THEORETICA CHIMICA ACTA

Theoret. Chim. Acta (Berl.) 53, 355–375 (1979)

Chemical Graphs

XXXIV. Five New Topological Indices for the Branching of Tree-Like Graphs [1]

Alexandru T. Balaban

Organic Chemistry Department, The Polytechnic, Bucharest, Roumania

In order to find the centre of an acyclic connected graph (of a tree), vertices of degree one (endpoints) are removed stepwise. The numbers δ_i of vertices thus removed at each step form a digit sequence S (pruning sequence) which reflects the branching of the tree. The sum of squares of digits in the sequence S affords a new topological *centric index* $B = \sum_i \delta_i^2$ for the branching of trees. Comparisons with other topological indices are presented evidencing that B induces an ordering of isomeric trees distinct from those induced by all other indices devised so far, because B emphasizes equally branches of similar length.

It is shown that Rouvray's index I is equivalent to Wiener's index w, and that the Gordon-Scantlebury index N_2 and Gutman *et al.*'s index M_1 belong to the same family, called *quadratic indices*, and induce the same ordering.

Since all topological indices vary both with the branching and the number of vertices in the tree, four new indices are devised from B and M_1 to account only (or mainly) for the branching, by normalization (imposing a lower bound equal to zero for chain-graphs, i.e. *n*-alkanes) or binormalization (same lower bound, and upper bound equal to one for star-graphs). Normalized and binormalized centric (C, C') and quadratic indices (Q, Q') are presented for the lower alkanes. From the five new topological indices, the centric indices (B, C, C') are limited to trees, but the quadratic indices (Q, Q') apply to any graph. Binormalized indices (C', Q') express the "topological shape" of the graph.

Key words: Chemical graphs – Topological indices

1. Notation

n = the number of vertices in the tree, or of carbon atoms in the alkane; however, in "*n*-alkane" n stands for normal (linear).

0040-5744/79/0053/0355/$04.20

q = the number of edges in the tree, or of C—C bonds in the alkane; $q = n - 1$.

v_i = the degree of vertex i, or the number of carbon atoms bonded to carbon i; if $v_i \leqslant 4$ for any i, the tree is called a carbon tree (carbon skeleton of an alkane); $\sum_i v_i = 2q = 2n - 2$.

V_v = the number of vertices with degree v in the tree (graph).

X = an ordered sequence of integer digits $x_1, x_1, x_1, x_2, x_3, \ldots$ which may also be written as $x_1^{y_1}, x_2^{y_2}, \ldots$ when there are y_1 digits equal to x_1, y_2 digits equal to x_2, etc. (in the above example $y_1 = 3$, $y_2 = 1$, etc.). Examples for such sequences are

P = the *graphical partition* of vertex degrees, $P = v_1^{Vv_1}, v_2^{Vv_2}, \ldots$. For a tree with n vertices, $\sum_i v_i Vv_i = 2n - 2$; or

S = the *pruning* (*lopping*) *partition* of endpoints deleted in each step towards (from) the centre of the tree, $S = \delta_1, \delta_2, \ldots, \delta_r = 1^{z_1}, 2^{z_2}, \ldots, j^{z_j}, \ldots$; for n-vertex trees, $\sum_i \delta_i = \sum_j jz_j = n$.

$p(A)$ = partitions of integer A.

$p(G, k)$ = number of ways in which k edges in graph G may be chosen so that no two of them be adjacent.

$U = [1 - (-1)^n]/2$, i.e. $U = 0$ for even n, and $U = 1$ for odd n.

Other notation will be defined in the text.

2. Introduction

Topological indices for organic compounds attempt to express numerically topological information for a given molecular or constitutional graph. Stereochemical features are usually being disregarded. This graph is the usual constitutional formula where atoms of valence $\geqslant 2$ are represented by vertices, and covalent bonds joining these atoms are represented by edges; for a hydrocarbon this is the graph of the carbon skeleton of the molecule with the hydrogen atoms omitted, i.e. the hydrogen-depleted, or hydrogen-suppressed graph.

From the various topological indices which have been proposed so far, and which have been reviewed in part by Rouvray [2], we mention in roughly chronological order the following:

1. The Wiener numbers [3]: these are the path number, w, i.e. the sum of the number of bonds between all pairs of vertices, and the polarity number, p, i.e. the number of pairs of vertices separated by three edges. For acyclic graphs (trees), one may calculate w simply by multiplying together for each edge the two numbers of vertices of the two moieties joined by that edge, and then summing the products for all edges (cf. also No. 4 below). Correlations with boiling points, heats of formation and vaporization, molecular volume and molar refraction for alkanes have been described [3, 4]. Platt [4] introduced a third index, f(first-neighbour sum) which is calculated by determining for each edge the number of adjacent edges, and then summing these numbers for all edges (cf. No. 3 below).

2. Smolenskii's index [5] for Tatevskii's functions [6]:

$$f(G) = a_0 + \sum_{k=1}^{q} a_k X_k,$$

where a_0 and a_k are experimentally determined constants for the additive property under investigation, X_k is any section which contains k edges of the graph G having a total number of q edges. Tatevskii's contributions [6] do not contain any graph-theoretical formalism (cf. also [7]) but Smolenskii showed [5] the equivalence of Tatevskii's approach to this formalism.

3. Gordon and Scantlebury's index [8] N_2, i.e. the number of distinct ways a C—C—C fragment may be superimposed on the carbon skeleton (hydrogen-depleted graph). It can be shown that $N_2 = f/2$ (cf. No. 1 above). On dividing N_2 by $2(n-1)$,[1] where n is the number of vertices in the graph, or of carbon atoms in the hydrocarbon, another index is obtained [8], which may be used for comparing the branching of alkanes having different n values, and which for linear alkanes approaches $\frac{1}{2}$ as n increases towards infinity.

4. Altenburg's modification [9] of Wiener's index $w = \sum_i i g_i$, consists in the expression $\sum_i g_i a_i$ depending on the indexed variable a_i where g_i is the number of pairs of vertices whose distance is i, i.e. which are separated by at least i edges. Wiener's and Altenburg's expressions, originally devised for acyclic graphs (trees) may be extended to cyclic graphs [10]. The variable a_i has no special significance; it is introduced only for enumeration in polynomial form.

5. Hosoya's index [10] Z, defined as: $Z = \sum_{k=0}^{q} p(G, k)$ where $p(G, k)$ is the number of ways in which k edges of graph G may be chosen so that no two of them are directly connected; alternatively, Z may be defined for trees as the sum of the absolute values of coefficients in the characteristic polynomial using as variable x:

$$\sum_{k=0}^{s} (-1)^k p(G, k) x^{n-2k} = (-1)^n \det |\boldsymbol{A} - x\boldsymbol{E}|,$$

where s is the largest number of edges disconnected to each other in the tree-like graph, and $\boldsymbol{E}$ is the unit matrix. As known from graph theory (Harary's book quoted below under Ref. [35] is the standard text), the characteristic polynomial may be easily obtained from the adjacency matrix $\boldsymbol{A}$ by inserting x's on the main diagonal and resolving the matrix as a determinant equal to zero.

Hosoya's index has found many applications (e.g. correlations with boiling points, entropies, or calculated bond orders, as well as for coding of chemical structures) [10–13] and has been reported for numerous acyclic [12] and cyclic graphs [13] in the form of extensive tables.

---

[1] Formulations used here differ from the original ones [8]: originally N_2 was the second moment of distributions for the number of units with degree of substitution i in the jth n-isomer, and $2(n-1)$ was the first moment of distributions.

