

---

**SOME TOPOLOGICAL INDICES DERIVED FROM THE  $v^m d^n$  MATRIX.  
PART 1. WIENER-LIKE INDICES OF THE  $BI_M$ -TYPE**

**Anton Perdih,\* Branislav Perdih**  
*Mala vas 12, SI-1000 Ljubljana, Slovenia*

*Received 06-06-2001*

**Abstract**

The matrix having nondiagonal elements  $v^m d^n$  ( $v^m d^n$  matrix) is a step in unification of several matrices used to derive topological indices, i.e. of the adjacency matrix, the distance matrix, the reciprocal distance matrix, etc. From the  $v^m d^n$  matrix were derived 225 Wiener-like indices  $W(m,n)$ . All of them are of the  $BI_M$ -type, i.e. they obey the *Methane-based* definition of branching. The index  $W(-\infty,-\infty)$  links the  $BI_M$  indices with the  $BI_A$  indices obeying the *n-Alkane-based* definition of branching. It is a simple, primitive, and degenerated, but a true branching index showing that the presence and the number of methyl groups (vertices of degree one) is of the highest importance in branching (complexity) of molecules. The index  $W(-\infty,0)$  contains also the information of the size of a molecule. The indices  $W(1,0)$ ,  $W(0,0)$ ,  $W(0,-\infty)$ , and  $W(-1,-\infty)$  only present information about the size of molecules. Most of the  $W(m,-\infty)$  and  $W(m,0)$  indices contain also the information that a branch on a quaternary carbon (i.e. on a vertex of degree four) contributes more to branching (complexity) than a branch on a tertiary carbon (i.e. on a vertex of degree three). The indices  $W(-\infty,n)$  and  $W(0,n)$  present, besides information about the size of a molecule and the number of branches, also information about the position of branches and the distance between them. Other  $W(m,n)$  indices present different, in some cases quite extreme combinations of all these parts of information.

**Introduction**

Several hundred topological indices have been developed and tested for their performance as branching indices or indices of substances' properties (Randić et al.<sup>1</sup> and references cited therein). A substantial part of them is derived from one or another matrix. Ivanciuc<sup>2</sup> mentions 14 different types of matrices as sources of topological indices. That these types of matrices are different has been considered also by other authors<sup>3</sup>. They made considerable efforts to define a general matrix.

The indices have been correlated with several physical, chemical, and biological properties of molecules. However, even several properties of alkanes still cannot be well

described with particular available indices<sup>4</sup> and combinations of them are to be used<sup>5</sup>. In spite of that, interest in topological indices has grown remarkably during the recent years. Therefore, the study of those topological indices that might be good branching indices remains an important area of research.

In previous papers<sup>6-8</sup> the suitability of 13 topological indices as branching indices, as well as of 25 physicochemical properties as reference properties for branching of alkanes was studied. The properties were divided into intrinsic and interaction-dependent ones and it was explained why most of the latter ones are not suitable as primary references for branching. The best reference properties to assess branching indices were found to be  $\Delta H_f^\circ$ , BP/Tc, and Tc/Pc. Two definitions of branching were presented, the *Methane based* definition as a general definition and the *n-Alkane based* definition as its subdefinition, more familiar to chemists working in this field. The indices obeying these definitions were labelled  $BI_M$  and  $BI_A$  indices, respectively. The search for true branching indices that would not necessarily fit particular properties but would only index branching has been encouraged. In the present paper, another set of indices that are derived from the distance matrix is evaluated.

The values of branching indices tested in previous papers<sup>6-8</sup> increase with the increasing carbon number of alkanes. Several of them, i.e. mostly those derived from the distance matrix, decrease with increasing branching, while some of them, i.e. mostly those derived from the adjacency matrix, increase. It is a general opinion that the best way to describe something is to do that in a simple and straightforward way, with direct relationship to the problem in question. A reverse relationship is acceptable if a direct one is lacking. According to this opinion, the indices increasing with branching would be more suitable branching indices than the others and the use of the adjacency matrix might give rise to better branching indices than the use of the distance matrix.

Studying this approach we became aware of another possibility of generalization of some of apparently different matrices. This generalization enables derivation of an infinite number of indices and we present here this possibility together with the properties of a limited set of 225 indices derived in this way.

### Data

The structures of alkanes are presented in shorthand, e.g. Hp is *n*-heptane, Oct is *n*-octane, 223M5 is 2,2,3-trimethylpentane, 3Et2M5 is 3-ethyl-2-methylpentane, etc. The other terms are explained on 2,2-, 2,3- and 2,5-dimethyl hexane (22M6, 23M6 and 25M6) as examples. The two branches (*2q*) in 22M6 are positioned on a quaternary carbon placed on the periphery (per.) of the molecule. The two branches in 23M6 and 25M6 are positioned on tertiary carbons (*2t*). In 23M6 the branches are adjacent and those in 25M6 are distant. The branches on carbons No. 2 and 5 are placed on the periphery of the molecule, and the one on carbon No. 3 is placed near the centre (ctr.) of the molecule. 25M6 is symmetric, whereas 22M6 and 23M6 are not.

