in t_b calculated on the basis of (15) was $\sim 10^\circ$. It should be noted that the boiling points of the 235 alkanes investigated here range between 9 and 715°C with uncertainties from 0.002 to 20°C (Kovats 1961). Recently Ogilvie and Abu-Elgheit (1981) reported an average error of 2°C in prediction of t_b . The results are firstly based on a sample of 50 alkanes (C_4 – C_{20}) and secondly involve as many as 8 substructural parameters.

For a set of 28 monoaliphatic alcohols Kier and Hall (1976a) reported a standard error of 0.86°C using five variables (${}^1\chi$, ${}^1\chi^v$, ${}^3\chi_p$, ${}^3\chi_p^v$, ${}^3\chi_c$). Our investigations on a sample of 62 alcohols led to equation (17), which shows a standard error of 5.62°C .

$$t_b = -345.429 {}^0\chi^{P(e)} + 869.933 {}^1\chi^{P(e)} + 171.715 {}^3\chi^{P(e)} + 194.796,$$

$$n = 62; r = 0.9871; s = 5.62; F = 735, \quad (17)$$

It was observed that the standard error reduced to 2.23 when seven variables (${}^0\chi^v$, ${}^1\chi^v$, ${}^2\chi^v$, ${}^3\chi_p^v$, ${}^3\chi_c^v$, ${}^4\chi_p^v$, ${}^4\chi_{pc}^v$) were included in the regression. It may be noted that Kier and Hall (1976a) studied a selected sample and used connectivity indices from two different sets *viz* simple connectivity indices and valence connectivity indices. As far as we know there is no correlation reported in literature between boiling points and χ indices of alkylbenzenes. For a sample of 47 such compounds we obtained the

following correlations using ${}^m\chi^{P(e)}$, equation (18), and ${}^m\chi^v$, equation (19), indices.

$$t_b = -139.449 {}^0\chi^{P(e)} + 362.281 {}^1\chi^{P(e)} - \frac{338.479}{{}^1\chi^{P(e)}} + 143.149,$$

$$n = 47; r = 0.9970; s = 5.44; F = 2345, \quad (18)$$

$$t_b = 31.833 {}^0\chi^v - 20.318 {}^2\chi^v - \frac{150.277}{{}^1\chi^v} + 65.241,$$

$$n = 47; r = 0.996; s = 5.82; F = 2049. \quad (19)$$

It was observed that a seven-variable (${}^1\chi^{P(e)}$, ${}^2\chi^{P(e)}$, ${}^3\chi^{P(e)}$, ${}^4\chi^{P(e)}$, ${}^4\chi_1^{P(e)}$, $1/{}^0\chi^{P(e)}$, $1/{}^1\chi^{P(e)}$) resulted in a standard error of 1.66°C and a maximum absolute deviation of 3.7°C. For a set of 22 aliphatic saturated monocarboxylic acids (Gombbar 1981) a 2-variable correlation, equation (20), in simple connectivity indices has been reported.

$$t_b = 65.658 + 37.785 {}^1\chi + \frac{5.208}{{}^3\chi_1},$$

$$n = 22; r = 0.996; s = 3.7; F = 1320, \quad (20)$$

It produces an average absolute deviation of $< 1.5^\circ\text{C}$.

4.1c *Density and molar volume* Kier and Hall (1976a) have obtained good correlations for density ρ of alkanes (82, 0.0046, ${}^3\chi_p$, $1/{}^1\chi$), aliphatic alcohols (40, 0.0030, ${}^2\chi$, ${}^3\chi_p$, ${}^5\chi_p$, ${}^3\chi_c$, $1/{}^1\chi$), aliphatic ethers (13, 0.0072, ${}^3\chi_p$, ${}^5\chi_p$) and aliphatic acids (20, 0.0137, ${}^3\chi$, $1/{}^1\chi$). However, variation in density of organic compounds is relatively small over a wide range of molecular weights. We have, therefore, chosen to correlate molar volume v_m ($= M/\rho$) with connectivity indices. We have performed regression analysis on molar volumes of 87 alkanes. The correlation obtained, equation (21), predicted v_m with an average absolute deviation below 0.3 % which is of the same order as obtained by Kier and Hall (1976b) from a study on densities of 46 alkanes using a five-variable (${}^3\chi_p$, ${}^5\chi_p$, ${}^4\chi_{pc}$, ${}^5\chi_{pc}$, $1/{}^1\chi$) equation.