6. Hosoya's distance polynomial [14] $(-1)^n \det |\boldsymbol{D} - x\boldsymbol{E}|$, where $\boldsymbol{D}$ is the distance matrix having as entry D_{ij} the number of edges in the shortest path between vertices: D_{ij} is called the distance between vertices i and j.

7. Gutman *et al.*'s [15] index M_1 (which was also denoted [16] as Σ)

$$M_1 = \sum_{i=1}^{n} v_i^2,$$

where v_i is the degree of vertex i in the hydrogen-suppressed graph. Muirhead's criterion [17] for the comparability of functions was subsequently used [18] to discuss the concept of branching and the possibility of comparing graphs with different branching patterns. It is worth mentioning that Ruch *et al.* [19] used Muirhead's criterion for comparing Young partition diagrams, and that the same criterion was employed by Randić *et al.* [20] for recognizing structural similarity in molecules.

8. Randić's index [16] chi, χ, which was also denoted [15] as M_2 because it has a common origin with the preceding index, is defined as

$$\chi = \sum_{q} (v_i v_j)^{-1/2},$$

where v_i and v_j are the degrees of the two ends of an edge in the hydrogen-suppressed graph, and the summation is extended over all q edges. For the uses of χ, see below under No. 10.

9. Randić's translation into binary notation of the adjacency matrix by a unique non-subjective numbering of vertices in a graph parallels the ordering induced in alkanes by χ. The underlying idea was to permute rows and columns of the adjacency matrix until on reading sequentially the rows, the smallest number results. Since entries in the adjacency matrix $\boldsymbol{A}$ are 1 for adjacent and 0 for non-adjacent vertices, the result is a binary number. Thus a unique form of the adjacency matrix is found for a graph, leading to a unique numbering of vertices and to a linear binary notation for the graph. This binary number is conjectured to be a unique cipher for the topology of the graph [21], but it is too large and cumbersome (even after conversion into decimal notation) to be used as a normal topological index.

10. Kier *et al.*'s [22–29] generalized index ${}^h\chi$ or M_{h+1}, where $h \geqslant 2$ is the length of a path (an edge is a path of length one). If the summation indicated above under No. 8 for Randić's index (which may also be written as ${}^1\chi$) is extended from an edge to a path of length h over all possible such paths in the hydrogen-suppressed graph, the generalized index results ${}^h\chi = \sum (v_1 v_2 \cdots v_{h+1})^{-1/2}$, where $v_1, v_2, \ldots, v_{h+1}$ are the degrees of vertices in the path of length h. Alternatively, ${}^h\chi$ may be obtained from the hth power of the adjacency matrix $\boldsymbol{A}^h$ [29].

In practice, only χ, ${}^2\chi$, and ${}^3\chi$ have been used: χ gives good correlations with boiling point [22], the oil–water partition coefficient [23], enzyme inhibition [24], and with physiological properties such as barnacle larvae narcosis [23], tadpole inhibition [23] and anesthetic activity [25–28]. For density correlations, either $1/\chi$ or ${}^3\chi$ had to be used [29].

11. Rouvray's index, I [30], obtained by summation of all entries in the distance matrix $\boldsymbol{D}$. Evidently, $I = 2w$ (cf. No. 1 above).

12. Lovasz and Pelikan's observation [31] that the largest eigenvalue of the characteristic polynomial is a measure of branching led to the use of this eigenvalue as a topological index.

13. Bonchev and Trinajstić's indices [32] based on the information content of a set of numbers associated with a graph: the information on polynomial coefficients, cf. No. 5 above,

$$I_{pc} = Z \log_2 Z - \sum_k p(G, k) \log_2 p(G, k);$$

the information on the distribution of distances in the graph G,

$$I_E^D = [n(n-1)/2] \log_2 [n(n-1)/2] - \sum_i k_i \log_2 k_i,$$

where distance i appears $2k_i$ times in the distance matrix; the information I_D^w on the partitioning of the Wiener number $w = \sum_i ig_i$, expressed by $I_D^w = w \log_2 w - \sum_i ig_i \log_2 i$, as well as the mean values of the above three indices: $\bar{I}_{pc}$, $\bar{I}_D^E$, and $\bar{I}_D^w$. It was found [32] that $\bar{I}_D^E$ reflects accurately the main features of branching, even for different n values, and that both I_D^E and I_D^w discriminate efficiently among isomeric alkanes: in fact, I_D^w is more discriminating than w.

The authors [32] present a very lucid discussion of the concept of branching and they outline several rules of branching, indicating how branching increases when the structure is modified by altering the number, length, and position of branches.

All the above indices are topological (or semitopological in the case of Smolenski's index). However, empirical indices have also been used, like Kovats' index for correlating retention times in gas–liquid chromatography [33], but such empirical indices will not be discussed further here.

With two possible exceptions, the above topological indices do not characterize uniquely the topology of a constitutional graph, but are more or less degenerate. The two exceptions are Randić's reordered adjacency matrix read sequentially (No. 9), and, according to Hosoya's conjecture, the distance polynomial and the numerical data obtained therefrom (No. 6). The degree of degeneracy for the remaining indices is lower with indices I_D^w, χ, and Z in this order, with I_D^w having the lowest degeneracy. Most indices may be applied to any graph, but in certain cases there may be restrictions. Some indices are by definition integers, while χ, ${}^h\chi$ and informational indices are usually non-integer numbers.

3. Partitions of the Number n of Vertices in a Tree by Deletion of Endpoints towards (from) Its Centre

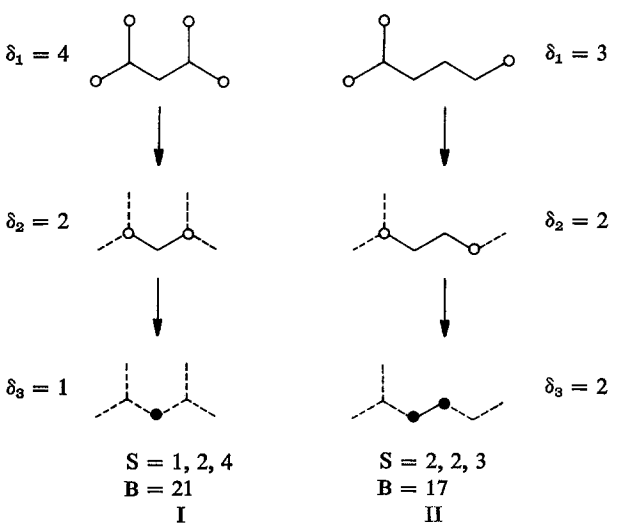
For trees it is well known in graph theory [34, 35] that a unique centre exists, be it a vertex or an edge. In the latter case, the pair of vertices connected by the edge may also be called a bicentre. The history of the independent discovery by Jordan and by Sylvester that every tree has a unique centre is presented in books by

König [34], and Harary [35a]. This property of trees was used by Neville [36], and more recently by Read [37] and Lederberg *et al.* [38] for codifying uniquely the structure of acyclic chemical compounds in linear notation systems or in computer programs, without resorting to elaborate conventions as in usual linear notation systems like Wiswesser's [39] or Dyson's [40]. As shown in the present paper, this property may also be used for obtaining a topological *centric index* for trees. In the following, the tree we shall use will be the hydrogen-suppressed constitutional graph of an alkane.