Since we only consider alkanes, we use the words "vertex" and "carbon" as synonyms that reflect different contexts. The use of "vertex" implies the graph-theoretical consideration, whereas the use of "carbon" implies the chemical consideration of the same thing. Their number is given as  $N_C$  or  $N_v$ , respectively.

#### *The physicochemical properties*

The data for the boiling point (BP), density ( $d$ ), the critical data  $T_c$ ,  $P_c$ ,  $V_c$ ,  $Z_c$ ,  $\alpha_c$ , and  $d_c$ , as well as the standard enthalpy of formation for the ideal gas ( $\Delta H_f^\circ$ ), the enthalpy of vaporisation ( $\Delta H_v$ ), the Antoine constants  $A$ ,  $B$ , and  $C$ , as well as the Pitzer's acentric factor ( $\omega$ ) and the refractive index ( $n_D$ ) were taken from the CRC Handbook<sup>9</sup> or from Lange's Handbook<sup>10</sup>. The data for the liquid molar volume ( $V_m$ ), the intrinsic molar volume ( $V_i$ ), the ratios  $T_c^2/P_c$  and  $T_c/P_c$  used instead of the van der Waals parameters  $a_0$  and  $b_0$ ,  $BP/T_c$ , and the molar refraction (MR) were calculated from data presented in the handbooks.

### Methods

#### *The susceptibility for the increase in carbon number and the susceptibility for branching*

The susceptibility is defined as the normalised difference of the indices' or properties' values, Eq. 1,

$$S_{ij} = X_j/X_i - 1 \quad (1)$$

where  $S_{ij}$  is the susceptibility,  $X$  is an index or a property, subscript  $i$  refers to the less branched structure and subscript  $j$  refers to the more branched structure. Which one is

more branched is concluded by intuition as presented in ref.<sup>11</sup> as well as by the *Methane based* definition and the *n-Alkane based* definition<sup>6</sup>. Two groups of susceptibilities are used. In the susceptibility for the increase in carbon number, denoted as  $S_{i,i+1}$ , subscript  $i$  refers mostly to the structure having the same number and type of branches if not explicitly shown otherwise. In these cases, the two alkanes taken into account differ in carbon number by one. For example, in  $S_{7,8}$  the digit 8 means any octane having the same number and type of branches as a heptane which is represented by the digit 7. In  $S_{\text{Hp,Oct}}$  the data of *n*-heptane (Hp) and *n*-octane (Oct) are used. In  $S_{2\text{M}6,2\text{M}7}$  the data of 2-methyl hexane (2M6) and 2-methyl heptane (2M7) are used, in  $S_{2\text{M}6,3\text{M}7}$  the data of 2-methyl hexane (2M6) and 3-methyl heptane (3M7), etc.

In the susceptibility for branching,  $S_{n,i}$ , the subscript  $n$  refers to the *n*-alkane and the subscript  $i$  refers to any alkane of the same carbon number. For example, in  $S_{\text{Hp},i}$  Hp refers to *n*-heptane and  $i$  to any heptane. In  $S_{\text{Hp},2\text{M}6}$ , 2M6 refers to 2-methyl hexane.

## Results and discussion

### *The $v^m d^n$ matrix*

The adjacency matrix and the distance matrix usually seem to be considered as two different types of matrices and also other types of matrices, e.g. the reciprocal distance matrix, seem to have been considered as different matrix types.<sup>3,12</sup> We noticed that the adjacency matrix, the distance matrix and the reciprocal distance matrix are special cases of a more general type of distance matrix. This more general distance matrix,  $\mathbf{G}$ , having  $\mathbf{g} \times \mathbf{g}$  elements, has the main (backslash) diagonal elements,  $\mathbf{g}_{ii} = 0$ , and all the others,  $\mathbf{g}_{ij} (i \neq j) = f(a, b, \dots, v, \dots) \times f(d_{ij})$ , where  $a, b, \dots, v$ , etc. are some properties of vertices  $i$  and  $j$ , and  $d_{ij}$  is the distance between the vertices  $i$  and  $j$ . In the present paper we consider the elements  $\mathbf{g}_{ij} (i \neq j) = v_j^m \times d_{ij}^n$ , where  $v_j$  is the degree of vertex  $j$  (in alkanes it is the number of the C-C bonds which the carbon in question is involved in) and  $d_{ij}$  is the shortest distance from vertex  $i$  to vertex  $j$  (in alkanes it is the lowest number of bonds between the carbons in question). We label this matrix the  $v^m d^n$  matrix. In the simple distance matrix,  $\mathbf{m} = 0$  and  $\mathbf{n} = 1$ , in the simple reciprocal distance matrix  $\mathbf{m} = 0$  and  $\mathbf{n} = -1$ , whereas in the simple adjacency matrix  $\mathbf{m} = 0$  and  $\mathbf{n} = -\infty$ . Besides the matrix having the elements  $\mathbf{g}_{ij} (i \neq j) = v_j^m \times d_{ij}^n$ , other more general matrices are possible,