$$v_m = 80.098 {}^1\chi^{P(e)} + 21.252 {}^2\chi^{P(e)} - 45.122 {}^3\chi^{P(e)} + 46.790,$$

$$n = 87; r = 0.9999; s = 0.65; F > 9 \times 10^4 \quad (21)$$

Jain and Gombbar (1979) obtained correlations between molar volumes of n -alkanes at different temperatures and simple connectivity indices. By studying the temperature dependence of the regression coefficients of these correlations they deduced a general relationship to estimate molar volume of n -alkanes at any temperature between melting and boiling points. The extrapolation of this general relationship led to limiting values for density and thermal expansion coefficients of polymethylenes.

Investigations on molar volumes of a sample of 54 aliphatic alcohols resulted in a highly significant ($r = 0.9993$) correlation with ${}^0\chi^{P(e)}$. A three-variable equation, equation (22), however, yields a standard error of $0.60 \text{ cm}^3 \text{ mol}^{-1}$

$$v_m = 40.444 {}^0\chi^{P(e)} - 35.219 {}^3\chi^{P(e)} + 10.562 {}^4\chi^{P(e)} + 5.523,$$

$$n = 54; r = 0.9998; s = 0.60; F > 4 \times 10^4, \quad (22)$$

and predicts molar volumes of only two compounds with a deviation more than $1 \text{ cm}^3 \text{ mol}^{-1}$.

We also studied molar volumes of 120 alkyl substituted benzenes. A highly significant correlation, equation (23), was obtained which gave an average absolute error of less than 0.3%. It produced maximum absolute deviation of $4 \text{ cm}^3 \text{ mol}^{-1}$. A three-variable ($^1\chi^v$, $^2\chi^v$, $^4\chi_{pc}^v$) correlation in valence connectivity indices also resulted in comparable statistics ($s = 1.12$).

$$v_m = 32.129 {}^0\chi^{P(e)} + 23.445 {}^2\chi^{P(e)} - 44.472 {}^4\chi^{P(e)} - 27.404, \\ n = 120; r = 0.9999; s = 1.05; F > 7 \times 10^5 \quad (23)$$

4.1d Enthalpy of atomisation For hydrocarbons the enthalpy of atomisation ΔH_a is linearly dependent upon the number of carbon atoms (Kier and Hall 1976a). Using n , the number of carbon atoms, along with seven connectivity indices viz $^1\chi$, $^4\chi_p$, $^4\chi_{pc}$, $^5\chi_p$, $^5\chi_c$, $^5\chi_{pc}$, and $^6\chi_{pc}$, Kier and Hall (1976a) obtained a correlation which yielded a standard error of 1.55 kJ mol^{-1} . We have observed that a 3-variable correlation, in terms of connectivity indices only, generated by a stepwise procedure is much inferior ($s = 8.8 \text{ kJ mol}^{-1}$) to that obtained with ${}^m\chi^{P(e)}$, equation (24).

$$\Delta H_a = 2587.8 {}^0\chi^{P(e)} + 124.6 {}^2\chi^{P(e)} - \frac{63.2}{{}^0\chi^{P(e)}} + 916.6, \\ n = 44; r = 0.9999; s = 2.82; F > 5 \times 10^6. \quad (24)$$

Studies on heat of atomisation of 21 alcohols also revealed that correlations in ${}^m\chi^{P(e)}$ are better than those, with the same degrees of freedom, in ${}^m\chi^v$. It appears that (25) is the best relationship reported so far as it predicts ΔH_a with a standard error of 0.52 kJ mol^{-1} which is of the order of the experimental error.