In order to find the centre or bicentre of a tree, each of the vertices of degree one (endpoints) and its incident edge are deleted, the operationg being repeated till the centre (bicentre) is left. We will call this operation *pruning*[2] the tree, and we shall denote by $\delta_1, \delta_2, \ldots, \delta_r$ the number of endpoints deleted in the first, second, ..., last (rth) stages of pruning. The number r of stages is called, following the usual graph-theoretical definitions [35a], the radius of the tree, or the minimum eccentricity. For instance, the two tree-graphs I and II (Table 1), both with seven vertices, which symbolize two isomers of C_7H_{16}, may be seen to lead in $r = 3$ stages to a centre for I, and to a bicentre for II.

If the numbers δ_i of endpoints deleted in each stage of the pruning are taken into account (as indicated on the margins of Table 1), then the reversed sequence $S = \delta_r, \delta_{r-1}, \ldots, \delta_2, \delta_1$ of these numbers, starting from the centre (bicentre) towards the initial graph, depends on the topology of the branching. Thus for I, $S = 1, 2, 4$ and for II, $S = 2, 2, 3$. The sequence S is automatically arranged in

Table 1. Finding the centre (bicentre), symbolized by black points, of trees. The dotted edges connect vertices (symbolized by white points) which were removed in previous steps. Trees I and II correspond to heptane isomers *d* and *g*, respectively, from Table 3



[2] Synonym: *lopping*.

non-decreasing order because the number of branches cannot decrease when starting from the centre of a connected tree (otherwise a disconnected graph i.e. a forest in graph-theoretical language, would result). Commas will be used to separate the digits (numbers) in sequence S and in subsequent sequences. We will call sequence S the *pruning partition*[3] of n (cf. item b below).

The following general remarks may be made for sequence S:
a) The first digit is always either 1 (for trees with a centre) or 2 (for trees with a bicentre). In the former case the longest chain is odd-membered whereas in the latter case it is even-membered.
b) Since all vertices must be accounted for in the pruning process, the pruning sequence S represents a partition of n. Any such partition may be translated into a tree by reverting the pruning process, provided that:

(i) There is at least one digit 1 (for a centre) or 2 (for a bicentre), cf. item a above; this condition follows from the fact that every tree has a centre or a bicentre.
(ii) There is at most one digit 1, since there is only one centre in a graph; this condition follows from the fact that when at a stage in the pruning process only one vertex is left, the pruning stops because the centre has been reached.

c) For carbon trees, in which vertices of degree higher than four are not allowed, there exist further restrictions for pruning partitions, among which we mention:
- partitions with two digits of type 1, x must have $x \leqslant 4$;
- partitions with two digits of type 2, x must have $x \leqslant 6$;
- partitions with three digits of type 1, 2, y must have $y \leqslant 8$;
- partitions with three digits of type 1, 3, y must have $y \leqslant 10$;
- partitions with three digits of type 1, 4, y must have $y \leqslant 12$;
- partitions with three digits of type 2, x, y must have $2 \leqslant x \leqslant 6$ and $y \leqslant 2x + 6$.

In a single formula, all the above restrictions may be formulated as follows for partitions S of type u, x, y (where y may be zero): u is either 1 or 2; $2 \leqslant x \leqslant 2u + 2$; $y \leqslant 2u + 2x + 2$.

Table 2 presents for all alkanes (carbon trees) with $n \leqslant 10$ the sequences S and, as exponent in brackets, the number of non-isomorphic trees having the same sequence S. When a sequence S corresponds only to one isomer, exponent (1) has been omitted from Table 2. The carbon trees with $n = 7$ are illustrated in Table 3, and those with $n = 8$ in Table 4.

4. A New Centric Index B for the Branching of Trees

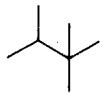
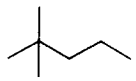
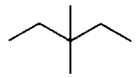
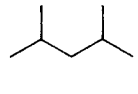
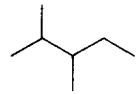
In order to obtain from the pruning partition (sequence S) an operational tool for comparing the branching of trees, two possibilities exist (analogously, on the basis of another sequence, the *graphical partition* of vertex degrees, symbolized by

[3] Synonym: *lopping partition*.

Table 2. Lopping partitions S, centric indices B, and the numbers of lopping and graphical partitions (S, and P, respectively) for trees with n vertices

n	Digit sequences S (and in brackets index B) for carbon trees (bracketed exponents indicate the number of isomeric alkanes with the same lopping sequence S)	Number of partitions			
		Carbon trees		All trees	
		S	P	S	P
2	2 (4)	1	1	1	1
3	1, 2 (5)	1	1	1	1
4	1,3 (10); 2,2 (8)	2	2	2	2
5	1,4 (17); 2,3 (13); 1,2,2 (9)	3	3	3	3
6	$2,4^{(2)}$ (24); $1,2,3^{(2)}$ (14); 2,2,2 (12)	3	4	4	5
7	2,5 (29); $1,2,4^{(4)}$ (21); 1,3,3 (19); $2,2,3^{(2)}$ (17); 1,2,2,2 (13)	5	5	6	7
8	2,6 (40); $1,2,5^{(4)}$ (30); $1,3,4^{(3)}$ (26) $2,2,4^{(6)}$ (24); 2,3,3 (22); $1,2,2,3^{(3)}$ (18); 2,2,2,2 (16)	7	7	8	11
9	1,2,6 (41); $1,3,5^{(3)}$ (35); $2,2,5^{(8)}$ (33); 1,4,4 (33); $2,3,4^{(4)}$ (29); $1,2,2,4^{(9)}$ (25); $1,2,3,3^{(2)}$ (23); $2,2,2,3^{(3)}$ 1,2,2,2,2 (17)	9	8	11	15
10	$1,2,7^{(2)}$ (54); $1,3,6^{(6)}$ (46); $2,2,6^{(7)}$ (44); $1,4,5^{(2)}$ (42); $2,4,4^{(2)}$ (36); $1,2,2,5^{(15)}$ (34); $1,2,3,4^{(9)}$ (30); $2,2,2,4^{(12)}$ (28); 1,3,3,3 (28); $2,2,3,3^{(4)}$ (26); $1,2,2,2,3^{(4)}$ (22); 2,2,2,2,2 (20)	13	10	15	22

Table 3. Isomers of heptane, i.e. carbon trees *a–i* with $n = 7$. First row under each tree, the lopping partition S and the derived indices B and C, in round and square brackets, respectively. Second row, the graphical partition P and the derived indices M_1 and Q, in round and square brackets, respectively. Further rows, binormalized centric index C', binormalized quadratic index Q', and other topological indices from the literature with their numbers as in the Introduction.