too. For example, Ivanciuc<sup>13</sup> tested some indices derived from a matrix having the elements  $\mathbf{g}_{ij} (i \neq j) = v_i^m \times d_{ij}^n$ , as well as from a matrix having the elements  $\mathbf{g}_{ij} (i \neq j) = v_i^m \times v_j^m \times d_{ij}^n$ . It can be seen that all these matrices are special cases of another more general matrix having the elements  $\mathbf{g}_{ij} (i \neq j) = v_i^a \times v_j^b \times d_{ij}^c$ , where in our case  $\mathbf{a} = 0$ ,  $\mathbf{b} = \mathbf{m}$ , and  $\mathbf{c} = \mathbf{n}$ . These more general matrices offer another possibility for a generalization of special matrices allowing the derivation of an infinite number of indices, in addition to that presented recently by Estrada<sup>3</sup>.

The  $v^m d^n$  matrix was chosen as one of the simple generalized matrices that include the distance matrix, the reverse distance matrix, and the adjacency matrix, as well as several variants of these. Formally it presents the influence of vertex  $j$  on vertex  $i$ , where  $v_j$  is one of the properties of vertex  $j$  that influence vertex  $i$  through the distance  $d_{ij}$ . In our case,  $v_i$  is set to the unit value. More general would be the study of indices derived from a matrix having the elements  $\mathbf{g}_{ij} (i \neq j) = v_i^a \times v_j^b \times d_{ij}^c$ , where  $v_i^a \times v_j^b$  represents the mutual interaction of a property of both vertices. A representative of this group of indices is one of the most useful indices, the Randić<sup>14</sup> index  $\chi$ , having the exponents  $\mathbf{a} = -1/2$ ,  $\mathbf{b} = -1/2$ ,  $\mathbf{c} = -\infty$ , whereas its zero order variant, the index  ${}^0\chi$ , has the exponents  $\mathbf{a} = -1/2$ ,  $\mathbf{b} = 0$ ,  $\mathbf{c} = -\infty$ . In general, the results of this type of the matrix would contain an additional dimension, and it was expected that the interpretation would be much more difficult than that of the results derived from the  $v^m d^n$  matrix. We therefore decided to study the simpler system, the  $v^m d^n$  matrix in this step.

The  $v^m d^n$  matrix is illustrated in Fig. 1, taking 2,2,3-trimethyl butane (223M4) as an example and exponents  $\mathbf{m} = 1$  and  $\mathbf{n} = 1$ . In 223M4 the vertices No. 1, 4, 5, 6, and 7 are of degree one and represent primary carbons (methyl groups). Vertex No. 2 is of degree four and represents a quaternary carbon. Vertex No. 3 is of degree three and represents a tertiary carbon. In 223M4 there is no vertex of degree two that would represent a secondary carbon (a methylene group).

#### *The $W(m,n)$ indices*

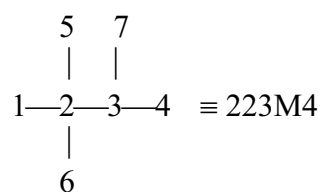
The indices derived from the generalized matrices introduced above can be e.g. those of the Wiener-type<sup>15</sup> or those based on the largest eigenvalues of the matrices. Wiener<sup>16</sup> index is an index derived from the distance matrix. It is defined<sup>17</sup> by Eq. 3.

$$W = \frac{1}{2} \sum d_{ij} \quad (3)$$

In the present paper we consider the indices of Wiener-type, derived from the  $v^m d^n$  matrix presented above. The Wiener type of indices was chosen since this type of indices is one of the most popular and one of the easiest to derive. We label them as  $W(m,n)$  indices and they are defined by Eq. 4:

$$W(m,n) = \frac{1}{2} \sum v_j^m \times d_{ij}^n \quad (4)$$

Several indices of this type are already known. We present here a few examples of these: the Wiener<sup>17</sup> index  $W \equiv W(0,1)$ , the Harary<sup>18,19</sup> index  $H \equiv W(0,-1)$ , and some Ivanciuc<sup>13</sup> indices, VTI-*i*. The position of these indices in the plane of exponents  $\mathbf{m}$  and  $\mathbf{n}$  used in order to derive the  $W(m,n)$  indices tested here is indicated in Fig. 30 in italics.