$$\Delta H_a = 2636.5 {}^0\chi^{P(e)} - 204.2 {}^3\chi^{P(e)} + 244.2 {}^3\chi_1^{P(e)} + 332.7 {}^4\chi^{P(e)} \\ - 0.808/{}^1\chi^{P(e)} - 217.7, \\ n = 21; r = 0.9999; s = 0.52; F > 1 \times 10^8, \quad (25)$$

For a sample of 33 alkylbenzenes a 3-variable (${}^0\chi^v$, ${}^1\chi^v$, ${}^2\chi^v$) equation yielded a standard error of $14.16 \text{ kJ mol}^{-1}$ which is once again much inferior to that obtained using corresponding perturbation connectivity indices, equation (26).

$$\Delta H_a = 2375.8 {}^0\chi^{P(e)} + 472 {}^1\chi^{P(e)} + 254.5 {}^2\chi^{P(e)} - 2769.4, \\ n = 33; r = 0.9999; s = 2.93; F > 1.75 \times 10^6. \quad (26)$$

This gives an average absolute deviation of 0.04%. According to this equation only 3 compounds show absolute deviation of more than 2 kJ mol^{-1} but for 1,2,3,4-tetramethylbenzene which showed rather large deviation of 10.7 kJ mol^{-1} .

4.1e Diamagnetic susceptibility We studied the diamagnetic susceptibility χ_M of the same samples of alkanes and alcohols as investigated by Kier and Hall (1976a). The three-variable (${}^0\chi$, ${}^1\chi$, ${}^2\chi$) equation reported by them is comparable ($s = 0.41$) to the one

obtained using corresponding perturbation terms, equation (27)

$$\chi_M = 88.430 {}^0\chi^{P(e)} - 131.022 {}^1\chi^{P(e)} - 22.636 {}^2\chi^{P(e)} - 24.607, \\ n = 27; r = 0.9999; s = 0.40; F > 2 \times 10^4. \quad (27)$$

Only four of the 27 alkanes are found to give an absolute deviation of more than 0.5.

Kier and Hall (1976a) have reported a three-variable (${}^0\chi$, ${}^1\chi^v$, ${}^3\chi_P^v$) correlation ($s = 0.853$) for χ_M of 15 alcohols. Our stepwise procedure has, however, picked another set of three variables which correlates far better, equation (28).

$$\chi_M = 10.012 {}^0\chi^v + 8.603 {}^1\chi^v + 1.951 {}^4\chi_{pc}^v + 3.260 \\ n = 15; r = 0.9999; s = 0.27; F > 1.00 \times 10^4. \quad (28)$$

The quality of the computer-selected 3-variable equation with ${}^m\chi^{P(e)}$ is found to be midway between ($s = 0.42$). For a sample of 12 alkyl benzenes the correlation with ${}^m\chi^{P(e)}$ were better than those in ${}^m\chi^v$. A highly significant correlation, equation (29)

$$\chi_M = 46.879 {}^0\chi^{P(e)} - 45.309 {}^1\chi^{P(e)} - 8.126 {}^4\chi^{P(e)} - 16.402 \\ n = 12; r = 0.9999; s = 0.33; F > 3.00 \times 10^4 \quad (29)$$

resulted in maximum absolute deviation of 0.56.

4.1f Critical constants Experimental determination of critical constants T_c , P_c and V_c , is tedious. Their estimation by extrapolation of correlations is, therefore, very useful. A survey of the methods to estimate critical constants of isomeric alkanes has been reported by Spencer and Daubert (1973). Various correlations obtained with connectivity indices are discussed below.