No.	Index	 *a*	 *b*	 *c*	 *d*	 *e*
—	$S\,(B)[C]$	2,5 (29)[8]	1,2,4 (21)[4]	1,2,4 (21)[4]	1,2,4 (21)[4]	1,2,4 (21)[4]
7	$P\,(M_1)[Q]$	$4,3,1^5$ (30)[4]	$4,2^2,1^4$ (28)[3]	$4,2^2,1^4$ (28)[3]	$3^2,2,1^4$ (26)[2]	$3^2,2,1^4$ (26)[2]
—	C'	0.667	0.333	0.333	0.333	0.333
—	Q'	0.4	0.3	0.3	0.2	0.2
1,11	$w = I/2$	42	46	44	48	46
1,3	$N_2 = f/2$	9	8	8	7	7
5	Z	13	14	16	15	17
8	χ	2.943	3.061	3.121	3.126	3.181
1	p	6	4	6	4	6

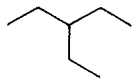
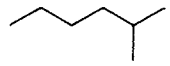
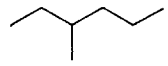
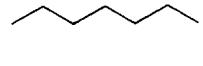
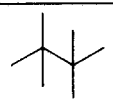
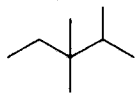
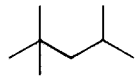
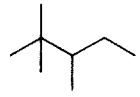
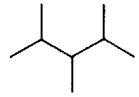
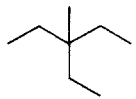
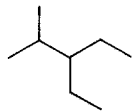
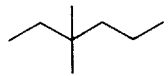
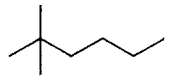
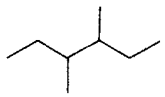
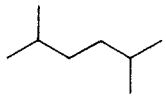
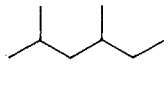
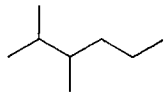
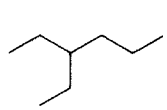
No.	Index	 *f*	 *g*	 *h*	 *i*
—	$S\,(B)[C]$	1,3,3 (19) [13]	2,2,3 (17)[2]	2,2,3 (17)[2]	1,2,2,2 (13)[0]
7	$P\,(M_1)[Q]$	$3,2^3,1^3$ (24)[1]	$3,2^3,1^3$ (24)[1]	$3,2^3,1^3$ (24)[1]	$2^5,1^2$ (22)[0]
—	C'	0.350	0.167	0.167	0
—	Q'	0.1	0.1	0.1	0
1,11	$w = I/2$	48	52	50	56
1,3	$N_2 = f/2$	6	6	6	5
5	Z	20	18	19	21
8	χ	3.346	3.270	3.308	3.414
1	p	6	4	5	4

Table 4. Isomers of octane, i.e. trees with 8 vertices, and some topological indices selected from those of Table 3

Sequence and/or index	 *a*	 *b*	 *c*	 *d*	 *e*
S (B)[C]	2,6 (40)[12]	1,2,5 (30)[7]	1,2,5 (30)[7]	1,2,5 (30)[7]	1,2,5 (30)[7]
P (M_1)[Q]	$4^2,1^6$ (38)[6]	4,3,2,1^5 (34)[4]	4,3,2,1^5 (34)[4]	4,3,2,1^5 (34)[4]	$3^3,1^5$ (32)[3]
C'	0.706	0.412	0.412	0.412	0.412
Q'	0.400	0.267	0.267	0.267	0.200
N_2	12	10	10	10	9
	 *f*	 *g*	 *h*	 *i*	
S (B)[C]	1,3,4 (26)[5]	1,3,4 (26)[5]	2,2,4 (24)[4]	2,2,4 (24)[4]	
P (M_1)[Q]	$4,2^3,1^4$ (32)[3]	$3^2,2^2,1^4$ (30)[2]	$4,2^3,1^4$ (32) [3]	$4,2^3,1^4$ (32)[3]	
C'	0.294	0.294	0.235	0.235	
Q'	0.200	0.133	0.200	0.200	
N_2	9	8	9	9	

	 *i*	 *k*	 *l*	 *m*	 *n*
$S\,(B)[C]$	2,2,4 (24)[4]	2,2,4 (24)[4]	2,2,4 (24)[4]	2,2,4 (24)[4]	2,3,3 (22)[3]
$P\,(M_1)[Q]$	$3^2,2^21^4$ (30)[2]	$3^2,2^2,1^4$ (30)[2]	$3^2,2^2,1^4$ (30)[2]	$3^3,2^2,1^4$ (30)[2]	$3,2^4,1^3$ (28)[1]
C'	0.235	0.235	0.235	0.235	0.176
Q'	0.133	0.133	0.133	0.133	0.067
N_2	8	8	8	8	7

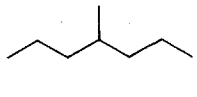
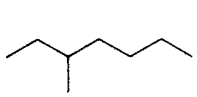
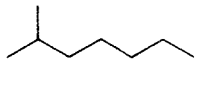
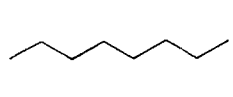
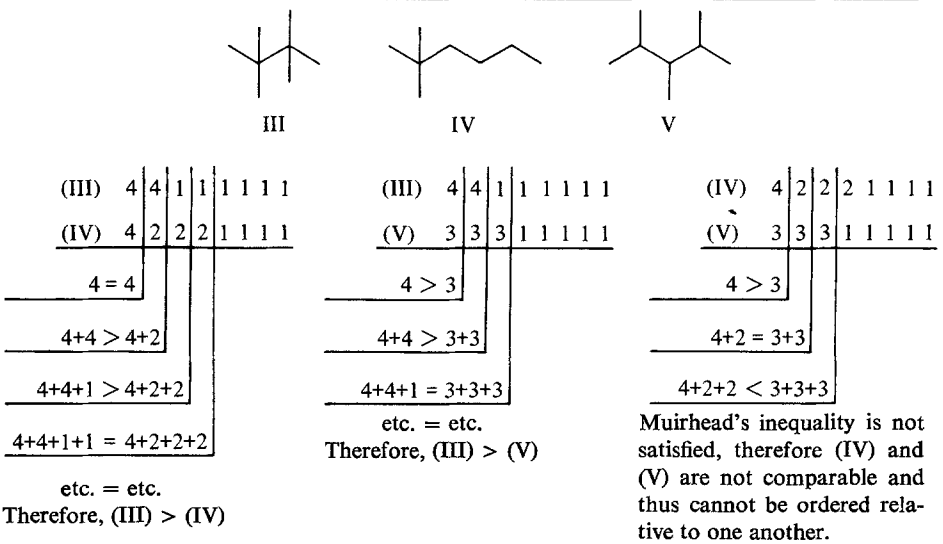
	 *o*	 *p*	 *q*	 *r*
$S\,(B)[C]$	1,2,2,3 (18)[1]	1,2,2,3 (18)[1]	1,2,2,3 (18)[1]	2,2,2,2 (16)[0]
$P\,(M_1)[Q]$	$3,2^4,1^3$ (28)[1]	$3,2^4,1^3$ (28)[1]	$3,2^4,1^3$ (28)[1]	$2^6,1^2$ (26)[0]
C'	0.059	0.059	0.059	0
Q'	0.067	0.067	0.067	0
N_2	7	7	7	6

Table 5. Ordering of trees with given graphical partitions P on the basis of Muirhead's criterion. Trees III, IV, and V correspond to octaneisomers *a*, *i*, and *e*, respectively, from Table 4



$P = v_1, v_2, \ldots, v_n$, both possibilities had been explored by Gutman *et al.* [15, 18]:

1) The first of these possibilities is to order trees according to a well-defined criterion, without defining a topological measure, or numerical index, for branching. It was argued [18] that only trees with the same n should be compared, and that even in this case an ordering of trees can only be made (in the case of partition P) if Muirhead's criterion for the comparability of functions is satisfied. This states that in two sequences in non-ascending order $a'_1, a'_2, \ldots, a'_k$, and $a_1, a_2, \ldots, a_k$, with $\sum_{i=1}^{k} a'_i = \sum_{i=1}^{k} a_i$, the former sequence is comparable to, and precedes, the latter if for all $1 \leqslant i \leqslant k$, Muirhead's inequality is obeyed, i.e. $\sum_1^i a'_1 \geqslant \sum_1^i a_i$. We illustrate in Table 5 this criterion for three isomers of octane taken from Ref. [18] with the graphical partitions $4^2, 1^6$ (for III), $4, 2^3, 1^4$ (for IV), and $3^3, 1^5$ (for V).[4] The result is expressed by the inequality III > (IV, V), i.e. on applying this criterion graph III precedes graphs IV and V, but the latter graphs IV and V cannot be ordered relative to one another.