$d_{ij}$  matrix

$$\begin{array}{|c|c|c|c|c|c|c|}
 \hline
 0 & 1 & 2 & 3 & 2 & 2 & 3 \\
 1 & 0 & 1 & 2 & 1 & 1 & 2 \\
 2 & 1 & 0 & 1 & 2 & 2 & 1 \\
 3 & 2 & 1 & 0 & 3 & 3 & 2 \\
 2 & 1 & 2 & 3 & 0 & 2 & 3 \\
 2 & 1 & 2 & 3 & 2 & 0 & 3 \\
 3 & 2 & 1 & 2 & 3 & 3 & 0 \\
 \hline
 \end{array}$$

$$\begin{array}{|c|c|c|c|c|c|c|}
 \hline
 1 & 4 & 3 & 1 & 1 & 1 & 1 \\
 \hline
 \end{array}$$

$v_j$  vector

$v^1 d^1$  matrix

$$\begin{array}{|c|c|c|c|c|c|c|}
 \hline
 0 & 4 & 6 & 3 & 2 & 2 & 3 \\
 1 & 0 & 3 & 2 & 1 & 1 & 2 \\
 2 & 4 & 0 & 1 & 2 & 2 & 1 \\
 3 & 8 & 3 & 0 & 3 & 3 & 2 \\
 2 & 4 & 6 & 3 & 0 & 2 & 3 \\
 2 & 4 & 6 & 3 & 2 & 0 & 3 \\
 3 & 8 & 3 & 2 & 3 & 3 & 0 \\
 \hline
 \end{array}$$

**Fig. 1.** 2,2,3-trimethyl butane: its formula, its simple distance matrix composed of distances  $d_{ij}$ , its vector of degrees of vertices,  $v_j$ , and the  $v^1 d^1$  matrix.

We tested the characteristics of indices of the Wiener-type derived from the more general type of matrices having the elements  $g_{ij} (i \neq j) = v_j^m \times d_{ij}^n$  for the values of  $\mathbf{m}$  and  $\mathbf{n}$  that are presented in Fig. 2. Among the infinite number of possible exponents those presented in Fig. 2 were chosen for several reasons. Some of them, e.g.  $\mathbf{m} = 0, \mathbf{n} = 1$  or

-1, give rise to some popular indices. Exponents 2 and 3 define an (apparent) surface or an (apparent) volume. Some of the negative exponents have some physical meaning. For example, the exponents -1 and -2 are observed in electrostatic interactions where they describe the dependence of interaction on the distance of charges. The exponent -6 is observed in the Lennard-Jones potential where it describes the dependence of the influence of dispersion forces on distance. The exponent  $-\infty$  was chosen as the lowest exponent possible. It was expected that it might give rise to some simple and easily comprehensible results.

Other exponents were chosen to obtain some information on what these types of exponents offer. Some exponents higher than 3 were tested, too, but they were found to over-exaggerate some details and for this reason they are not included here.

There seems to be some disagreement about the meaning of the distance raised to a negative exponent. Schultz and Schultz<sup>20</sup>, for example, state that there are no constant reciprocal edge count intervals between vertices of reciprocal distance matrices since the reciprocals become increasingly smaller in value as the corresponding integers increase in size. Our perception is that the bond lengths in molecules remain the same regardless of the exponent we use. The exponent does not influence the distance. It shows the dependence of an interaction on distance. From this point of view the  $v^m d^n$  matrix is a "vertex-property interaction over vertex-distance" matrix.

### *Characteristics of $W(m,n)$ indices*

The indices  $W(m,n)$  of methane are equal to zero since  $g_i = 0$  by definition. The  $W(m,n)$  indices of ethane, being equal to  $\frac{1}{2} \times (1^m \times 1^n + 1^m \times 1^n)$ , are in any case equal to 1. Other Wiener-type indices,  $W(m,n)$ , derived from the generalised distance matrix having the non-diagonal elements  $v_j^m \times d_{ij}^n$ , have the following characteristics.