Critical temperature: For a sample of 49 alkanes we obtained equation (30) in terms of simple connectivity indices.

$$T_c = 24.433 {}^0\chi - 11.017 {}^2\chi - \frac{348.782}{{}^1\chi} + 531.913 \\ n = 49; r = 0.9977; s = 6.03; F = 3215. \quad (30)$$

A three-variable correlation in $\chi^{P(e)}$ (${}^0\chi^{P(e)}$, ${}^3\chi^{P(e)}$, $1/{}^1\chi^{P(e)}$) is also found to be comparable ($s = 7.83$). Equation (30) predicts T_c within 1–2% which is of the same order as reported by Lydersen (1955) and Nokay (1959). Spencer and Daubert (1973) predicted T_c of alkanes with an error of < 0.5%. Average absolute deviations of 0.8% and 0.2% have been reported for two different samples of alkanes (Schmidt and Wenzel 1981). We observed that a 7-variable correlation (${}^0\chi^{P(e)}$, ${}^1\chi^{P(e)}$, ${}^3\chi^{P(e)}$, ${}^3\chi_1^{P(e)}$, ${}^4\chi_1^{P(e)}$, $1/{}^0\chi^{P(e)}$, $1/{}^1\chi^{P(e)}$) resulted in an absolute average error of 0.6% with maximum absolute deviation of 10°K.

For a sample of 13 alcohols we observed that a 3-variable equation in ${}^m\chi^{P(e)}$, equation (31), resulted in an average deviation of less than 1%. The deviation is found to fall below 0.25% when some more terms are included in the regression.

$$T_c = -18.443 {}^2\chi^{P(e)} + 652.607 {}^3\chi^{P(e)} - 737.099 {}^4\chi_1^{P(e)} + 515.039, \\ n = 13; r = 0.9964; s = 5.18; F = 414. \quad (31)$$

Schmidt and Wenzel (1981) estimated T_c of 13 alkyl benzenes with an average error of 0.8%. We found that a 2-variable equation in χ^v , equation (32), yields an average error of 1.3%.

$$T_c = 50.454 \, {}^3\chi_p^v - \frac{181.759}{{}^1\chi^v} + 619.830$$

$$n = 21; r = 0.9794; s = 8.49; F = 212. \quad (32)$$

Critical pressure: Kudchadkar *et al* (1968) estimated P_c of isoalkanes fairly accurately (maximum deviation = ± 0.14 atm) using vapour pressure-temperature data. By group contribution methods (Lydersen 1955; Schmidt and Wenzel 1981) errors from 3 to 4% have been reported for alkanes. Equation (33) also gives an average deviation of about 3%.

$$P_c = -7.522 \, {}^1\chi^{P(e)} + 20.872 \, {}^3\chi^{P(e)} + 38.418 \, {}^0\chi^{P(e)} + 20.722,$$

$$n = 49; r = 0.9932; s = 0.757; F = 1092. \quad (33)$$

This equation is, however, valid only for molecules containing upto 30 carbon atoms.

The P_c of 13 alcohols was found to be correlated with valence connectivity indices, equation (34).

$$P_c = -2.554 \, {}^0\chi^v + 4.164 \, {}^4\chi_{pc}^v + \frac{75.644}{{}^0\chi^v} + 30.942$$

$$n = 13; r = 0.9976; s = 1.17; F = 615. \quad (34)$$

This leads to an average error of 2%. A 5-variable (${}^0\chi^v$, ${}^3\chi_c^v$, ${}^4\chi_{pc}^v$, $1/{}^0\chi^v$, $1/{}^1\chi^v$) correlation reduces the error to merely 1% with maximum absolute deviation of 0.75 atm. There are no correlations available for comparison.

Forman and Thodos (1958, 1960) estimated P_c of alkylbenzenes with an average deviation of 3–4%. Schmidt and Wenzel (1981) estimated P_c for 13 alkylbenzenes with an error of 4.4%. Equation (35) predicts the values within 2% error and can be considered the best available so far for prediction of P_c .

$$P_c = 5.977 \, {}^0\chi^{P(e)} - 3.006 \, {}^4\chi^{P(e)} + \frac{147.098}{{}^0\chi^{P(e)}} - 61.816$$

$$n = 21; r = 0.9912; s = 0.78; F = 318. \quad (35)$$

This gives a maximum absolute deviation of 1.46 atm.