2) The second possibility is to use the numbers in the given sequence by converting them into a topological index, which can be used both for ordering as in (1), and in quantitative correlations. Thus from the graphical partition P, index M_1 was obtained [15] (cf. No. 7 above).

There exists a correlation between the orderings derived from partition P and those introduced in alkanes by procedures 1) and 2), but the correlation coefficient is poor, as evidenced by the results in Table 6.

[4] In this abbreviated notation, the exponent indicates how many times one and the same digit is repeated.

Table 6. Differences between the ordering of trees according to Gutman *et al.*'s topological index [15] M_1, and according to Muirhead's criterion by Gutman and Randić [18]

n	M_1 (Ref. [15])	(Ref. [18])	Pairs of graphical partitions P
8	Same	Same	$4^2, 1^6$ and $5, 2^2, 1^5$; $4, 2^3, 1^4$ and $3^3, 1^5$
9	Same	Same	$5, 2^3, 1^5$ and $4^2, 2, 1^6$; $4, 2^4, 1^4$ and $3^3, 2, 1^5$
10	Same	Same	$4, 2^5, 1^4$ and $3^3, 2^2, 1^5$; $5, 3, 2^2, 1^6$ and $4^2, 3, 1^7$
9	Different	Same	$6, 2^2, 1^6$ and $5, 4, 1^7$
10	Different	Same	$7, 2^7, 1^7$ and $6, 4, 1^8$; $6, 3, 2, 1^7$ and $5^2, 1^8$
10	Different	Same	$6, 2^3, 1^6$ and $5, 4, 2, 1^7$; $4, 3, 2^3, 1^5$ and $3^4, 1^6$
10	Same	Different	$4, 3, 2^3, 1^5$ and $4, 3^2, 2, 1^6$

In the present case we want to replace sequence P by S. Both approaches 1) and 2) were tested, and it appeared that the latter gives better results. For carbon trees with $n \leqslant 9$ vertices, procedures 1) and 2) induce the same ordering; by using Muirhead's criterion 1), non-comparable partitions S which cannot be ordered relative to one another appear for carbon trees with $n \geqslant 9$: 5, 2, 2 and 4, 4, 1 ($n = 9$, same B). For $n = 10$, however, 2) appears to be more discriminating as indicated by the following non-comparable partitions S: 4, 2, 2, 2 and 3, 3, 3, 1 (same B); 6, 2, 2 and 5, 4, 1 (different B values); 4, 4, 2 and 5, 2, 2, 1 (different B values). Partitions S were rearranged in non-increasing order for applying Muirhead's criterion. Indices B may be found in Table 2 for all above partitions, according to the definition of B which follows. It thus appears that procedure 1) is less discriminating than 2).

In order to obtain a convenient, numerical, transform of the pruning partition (digit sequence S) a new topological index B for the branching of trees is defined by summing the squares of all numbers in sequence S: $B = \sum_{i=1}^{I} \delta_i^2$.

Tables 1–4 display, in addition to sequences S, the index B obtained according to the above formula. It may be seen that for a given n, the higher the value of B, the more branched the tree. Therefore B may serve as a topological index for the branching of trees, as will be presented in more detail in the next section. Because of the way it is derived, it will be called the *centric index B*.

5. Comparison between Topological Indices

Table 3 presents several other topological indices, including index B for each isomer of heptane. Some indices from Table 3 increase (N_2, f, M_1, B), while others ($w, I, Z,$) decrease with increasing branching (index p is also included in order to show that it does not correlate with branching, but rather with steric interference [4].

There exists a rich literature on graphs with a given partition of vertex degrees, i.e. with a given sequence P, e.g. a whole chapter in Harary's book [35b]. The sum of vertex degrees v_i for a tree on n vertices is $\sum_i v_i = 2q = 2n - 2$; however, only relatively few of the possible $p(2n - 2)$ partitions are graphical partitions $P = v_1, v_2, \ldots, v_n$, i.e. correspond to actual graphs. The sequence P of vertex degrees,

Table 7. Comparison of indices B. M_1, C and Q for the most branched (star-graph) and the least branched (chain-graph) trees with n vertices[a]

Partition and index	Star-graph	Chain-graph
Graphical partition P	$n-1, 1^{n-1}$	$2^{n-2}, 1^2$
Index M_1	n^2-n	$4n-6$
Index Q	$(n-2)(n-3)/2$	0
Lopping partition S	$1, n-1$	$U, 2^{(n-U)/2}$
Index B	$(n-1)^2+1$	$2n-U$
Index C	$[(n-2)^2-2+U]/2$	0

[a] $U = [1-(-1)^n]/2$, i.e. $U = 0$ for even n, and $U = 1$ for odd n.

which leads to index M_1, is appreciably longer than the sequence S which leads to index B; therefore the latter sequence S is written in full in Table 3, whereas the former sequence P has been abbreviated, e.g. 3, 2, 2, 2, 1, 1, 1 to $3, 2^3, 1^3$. Note that in this conventional abbreviation the exponents have no brackets.

The topological index B has, among all other indices known earlier and included in Table 3, the closest similarity to M_1. Both B and M_1 are derived analogously from sequences of digits: M_1 is based on the graphical partition P of vertex degrees, while B is based on the pruning partition S. For trees with n vertices, the sums of digits in sequences are $2n-2$ for P, and n for S. Obviously, for a given n, $p(n) < p(2n-2)$, in other words the numbers of mathematically possible sequences are larger for P than for S. It would be interesting to study as extensively the construction of graphs with given pruning partition S, as the analogous problem for given graphical partition P. It seems that the rules for knowing which pruning partitions can, or cannot, be converted into trees (or into carbon trees, with further restrictions) are much simpler than for graphical partitions. To find an enumeration formula for pruning partitions with given n either for trees or for carbon trees is an interesting and unsolved problem. Another similar problem would be to find an enumeration formula, and a construction procedure (which is an easier task), for obtaining all of the possible non-isomorphic trees with the same pruning partition.

Table 8. Ordering of carbon trees a–i with $n = 7$ induced by topological indices from Table 3

Indices	Order of carbon trees with $n = 7$
B, C, C'	$a(bcde)f(gh)i$
M_1, f, N_2, Q, Q'	$a(bc)(de)(fgh)i$
w, I	$ac(be)(df)hgi$
I_D^v	$acbedfhgi$
Z, I_{pc}	$abdceghfi$
χ	$abcdeghfi$
I_D^E	$acebfdhgi$

Examination of Table 2 reveals that for carbon trees with $n \leqslant 8$ the numbers of pruning and graphical partitions are equal (excepting $n = 6$) and that for $n = 9$–10 the number of pruning partitions is larger than the number of graphical partitions. Since each partition corresponds respectively to a certain value of index B or M_1, if follows that for a given n in the range 8 to about 12 (the highest n value we tested), the index B discriminates more effectively among isomeric carbon trees than the index M_1. Actually, if all trees (not only carbon trees) are taken into account, the numbers of allowed partitions for a given n are always greater for P than for S, as shown in the last two columns of Table 2. The reason why the converse is true for carbon trees with $n = 8$–10 (and presumably also for a few higher n values) is as follows: whereas carbon trees may have no digits larger than four in sequence P, they may in sequence S. On passing from all trees to carbon trees, one has to exclude those trees whose sequences S do not obey the rules indicated under item (c) in the Sect. 3, e.g. sequences of types 1, $x\ldots$ and 2, $y,\ldots$ with $x > 4$ and $y > 6$. These are fewer in number than sequences in P which contain numbers larger than 4, and which are also to be excluded on passing from trees to carbon trees. Indeed, on comparing the last four columns of Table 2, one may see that the differences are larger for P than for S on passing from trees to carbon trees. Thus for carbon trees with n lower than about 12, the sequences S (and hence the index B derived therefrom) offer certain advantages over sequences P and the derived index M_1.