#### *$W(m,n)$ indices which are integers*

The  $W(m,n)$  indices which are integers are presented in Fig. 2, including those that occur as halves of integers. The former ones occur when the exponents are integers, i.e. when  $m$  and  $n$  are integers and  $\geq 0$ , as well as when  $m$  is an integer,  $\geq 0$ , and  $n = -\infty$ . The latter ones occur when  $n$  is an integer,  $\geq 0$ , or equal  $-\infty$  and  $m = -\infty$ , as well as when  $m = -1$  and  $n = -\infty$ . In view of these facts it is unclear whether  $W(m,n)$  should rather be

<b>m</b>															
<b>3</b>	I								I				I	I	I
<b>2</b>	I								I				I	I	I
<b>1</b>	I								I				I	I	I
$\frac{1}{2}$															
$\frac{1}{3}$															
$\frac{1}{4}$															
<b>0</b>	I								I				I	I	I
$-\frac{1}{4}$															
$-\frac{1}{3}$															
$-\frac{1}{2}$															
<b>-1</b>	$I/2$														
<b>-2</b>															
<b>-4</b>															
<b>-6</b>															
<b><math>-\infty</math></b>	$I/2$								$I/2$				$I/2$	$I/2$	
	<b><math>-\infty</math></b>	<b>-6</b>	<b>-4</b>	<b>-2</b>	<b>-1</b>	$-\frac{1}{2}$	$-\frac{1}{3}$	$-\frac{1}{4}$	<b>0</b>	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	<b>1</b>	<b>2</b>	<b>3</b>
	<b>n</b>														

Fig. 2.  $W(m,n)$  indices which are integers (I) or halves of integers ( $I/2$ ).

<b>m</b>															
<b>3</b>	24								24						
<b>2</b>	24								25				8	2	
<b>1</b>	31								37				8	2	
$\frac{1}{2}$	24								24						
$\frac{1}{3}$	24								24						
$\frac{1}{4}$	24								24						
<b>0</b>	37				2				37				8		
$-\frac{1}{4}$	24								24						
$-\frac{1}{3}$	24								24						
$-\frac{1}{2}$	24								24						
<b>-1</b>	37								24				10	2	
<b>-2</b>	24			2	2				24						
<b>-4</b>	24								24						
<b>-6</b>	24								24						
<b><math>-\infty</math></b>	37			4	6				29				8	6	
	<b><math>-\infty</math></b>	<b>-6</b>	<b>-4</b>	<b>-2</b>	<b>-1</b>	$-\frac{1}{2}$	$-\frac{1}{3}$	$-\frac{1}{4}$	<b>0</b>	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	<b>1</b>	<b>2</b>	<b>3</b>
	<b>n</b>														

Fig. 3. Degeneration of  $W(m,n)$  indices for alkanes from propane through octanes expressed as the number of data equal to at least one more among the 38 possible ones.



defined as  $W(m,n) = \sum v_j^m \times d_{ij}^n$  and not as  $W(m,n) = \frac{1}{2} \sum v_j^m \times d_{ij}^n$ . We would tend to choose the former possibility in spite of being aware of the long tradition of using the latter one.

*The degeneration of  $W(m,n)$  indices*

How many data of the tested  $W(m,n)$  indices are degenerated, i.e. equal to at least one more, is presented in Fig. 3 for all alkanes from propane through octanes. Highly degenerated are the indices  $W(m,-\infty)$  and  $W(m,0)$ , and some degeneration is observed also among some of the indices of the groups  $W(m,1) > W(m,2) > W(m,-1) > W(m,-2)$ . The "degeneration causing" exponents are  $\mathbf{m} = -\infty, -2, -1, 0, 1$ , and  $2$ , as well as  $\mathbf{n} = -\infty, -2, -1, \mathbf{0}, 1$ , and  $2$ . The unconditionally degeneration causing exponents are presented in bold. A substantial part of degeneration is observed among isomers.

*The increase of values of  $W(m,n)$  indices of  $n$ -alkanes due to the increase of size of the molecule*

How the values of  $W(m,n)$  indices for  $n$ -alkanes,  $W(m,n)_n$ , increase when the size of the molecule increases, i.e. when the carbon number ( $N_C$ ) of  $n$ -alkanes increases from  $C_2$  to  $C_{15}$ , is indicated in Fig. 4. None of them decreases with increasing carbon number. The slope of linearly increasing  $W(m,n)_n$  indices, i.e. of  $W(m,-\infty)_n$ , is equal to  $2^m$ . If  $\mathbf{n} \geq 0$ , the dependence is parabolic. It is presented in Fig. 4 as the degree of the simplest polynomial having  $r^2 > 0.99995$ . If  $-\infty < \mathbf{n} < 0$ , the dependence is hyperbolic. The asymptote of hyperbola is in general a parabola. At very low  $\mathbf{n}$  and high carbon number the hyperbola asymptotically approaches an almost linear increase. The form of hyperbola depends on the combination of  $\mathbf{m}$  and  $\mathbf{n}$ . It is indicated in Fig. 4 by the symbols  $+$ ,  $+/-$ , and  $-$ . These symbols are in Fig. 5 and 6 in parentheses, in order to be more visible. Fig. 5 presents the typical hyperbolas, and their properties are illustrated in Fig. 6, showing the increments of  $W(m,n)_n$  with  $N_C$  ( $N_v$ ).