Critical volume: For a sample of 49 alkanes, whose critical volume, V_c , ranges between $255 \text{ cm}^3 \text{ mol}^{-1}$ and $1200 \text{ cm}^3 \text{ mol}^{-1}$, we obtained a highly significant correlation with ${}^m\chi^{P(e)}$, equation (36).

$$V_c = 347.458 \, {}^1\chi^{P(e)} - 277.823 \, {}^3\chi^{P(e)} + 141.754 \, {}^4\chi_1^{P(e)} + 46.323$$

$$n = 49; r = 0.9989; s = 10.8; F = 7026. \quad (36)$$

This is the best relationship known so far to estimate critical volume of alkanes. The earlier reported relationships yield average deviations of 2.3% (Lydersen 1955), 2% (Reidel 1954) and 1.78% (Vetere 1976) whereas equation (36) gives an average error of 1.4%. Vetere (1976) estimated critical volumes of 10 alcohols within 3% of the experimental values. We found that a 2-variable equation, equation (37), in ${}^m\chi^{P(e)}$ gives an average error of 1% only; that too for a larger sample.

$$V_c = 273.945 \, {}^1\chi^{P(e)} - \frac{0.368}{{}^3\chi^{P(e)}} + 67.370$$

$$n = 13; r = 0.9997; s = 3.29; F = 8215. \quad (37)$$

Using a large number of group contributions Vetere (1976) reported an average error of 0.96% for a sample of 18 alkylbenzenes. Our studies on 21 alkylbenzenes have led to equation (38) which gives comparable quality (average absolute error 1.1%) in terms of three ${}^m\chi^{P(e)}$ terms only.

$$V_c = 255.983 {}^1\chi^{P(e)} - 24.649 {}^4\chi_1^{P(e)} - \frac{153.108}{{}^1\chi^{P(e)}} - 54.652,$$

$$n = 21; r = 0.9980; s = 4.91; F = 1424. \quad (38)$$

4.1g Second virial coefficient Even the most carefully determined values of second virial coefficient B show experimental error of $\pm 50 \text{ cm}^3 \text{ mol}^{-1}$ for n -alkanes (Dymond and Smith 1969). Theoretical prediction of B is, therefore, very useful. Kunz and Kapner (1971) estimated B within 10–15% of the observed values. We observed a significant correlation between B and ${}^m\chi$, equation (39)

$$B = 46.322 - 229.820 ({}^1\chi)^2$$

$$n = 11; r = 0.9977; s = 41.74; F = 1942. \quad (39)$$

This equation predicts values with deviation less than 10%. For 2 of the 11 alkanes the deviation has been found to be more than $50 \text{ cm}^3 \text{ mol}^{-1}$.

The B values with experimental error of $\pm 16 \text{ cm}^3 \text{ mol}^{-1}$ are available for 8 alcohols (Dymond and Smith 1969). Polak and Ylu (1972) estimated B values for these alcohols with an average error of 3.2%. Our study using ${}^m\chi^{P(e)}$ indices led to equation (40) which gives an average error of 4%.

$$B = -718.847 {}^1\chi^{P(e)} - 2192.580 {}^3\chi^{P(e)} - 298.738$$

$$n = 8; r = 0.9914; s = 34.2; F = 143. \quad (40)$$

For a sample of 5 alkylbenzenes we obtained equation (41) which gives a maximum absolute deviation of $5.18 \text{ cm}^3 \text{ mol}^{-1}$.

$$B = 405.171 - 276.234 [{}^1\chi^{P(e)}]^3,$$

$$\bar{n} = 5; r = 0.9999; s = 4.9; F > 3 \times 10^4. \quad (41)$$

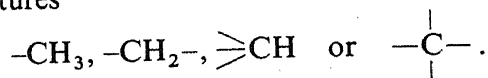
Equation (41), therefore, estimates B values within experimental uncertainties ($\pm 8 \text{ cm}^3 \text{ mol}^{-1}$).