Moreover, as it emerges from Tables 3 and 4, for a given n the range of variation for B is about twice as large as that for M_1. This is demonstrated by Table 7 which presents the indices B and M_1 for the most branched tree on n vertices (the star-graph formed by one vertex of degree $n - 1$, linked only to vertices of degree one) and the least branched tree (the chain-graph, i.e. an n-alkane). For convenience, in Table 7 both sequences (pruning and graphical partitions) are presented in abbreviated form, with exponents indicating how many times the same digit is repeated in the sequence.

It is apparent from Table 7 that for chain graphs the B index is about half as large as the M_1 index, whereas for highly branched graphs like the star graph both indices have almost equal values; asymptotically this becomes rigorously true, i.e. when $n \to \infty$ one has to replace "about" and "almost" by "exactly" in the previous sentence.

In other respects, B and M_1 are rather similar: their range of variation for a given n overlaps with ranges for higher and lower n, but when comparisons are made for the same n value this feature is no drawback.

A comparison between all topological indices from Table 3 is provided by Table 8 which illustrates the ordering induced among carbon trees with $n = 7$. It may be seen that indices M_1 and N_2 on one hand, and completely equivalent indices on the other hand (N_2 and f, w and I, which are written on the same line to save space) induce the same ordering. Also, it appears that though no index from Table 3 determines unambiguously the topology of a graph, indices Z, χ, and the informational indices have the smallest degeneracy, i.e. are the most discriminating ones.

Though the seven distinct orderings from Table 8 all start with tree *a* and end with tree *i* (both with $n = 7$), there exist definite differences between these orderings. Thus, indices *w*, *I*, I_D^w and I_D^E place *c* next after *a*, indices *Z*, I_{pc} and χ place *b* next after *a*, while the remaining indices (*B*, M_1, N_2, *f*) do not discriminate among *b* and *c* in this respect. Furthermore, indices *w*, *I*, I_D^w and I_D^E place *g* immediately before *i*, indices *Z*, I_{pc} and χ place *f* immediately before *i*, while N_2, *f*, and M_1 do not discriminate among *f* and *g*, but index *B* definitely places *f* before *g* or *h*.

Similar conclusions may be reached by examining the carbon trees with $n = 8$ presented in Table 4, which concentrates on fewer indices than Table 3, namely *B*, M_1 and N_2 (in addition to new indices which will be presented in the next section). The parallelism between Gutman *et al.*'s index M_1 and Gordon-Scantlebury's index N_2 is again evident. While index *B* induces the following ordering of trees with $n = 8$: *a*(*bcde*)(*fg*)(*hijklm*)*n*(*opq*)*r*, indices M_1 and N_2 induce the order: *a*(*bcd*)(*efhi*)(*gjklm*)(*nopq*)*r*; for comparison, other indices induce the following orderings: *Z*: *acd*(*bi*)*e*(*hk*)*lm*(*fg*)(*jq*)*opnr*; I_D^w: *abdfechgjmilnkopqr*; I_D^E: *abdfecghjm lnikopqr*; I_{pc}: *acdibekhlmgqofjpnr*.

Thus, while indices M_1 and N_2 result in the same weight for branches of lengths 1 and 2 on a longer chain, which is certainly wrong, index *B* gives a larger weight to the longer branch (compare trees *f*, *g*, *h* with $n = 7$, and trees *n*, *o*, *p*, *q* with $n = 8$). On the other hand, index *B* is the same for two isomeric alkanes which have the same number of equal short branches on a longer chain, no matter how these branches are attached; however, indices M_1 and N_2 in this case are lower when the branches are attached to different vertices, and higher when they are attached to the same vertex (geminal branches); on the other hand, neither *B*, nor M_1 and N_2, are able to discriminate between trees having the same types of branches in various positions (e.g. trees *j*, *k*, *l*, *m* with $n = 8$ from Table 4). For the purpose of such a finer discrimination, *Z*, χ, or I_D^w must be employed.

The above discussion illustrates the value of the new index *B*. Since the sequence *S* on which it is based starts with the centre of the tree, and includes the numbers δ_i of endpoints deleted at each pruning stage, *B* reflects the topology of the tree as viewed from the centre; this is why the name *centric index* was adopted.

The reason why indices M_1 and N_2 always induce the same ordering of trees is that they ascribe a larger weight to vertices of higher degree. Let V_v denote the number of vertices with degree *v* (for carbon trees, $1 \leqslant v \leqslant 4$). Evidently, $\sum_{v=1}^{4} V_v = 2n - 2$. According to the definition (cf. No. 7 in the introductory section),

$$M_1 = 16V_4 + 9V_3 + 4V_2 + V_1.$$

It can be demonstrated that the number of ways a C—C*—C fragment may be superimposed on a carbon tree so that the central C* coincides with a vertex of degree v_i is $\binom{v_i}{2}$, i.e. 6 for $v_i = 4$ and 3 for $v_i = 3$. Therefore we may write (cf. No. 3 in the Introduction)

$$N_2 = 6V_4 + 3V_3 + V_2.$$

Thus, both M_1 and N_2 belong to the same class of topological indices whose general expression is $\sum_{v=1}^{4} V_v f(v)$ where $f(v)$ is a quadratic polynomial in *v*: for M_1,

$f(v) = v^2$; for $N_2, f(v) = (v^2 - v)/2$. We call therefore both M_1 and N_2 as belonging to the class of *quadratic indices*. The polynomial $f(v)$ must be at least quadratic, since otherwise there would be no discrimination among isomers, e.g. with $f(v) = v$, one obtains the graphical partition P instead of a topological index, and it does not discriminate among isomers because its sum is $2n - 2$, the same for all isomers. The referee has pointed out that the use of higher powers for v than 2 may also lead to conflicting results, e.g. for the sequences $S = 6, 2, 2, 1$ and $5, 5, 1$ index B is greater for the second sequence, but if the fourth power were used instead of the square, the situation would be reversed (by the Muirhead–Ruch criterion, these two partitions are incomparable).

6. Normalized and Binormalized Centric (C, C') and Quadratic (Q, Q') Indices

All topological indices discussed in the Introduction vary both with the branching and with the number n of vertices, e.g. $w^{1/3}$ is a sort of mean molecular diameter of an alkane [4]. In some applications, as will be indicated below, we need to eliminate the variation due to n, i.e. we wish to know if the tree looks like a snake (chain-graph), a hedgehog (star-graph), or a hybrid therefrom. Dividing w by $n - 1$ as indicated [8] does not solve completely the problem.

We *normalize* a topological index by imposing for all graphs (regardless of n) the same lower bound, equal to zero for the least branched graph (chain-graph). Normalized indices, which are integers when they are derived from integer initial indices, still show a variation with n and with shape, therefore we *binormalize* topological indices by imposing, in addition to the above lower limit, the same upper bound equal to one, for the most branched tree (star-graph). Binormalized indices are integers only for the chain and star graphs, and for carbon trees they tend to decrease with increasing n when they have the same type of branching (this is because the higher the n value, the higher the denominator in the binormalized carbon-tree index). A prime symbol ($'$), added to the symbol of the normalized index, will be used to indicate binormalized indices.