The hyperbolas marked in Fig. 4-6 by  $(+)$  increase relatively fast, while their increments are increasing slower and slower. The increments of hyperbolas marked in Fig. 4-6 by  $(-)$  are decreasing to a constant value. The increments of hyperbolas marked in Fig. 4-6 by  $(+/-)$  decrease at low  $N_C$  and then increase to an upper limit.

<b>m</b>																
<b>3</b>	1	+	+	+	+	+	+	+	+	2	3	3	3	3	4	4
<b>2</b>	1	+	+	+	+	+	+	+	+	2	3	3	3	3	4	4
<b>1</b>	1	+	+	+	+	+	+	+	+	2	3	3	3	3	4	4
$\frac{1}{2}$	1	+	+	+	+	+	+	+	+	2	3	3	3	3	4	4
$\frac{1}{3}$	1	+	+	+	+	+	+	+	+	2	3	3	3	3	4	4
$\frac{1}{4}$	1	+	+	+	+	+	+	+	+	2	3	3	3	3	4	4
<b>0</b>	1	+	+	+	+	+	+	+	+	2	3	3	3	3	4	4
$-\frac{1}{4}$	1	+/-	+	+	+	+	+	+	+	2	3	3	3	3	4	4
$-\frac{1}{3}$	1	+/-	+/-	+	+	+	+	+	+	2	3	3	3	3	4	4
$-\frac{1}{2}$	1	+/-	+/-	+	+	+	+	+	+	2	3	3	3	3	4	4
<b>-1</b>	1	+/-	+/-	+/-	+	+	+	+	+	2	3	3	3	3	4	4
<b>-2</b>	1	-	-	+/-	+/-	+	+	+	+	2	2	2	3	3	4	4
<b>-4</b>	1	-	-	-	-	+/-	+/-	+/-	+	2	2	2	2	3	3	4
<b>-6</b>	1	-	-	-	-	-	-	-	+	2	2	2	2	2	3	4
<b><math>-\infty</math></b>	1(0)	-	-	-	-	-	-	-	-	2(1)	2	2	2	2	3	4
	<b><math>-\infty</math></b>	<b>-6</b>	<b>-4</b>	<b>-2</b>	<b>-1</b>	$-\frac{1}{2}$	$-\frac{1}{3}$	$-\frac{1}{4}$	<b>0</b>	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	<b>1</b>	<b>2</b>	<b>3</b>	
									<b>n</b>							

**Fig. 4.** The degree of the simplest polynomial describing well ( $r^2 > 0.99995$ ) the dependence of  $W(m,n)$  of  $n$ -alkanes ( $W(m,n)_n$ ) on  $N_C$ .  $n$ -Alkanes from  $C_2$  to  $C_{15}$  were considered.

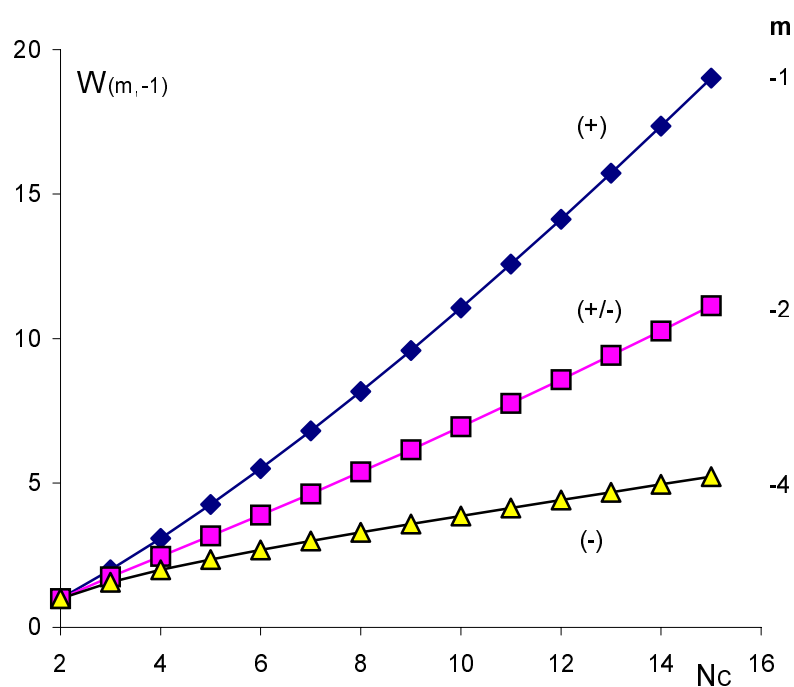
+: The increment of increase of  $W(m,n)_n$  with  $N_C$  increases and approaches a constant value