5. Present status and future

The internal consistency and predictive capabilities of various correlations mentioned above loudly speak for the success of valence and perturbation connectivity indices in the topology-property correlation studies. However, at this stage it is difficult to commit which one of the connectivity indices leads to better correlations for all the physicochemical properties. It seems that perturbation connectivity indices correlate marginally better.

In spite of such good correlations which in many cases predict the values of various properties within experimental accuracy, there is no sound physical basis for them. Recently after a careful study on molar volume and enthalpy of vaporisation Edward (1982) has shown that different χ values give different weights to the contribution of

substructures



He also pointed out that the information in various higher order χ terms regarding degree of folding (Jain and Gombar 1979) is a function of the mole fraction, Z_g , of gauche conformations. Similar studies on other physicochemical properties should be interesting to reinforce the observation.

The connectivity indices are derived from the adjacency relationship of nonhydrogen atoms. These topology quantifiers, therefore, cannot numerically represent the three-dimensional features of molecules. We have recently attempted (Gombar *et al* 1983) to incorporate the effects of nearest neighbour atoms by defining perturbation connectivity indices. In the present study we have arbitrarily assumed $\gamma_{ij} = \gamma_{ji} = 0.10$ and in an earlier communication (Gombar *et al* 1984) we have optimized γ ($= \gamma_{ji} = \gamma_{ij}$) against statistical quality. It should be worth investigating to define γ as a function of bond length and/or electronic environments. As such all the χ indices do not encode electronic features of molecules. It is evident from the missing correlations between these indices and properties like dissociation constant and ionisation potential.

References

- Allen T L 1959 *J. Chem. Phys.* **31** 1039
 Aranda V G 1936 *An. Quim.* **34** 513
 Basak S C, Roy A B and Ghosh J J 1979 in *Proc. Second Int. Conf. on Mathematical Modelling*, Univ. of Missouri, USA, Vol. 2, pp. 851-856
 Basak S C, Gieschen D P, Magnuson V R and Harris D K 1982a *IRCS Med. Sci.* **10** 619
 Basak S C, Ray S K, Raychaudhury C, Roy A B and Ghosh J J 1982b *IRCS Med. Sci.* **10** 145
 Basak S C and Magnuson V R 1983 *Arzneim-Forsch/Drug Res.* **33** 501
 Benson S W 1976 *Thermochemical kinetics*, 2nd edn (New York: John Wiley) Ch. 1
 Benson S W and Buss J H 1958 *J. Chem. Phys.* **29** 546
 Boyd J C, Millership J S and Woolfson A D 1982 *J. Pharm. Pharmacol.* **34** 364
 Bremner J G M and Thomas G D 1948 *Trans. Faraday Soc.* **44** 230
 Bruhl J W 1891 *Z. Phys. Chem.* **7** 140
 Cammarata A 1979 *J. Pharmacol. Sci.* **68** 839
 Colingoert G, Beaty H A, Juoler R C and Thomos G W 1941 *Ind. Eng. Chem.* **33** 103
 Dubois J E, Laurent D and Viellard H 1967 *C. R. Acad. Sci. Paris* **C264** 1019
 Dubois J E, Laurent D and Aranda A 1973 *J. Chim. Phys.* **70** 1608
 Dubois J E, Laurent D, Bost P, Chambaud S and Mercier C 1976 *Eur. J. Med. Chem.* **11** 225
 Dymond J H and Smith E B 1969 *The virial coefficients of gases—A critical comparison* (Oxford: Clarendon Press)
 Edward J T 1982 *Can. J. Chem.* **60** 480
 Eisenlohr F 1910 *Z. Phys. Chem.* **75** 585
 Fajans K and Knorr J 1926 *Chem. Ber.* **59** 249
 Forman J C and Thodos G 1958 *A.I.Ch. E. J.* **4** 358
 Forman J C and Thodos G 1960 *A.I.Ch. E. J.* **6** 206
 Francis A W 1941 *Ind. Eng. Chem.* **33** 554
 Franklin J L 1949 *Ind. Eng. Chem.* **41** 1070
 Glasstone S 1946 *Textbook of physical chemistry* 2nd edn (Princeton, New Jersey: van Nostrand Reinhold) Ch. 8
 Gombar V K, Kapoor V K and Jain D V S 1980 *Indian J. Chem.* **A19** 715
 Gombar V, Jain D V S and Singh S 1983 Abstracts of the 35th Indian Pharmaceutical Congress, Paper A-18
 Gombar V, Jain D V S and Singh S 1984 *Q. S. A. R.* (communicated)