Let X be a given sequence of integers: $X = x_1^{y_1}, x_2^{y_2}, \ldots$ if y_1 values equal to x_1 are present, y_2 values equal to x_2, etc. Let us assume that from this sequence a topological index was obtained by using a quadratic formula, e.g. index B from sequence S, or indices M_1 and N_2 from sequence P. From sequence X, a normalized index can be obtained in two ways: either by devising a new quadratic formula for obtaining automatically the value zero for the chain graph, or by subtracting from the index obtained by the usual, earlier, definition of the quadratic formula, the index of the chain-graph obtained by the same procedure. We shall demonstrate that both ways lead to the same result, allowing the definition of new indices.

For any graphical partition $P = 1^{V_1}, 2^{V_2}, \ldots, i^{V_i}, \ldots$ we have $\sum_i iV_i = 2n - 2$. The chain-graph (Table 7) has partition P (chain) $= 2^{n-2}, 1^2$ and its index after Gutman *et al.* [15] is $M_1 = \sum_i iV_i^2 = 2 + 4(n - 2) = 4n - 6$.

If we wish to find a *normalized quadratic index*, which will be denoted by Q, we have to find a quadratic function of the general form (where the constants A and

c must be determined): $Q = \sum_i (i^2 + ci)V_i + A$ which for the chain-graph should give $Q = 0$. It results that: $(n - 2)(4 + 2c) + 2(1 + c) + A = 0$ or, on rearranging, $n(4 + 2c) - (2c + 6 - A) = 0$. In order that this expression be independent of n, we must have $4 + 2c = 0$ and $2c + 6 - A = 0$. It follows that $c = -2$ and that $A = 2$. Thus the normalized quadratic index may be defined as

$$Q = \left[\sum_i (i^2 - 2i)V_i + 2\right]/2$$

where V_i is the number of vertices with degree i in the graph.
It follows that $2Q = \sum_i i^2V_i + 2 - 2\sum_i iV_i = \sum_i i^2V_i + 2 - 2(2n - 2) = \sum_i i^2V_i - (4n - 6) = M_1(\text{graph}) - M_1(\text{chain-graph})$, proving that both approaches have led to the same result.

For comparison with the expressions derived in the preceding section for M_1 and N_2, it should be mentioned that for carbon trees, the normalized quadratic index is $Q = (8V_4 + 3V_3 - V_1 + 2)/2$.

Since for any carbon tree on n vertices we have

$$V_1 = 2 + V_3 + 2V_4 \quad \text{and} \quad V_2 = n - 2 - 2V_3 - 3V_4,$$

(as it results immediately from $V_1 + V_2 + V_3 + V_4 = n$, and from $4V_4 + 3V_3 + 2V_2 + V_1 = 2q = 2n - 2$), the three expressions for M_1, N_2 and Q can be written as functions of only V_3 and V_4:

$$\begin{aligned} M_1 &= 16V_4 + 9V_3 + 4V_2 + V_1 = 2(3V_4 + V_3) + 4n - 6; \\ N_2 &= 6V_4 + 3V_3 + V_2 = 3V_4 + V_3 + n - 2; \\ Q &= 4V_4 + \tfrac{3}{2}V_3 + 1 - \tfrac{1}{2}V_1 = 3V_4 + V_3. \end{aligned}$$

It may be seen that all three quadratic indices M_1, N_2 and Q are interrelated by the equations:

$$Q = N_2 - n + 2 = 3 - 2n + M_1/2; \qquad M_1 = 2(N_2 + n - 1).$$

For any pruning partition $S = 1^{z_1}, 2^{z_2}, \ldots, j^{z_j}, \ldots$ we have $\sum_j jz_j = n$. The chain graph has partition $S(\text{chain}) = U, 2^{(n-U)/2}$, where $U = [1 - (-1)^n]/2$, and its index $B = \sum_j jz_j^2$ is $B(\text{chain}) = U + 2(n - U) = 2n - U$. It can be demonstrated as above that the expression for a *normalized centric index*, denoted by C, is

$$C = \left[\sum_j (j^2 - 2j)z_j + U\right]/2 = (B - 2n + U)/2.$$

It follows that $2C = \sum_j j^2z_j - 2\sum_j jz_j + U = \sum_j j^2z_j - (2n - U) = B(\text{graph}) - B(\text{chain})$, again demonstrating convergence of both approaches.

For obtaining binormalized indices, we divide Q and C by their respective values for the star-graph (cf. Table 7), and obtain:

$$\text{Binormalized centric index } C' = \frac{\sum_j (j^2 - 2j)z_j + U}{(n - 2)^2 - 2 + U} = \frac{B - 2n + U}{(n - 2)^2 - 2 + U};$$

$$\text{Binormalized quadratic index } Q' = \frac{\sum_i (i^2 - 2i)V_i + 2}{(n - 2)(n - 3)}.$$

Values for all four new normalized and binormalized indices C, C', Q, and Q' for the trees with $n = 7$ and 8 vertices are presented in Tables 3 and 4, respectively. The centric indices (C, C') parallel the ordering induced by index B, while the quadratic indices Q and Q' induce orderings which parallel those due to the other quadratic indices, M_1 and N_2.

Whereas the binormalized centric index C' is limited to trees (because only trees have a unique centre), the binormalized quadratic index Q' may be calculated for any graph. In the case of graphs with cycles, Q' values higher than 1 may be possible, e.g. for a ring on n vertices, $Q' = 8n/(n-2)(n-3)$, and for a cubic graph where all vertices have degree 3 like the valence isomers of annulenes [41], $Q' = 18n/(n-2)(n-3)$. Thus for regular graphs of degree 2 or 3 on less than 12 vertices, $Q' > 1$.

There is a discussion in the literature about the theoretical basis of determining the geometrical structure and shape of molecules [42]. The quadratic and centric binormalized indices Q' and C' for trees give information on the "topological shape" and on the "molecular volume" of the tree: for both binormalized indices, their values indicate how much hybridization between the "snake shape" (chain-graph, with $C' = Q' = 0$) and the "hedgehog shape" (star-graph, with $C' = Q' = 1$) is present in the given tree. Differences between C' and Q' mirror the differences between indices B and M_1 which were discussed in the preceding section.

7. Applications and Extensions

The anti-knock ability of alkanes in combusion engines is measured by their octane numbers, which depend both on the number n of carbon atoms and on branching. Linear correlations between each topological index discussed in the present paper and the experimentally determined octane number were tested separately for all isomers of heptane on one hand, and of octane on the other hand (these are the largest alkanes where octane numbers are known for practically all their isomers). The results, which will be presented in detail elsewhere [43] show that among all indices discussed in the present paper the centric indices give the highest correlation coefficients.

For comparing octane numbers of all alkanes with $n = 4$–8, linear biparametric correlations with n and a toplogical index were tested with all above indices. The statistical analysis (correlation coefficient, standard deviation, Fischer statistics, and explained variance) shows satisfactory results with several indices including the centric and quadratic indices; the normalized and binormalized indices have for n the smallest coefficients in these biparametric correlations, proving that the (bi)normalization eliminates most (though not all) of the variation of octane number due to the number of carbon atoms [43].

An extension of the notion of centre to cyclic graphs is possible. It will be published separately together with derived centric topological indices for cyclic graphs, including informational cyclic indices [44]. Applications of this extension of centre

to cyclic graphs for the purpose of chemical coding and documentation will be considered [45].