-: The increment of increase of  $W(m,n)_n$  with  $N_C$  decreases and approaches a constant value

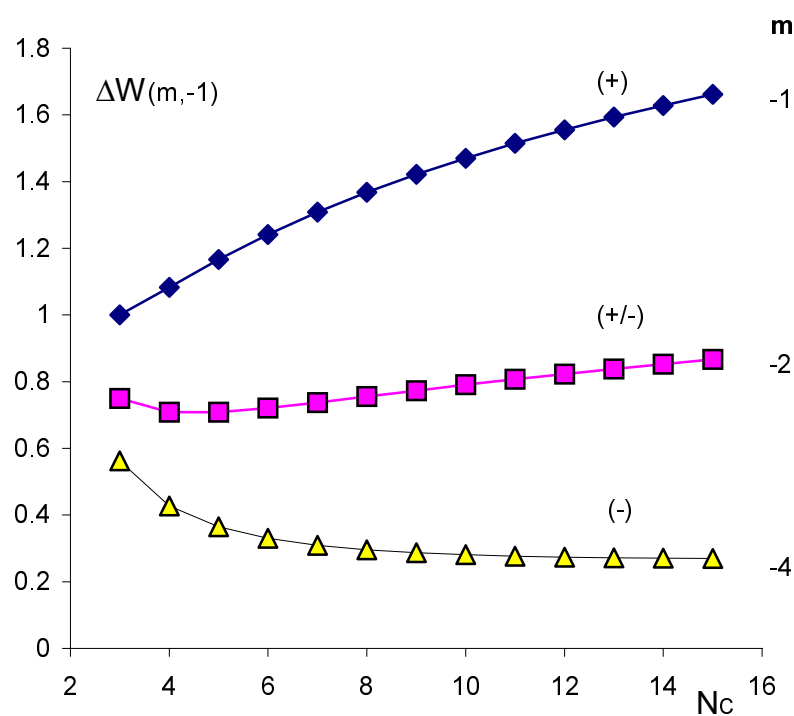
+/-: The increment of increase of  $W(m,n)_n$  with  $N_C$  decreases at low  $N_C$  and then increases slightly, approaching a constant value

In parentheses: the apparent exponent of the polynomial

Another possibility of presenting the influence of exponents **m** and **n** on the dependence of  $W(m,n)$  indices of  $n$ -alkanes due to the increase in carbon number is to use the susceptibility for the increase of carbon number by one,  $S_{i,i+1}$ , cf. ref.<sup>7</sup> This possibility is presented in Fig. 7 with  $S_{7,8}$ , i.e. for the increase of carbon number from 7 to 8. It can be seen that  $W(-\infty,-\infty)_n$  does not increase with increasing carbon number. The dependence on the exponent of the degree of vertex, **m**, is sigmoid-like, whereas the dependence on the exponent of distance, **n**, is parabolic. The highest increase with carbon number has the index  $W(3,3)_n$ .

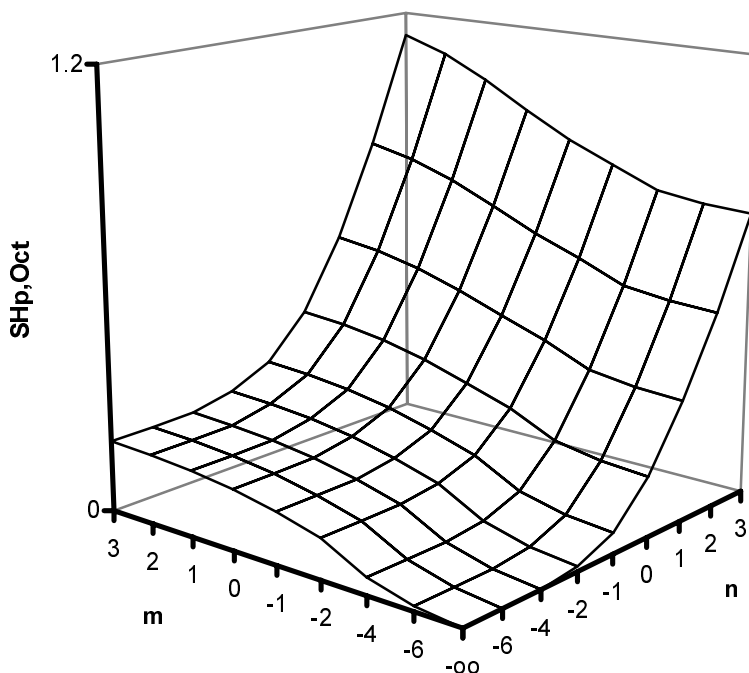


**Fig. 5.** The dependence of  $W(m,-1)_n$  on  $N_C$  ( $N_v$ ) in the transition region marked in Fig. 4 by symbols +, +/-, and -, i.e. for  $m = -1, -2,$  and  $-4$ .