- Gombar V 1981 *The Pharmacos* 25 46
- Gombar V and Wadhwa L 1982 *Arzneim-Forsch./Drug Res.* 32 715
- Gordon M and Scantlebury G R 1964 *Trans. Faraday Soc.* 60 604
- Gupta S P and Singh P 1978 *Indian J. Chem.* B16 709
- Gutman I, Ruscic B, Trinajstić N and Wilcox C F Jr 1975 *J. Chem. Phys.* 62 3399
- Hall L H, Kier L B and Murray W J 1975 *J. Pharmacol. Sci.* 64 1974
- Hoshino D, Nagahama K and Hirata M 1978 *J. Chem. Eng. Jpn* 11 403
- Hosoya H 1971 *Bull. Chem. Soc. Jpn* 44 2332
- Huggins M L 1941 *J. Am. Chem. Soc.* 63 116
- Huggins M L 1954 *J. Am. Chem. Soc.* 76 343
- Jain D V S and Gombar V 1979 *J. Chem. Soc. Faraday Trans. I* 75 1132
- Klages F 1949 *Chem. Ber.* 82 358
- Kier L B, Hall L H, Murray W J and Randić M 1975a *J. Pharmacol. Sci.* 64 1971
- Kier L B, Murray W J and Hall L H 1975b *J. Med. Chem.* 18 1272
- Kier L B, Murray W J, Randić M and Hall L H 1976 *J. Pharmacol. Sci.* 65 1226
- Kier L B and Hall L H 1976a *Molecular connectivity in chemistry and drug research* (New York: Academic Press)
- Kier L B and Hall L H 1976b *J. Pharmacol. Sci.* 65 1806
- Kier L B and Hall L H 1977 *Eur. J. Med. Chem. Chim. Therapeut.* 12 307
- Kier L B and Hall L H 1979 *J. Pharmacol. Sci.* 68 120
- Kier L B 1980 in *Physical chemical properties of drugs* (eds) S H Yalkowsky, A A Sinkula and S C Valvani (New York: Marcel Dekker) Ch. 9
- Kier L B and Hall L H 1983 *J. Pharmacol. Sci.* 72 1170
- Komoshilov N F 1939 *Zh. Obshch. Khim.* 9 701
- Komoshilov N F 1940 *Zh. Obshch. Khim.* 10 945
- Kovats E 1961 *Z. Anal. Chem.* 181 351
- Kudchadkar A P, Holcomb W D and Zwolinski B J 1968 *J. Chem. Eng. Data* 13 182
- Kunz R G and Kapner R S 1971 *A.I.Ch. E. J.* 17 562
- Kurtz S S and Lipkin M R 1941 *Ind. Eng. Chem.* 33 779
- Kurtz S S and Sankin A 1954 *Ind. Eng. Chem.* 46 2186
- Laidler K J 1956 *Can. J. Chem.* 34 626
- Li K, Arnett R L, Epstein M B, Ries R B, Bitter L P, Lynch J M and Rossini F D 1956 *J. Phys. Chem.* 60 1400
- Lydersen A L 1955 Estimation of critical properties of organic compounds by the method of group contributions, Coll. Eng. Univ. Wisconsin, Eng. Expt. Sta. Rept.
- Macknick A B and Prausnitz J M 1979 *Ind. Eng. Chem. Fundam.* 18 348
- Mercier C and Dubois J E 1979 *Eur. J. Med. Chem.* 14 415
- Mercier C, Sobel Y and Dubois J E 1981 *Eur. J. Med. Chem.* 16 473
- Mibashan A 1945 *Trans. Faraday Soc.* 41 374
- Murray W J, Hall L H and Kier L B 1975 *J. Pharmacol. Sci.* 64 1978
- Murray W J, Kier L B and Hall L H 1976 *J. Med. Chem.* 19 573
- Nokay R 1959 *Chem. Eng.* 66 147
- Ogilvie J F and Abu-Elgheit M A 1981 *Comput. Chem.* 5 33
- Parks G S and Huffman H M 1932 Am. Chem. Soc. Monograph No. 60 (New York: Chem. Catalog. Co.)
- Pauling L 1960 *The nature of the chemical bond* 3rd edn (Ithaca, New York: Cornell Univ. Press) Ch. 3
- Platt J R 1947 *J. Chem. Phys.* 15 419
- Polak J and Ylu B C 1972 *Can. J. Chem. Eng.* 50 553
- Randić M 1975 *J. Am. Chem. Soc.* 97 6609
- Ray S K, Basak S C, Raychaudhury C, Roy A B and Ghosh J J 1981 *Indian J. Chem.* B20 894
- Ray S K, Basak S C, Raychaudhury C, Roy A B and Ghosh J J 1982 *Arzneim-Forsch./Drug Res.* 32 322
- Ray S K, Basak S C, Raychaudhury C, Roy A B and Ghosh J J 1983 *Arzneim-Forsch./Drug Res.* 33 353
- Raychaudhury C, Basak S C, Roy A B and Ghosh J J 1980 *Indian Drugs* 18 97
- Reidel L 1954 *Chem. Eng. Tech.* 26 83
- Schmidt C and Wenzel H 1981 *Can. J. Chem. Eng.* 59 527
- Singh S, Jain D V S and Gombar V 1984 (to be published)
- Smithenberg J and Mulder D 1948 *Rec. Trav. Chem.* 67 813
- Smolenski E A 1964 *Russ. J. Phys. Chem.* 38 700
- Smyth C P 1925 *Philos. Mag.* 50 361, 715