A review on correlations between chemical structure and biological activity, including a chapter on correlations with topological indices, is in press [28]. A review on mathematical models of branching is in press [46].

Acknowledgements. Thanks are expressed to Drs. Milan Randić, Dennis H. Rouvray, and Danail Bonchev for their valuable comments on the manuscript of the present paper.

References

1. Preceding part, Balaban, A. T.: Rev. Roumaine Chim. **23**, 733 (1978)
2. Rouvray, D. H.: Am. Sci. **61**, 729 (1973)
3. Wiener, H.: J. Am. Chem. Soc. **69**, 17, 2636 (1947); J. Chem. Phys. **15**, 766 (1947); J. Phys. Chem. **52**, 425, 1082 (1948)
4. Platt, J. R.: J. Chem. Phys. **15**, 419 (1947); J. Phys. Chem. **56**, 328 (1952)
5. Smolenskii, E. A.: Zhur. Fiz. Khim. **38**, 1288 (1964)
6. Tatevskii, V. M.: Khimicheskoe stroenye uglevodorodov i zakonomernosti v ikh fiziko-khimicheskikh svoistv. Moscow 1953; Tatevskii, V. M., Benderskii, V. A., Yarovoi, S. S.: Metody rascheta fiziko-khimicheskikh svoistv parafinovykh uglevodorodov. Moscow 1960; V. M. Tatevskii and Yu. G. Papulov, Zhur. Obshchei Khim. **34**, 241, 489, 708 (1960)
7. Papulov, Yu. G., Chulakova, L. V., Levin, V. P., Smolyakov, V. M.: Zhur. Fiz. Khim. **48**, 31 (1974); Bernstein, H. J.: J. Chem. Phys. **19**, 140 (1951); **20**, 263, 351 (1953); J. Phys. Chem. **69**, 1550 (1965)
8. Gordon M., Scantlebury, G. R.: Trans. Faraday Soc. **60**, 605 (1964)
9. Altenburg, K.: Kolloid Z. **178**, 112 (1961); Brennstoff Chem. **47**, 100, 331 (1966)
10. Hosoya, H.: Bull. Chem. Soc. Japan **44**, 2332 (1971); Internat. J. Quantum Chem. **6**, 801 (1972); J. Chem. Doc. **12**, 181 (1972); Chem. Letters (Japan) 65 (1972); Fibonacci Quart. **3**, 255 (1973)
11. Hosoya, H., Kawasaki, K., Mizutani, K.: Bull. Chem. Soc. Japan **45**, 3415 (1972)
12. Mizutani, K., Kawasaki, K., Hosoya, H.: Nat. Sci. Rept. Ochanomizu Univ. (Japan) **22**, 39 (1971)
13. Kawasaki, K., Mizutani, K., Hosoya, H.: Nat. Sci. Rept. Ochanomizu Univ. (Japan) **22**, 181 (1971)
14. Hosoya, H., Murakami, M., Gotoh, M.: Nat. Sci. Rept. Ochanomizu Univ. (Japan) **24**, 27 (1973)
15. Gutman, I., Ruscić, M., Trinajstić, N., Wilcox, C. F., Jr.: J. Chem. Phys. **62**, 3399 (1975)
16. Randić, M.: J. Am. Chem. Soc. **97**, 6609 (1975)
17. Muirhead, R. F.: Proc. Edinburgh Math. Soc. **19**, 36 (1901); **21**, 144 (1903); **24**, 45 (1906)
18. Gutman, I., Randić, M.: Chem. Phys. Letters **47**, 15 (1977)
19. Ruch, E., Schönhofer, A.: Theoret. Chim. Acta (Berl.) **19**, 225 (1970); Ruch, E.: Theoret. Chim. Acta (Berl.) **38**, 167 (1975); Ruch, E., Mead, A.: Theoret. Chim. Acta (Berl.) **41**, 95 (1976)
20. Randić, M., Wilkins, C. L. to appear
21. Randić, M.: J. Chem. Phys. **60**, 3920 (1974); **62**, 309 (1975); J. Chem. Info. Comp. Sci. **15**, 105 (1975); Chem. Phys. Letters **42**, 283 (1976)
22. Hall, L. H., Kier, L. B., Murray, W. J.: J. Pharm. Sci. **64**, 1974 (1975)
23. Murray, W. J., Hall, L. H., Kier, L. B.: J. Pharm. Sci. **64**, 1978 (1975)
24. Kier, L. B., Murray, W. J., Hall, L. H.: J. Med. Chem. **18**, 1272 (1975)
25. Kier, L. B., Hall, L. H., Murray, W. J., Randić, M.: J. Pharm. Sci. **64**, 1971 (1975)
26. Kier, L. B., Hall, L. H.: Molecular connectivity in chemistry and drug research. New York: Academic Press 1976

27. Nizhnii, S. V., Epstein, N. A.: Usp. Khim. **47**, 739 (1978)
28. Balaban, A. T., Chiriac, A., Motoc, I., Simon, Z.: Lecture Notes in Chemistry, in press
29. Kier, L. B., Murray, W. J., Randić, M., Hall, L. H.: J. Pharm. Sci. **65**, 1226 (1976)
30. Rouvray, D. H., Crafford, B. C.: South African J. Sci. **72**, 47 (1976)
31. Lovasz, L., Pelikan, J.: Period. Math. Hung. **3**, 175 (1973)
32. Bonchev, D., Trinajstić, N.: J. Chem. Phys. **67**, 4517 (1977)
33. Kováts, E.: Z. Analyt. Chem. **181**, 351 (1961)
34. König, D.: Theorie der endlichen und unendlichen Graphen, p. 64. Leipzig: Akad. Verlagsges 1936; reprinted Chelsea, New York, 1950
35. Harary, F.: Graph theory. Reading, Mass.: Addison-Wesley 1972: *a*), p. 35; *b*), p. 57; *c*), p. 233.
36. Neville, E. H.: Proc. Cambridge Phil. Soc. **49**, 381 (1953)
37. Read, R. C.: The coding of trees and tree-like graphs; Read, R. C., Milner, R. S.: A new system for the designation of chemical compounds for the purposes of data retrieval. I. Acyclic compounds" (Preprints, University of the West Indies, Jamaica, 1969)
38. Lederberg, J., Sutherland, G. L., Buchanan, B. G., Feigenbaum, E. A., Robertson, A. V., Duffield, A. M., Djerassi, C.: J. Am. Chem. Soc. **91**, 2973 (1969); Lederberg, J., in: The mathematical sciences, p. 37. Cambridge, Mass.: The M.I.T. Press 1969
39. Wiswesser, W. J.: A line-formula chemical notation. New York: T. Y. Crowell 1954
40. Dyson, G. M.: A new notation for organic chemistry. London: Chem. Soc. 1946; I.U.P.A.C.: Rules for I.U.P.A.C. notation for organic compounds. London: Longmans, Green and Co. 1961
41. Balaban, A. T.: Rev. Roumaine Chim. **11**, 1097 (1966); **15**, 463 (1970)
42. Woolley, R. G.: Advan. Phys. **25**, 27 (1976); J. Am. Chem. Soc. **100**, 1073 (1978)
43. Balaban, A. T., Moţoc, I.: Math. Chem., **5**, 197 (1979)
44. Bonchev, D., Balaban, A. T., Mekenjan, O.: to be published
45. Balaban, A. T., Bonchev, D., Mekenjan, O.: to be published
46. Bonchev, D., Knop, J. V., Trinajstić, N.: to appear

Received March 6, 1979