- Sobel Y, Mercier C and Dubois J E 1981 *Eur. J. Med. Chem.* **16** 477
 Somayajulu G R and Zwolinski B J 1966 *Trans. Faraday Soc.* **62** 2327
 Somayajulu G R and Zwolinski B J 1972 *Trans. Faraday Soc.* **68** 1971
 Somayajulu G R and Zwolinski B J 1974 *J. Chem. Soc. Trans. Faraday II* **70** 967, 973
 Souders M, Mathews C S and Hurd C O 1949 *Ind. Eng. Chem.* **41** 1037
 Spencer C F and Daubert T E 1973 *A.I.Ch. E. J.* **19** 482
 Sugden J 1924 *J. Chem. Soc.* **125** 32
 Sylvester J J 1878 *Am. J. Math.* **1** 64
 Tatevskii V M and Papulov 1960 *Russ. J. Phys. Chem.* **34** 115, 231, 335
 Tatevskii V M 1950 *Doklady Akad. Nauk SSSR* **75** 819
 Tatevskii V M 1951 *Zh. Fiz. Khim.* **25** 241
 Tatevskii V M, Banderskii V A and Yarovski 1961 *Rules and methods for calculating physicochemical properties of paraffinic hydrocarbons* (Oxford: Pergamon Press)
 van Krevelen D W and Chermin H A G 1951 *Chem. Eng. Sci.* **1** 66
 Vetere A 1976 *A.I.Ch. E. J.* **22** 950
 von Steiger A L 1921 *Chem. Ber.* **54** 1381
 Wiener H 1947 *J. Am. Chem. Soc.* **69** 2636, 17
 Wiener H 1948 *J. Phys. Chem.* **52** 425, 1082
 Zahn C T 1934 *J. Chem. Phys.* **2** 671