**Fig. 6.** The increments of increase of  $W(m,-1)$  with  $N_C$  ( $N_v$ ), illustrating of the meaning of symbols +, +/-, and - in Fig. 4.

$$\Delta W(m,-1) = W(m,-1)_{N_C} - W(m,-1)_{N_C-1}.$$



**Fig. 7.** The dependence of susceptibility for the increase in carbon number from 7 to 8,  $S_{Hp,Oct}$ , of  $W(m,n)_n$  on exponents  $m$  and  $n$ .

*The increase of values of  $W(m,n)$  indices due to the increase of the size of molecules of other alkane isomers*

How the values of  $W(m,n)$  indices of other isomers increase with the size of a molecule is the most easily presented by the susceptibility of typical structural types for the increase in carbon number by one,  $S_{i,i+1}$ , cf. ref.<sup>7</sup> We consider  $S_{5,6}$ ,  $S_{6,7}$ , and  $S_{7,8}$ .

*The branched structure is shifted to the periphery of the molecule*

The following transitions were considered first:

In  $S_{5,6}$ : Pe  $\rightarrow$  Hx, 2M4  $\rightarrow$  2M5, 22M3  $\rightarrow$  22M4

In  $S_{6,7}$ : Hx  $\rightarrow$  Hp, 2M5  $\rightarrow$  2M6, 3M5  $\rightarrow$  3M6, 23M4  $\rightarrow$  23M5, 22M4  $\rightarrow$  22M5

In  $S_{7,8}$ : Hp  $\rightarrow$  Oct, 2M6  $\rightarrow$  2M7, 3M6  $\rightarrow$  3M7, 3Et5  $\rightarrow$  3Et6, 24M5  $\rightarrow$  24M6, 23M5  $\rightarrow$  23M6, 22M5  $\rightarrow$  22M6, 33M5  $\rightarrow$  33M6, and 223M4  $\rightarrow$  223M5.

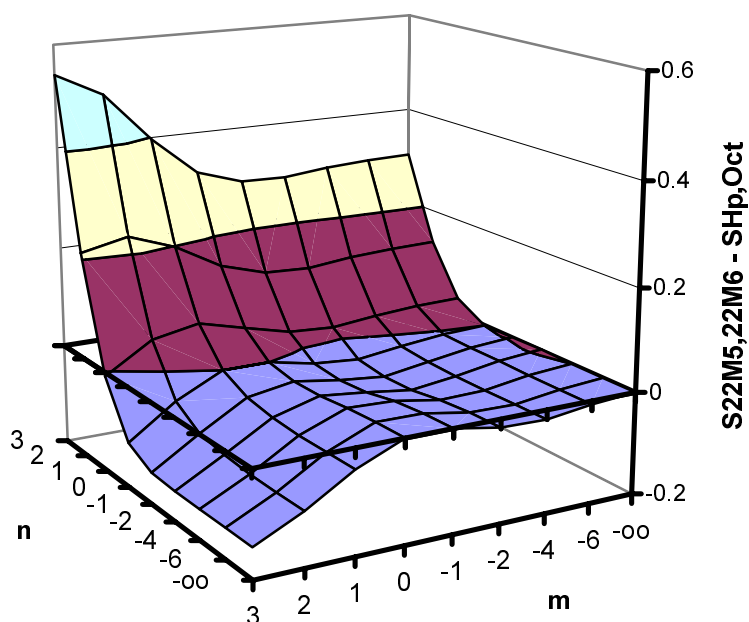
In these transitions the main chain of the alkane is elongated by insertion of a methylene group, so that the branched structure is preserved as much as possible although it is shifted to the periphery of the molecule. Since the influence of elongation of the  $n$ -alkane is presented in Fig. 7, only the difference between  $S_{i,i+1}$  of other alkane

isomers and  $S_{i,i+1}$  of  $n$ -alkanes is presented in Fig. 8-12. Fig. 8 illustrates  $S_{7,8}$ , i.e. the case when carbon number increases from 7 to 8, as exemplified by alkanes 22M5 and 22M6. A similar shape of the difference of  $S_{i,i+1}$  and  $S_{\text{Hp,Oct}}$  as the structure 22M presented in Fig. 8 can be seen in the structures 2M, 3M, 3Et, 23M, 33M, and 223M. In Fig. 8 it can be seen that for the majority of isomers the  $W(m,n)$  increases with carbon number at  $n \geq 0$  faster than  $W(m,n)_n$ . On the other hand, at lower  $n$  and especially when  $-\infty < m < -1$  or when  $m > 0$ , the  $W(m,n)$  indices of other isomers increase with carbon number slower than  $W(m,n)_n$ .

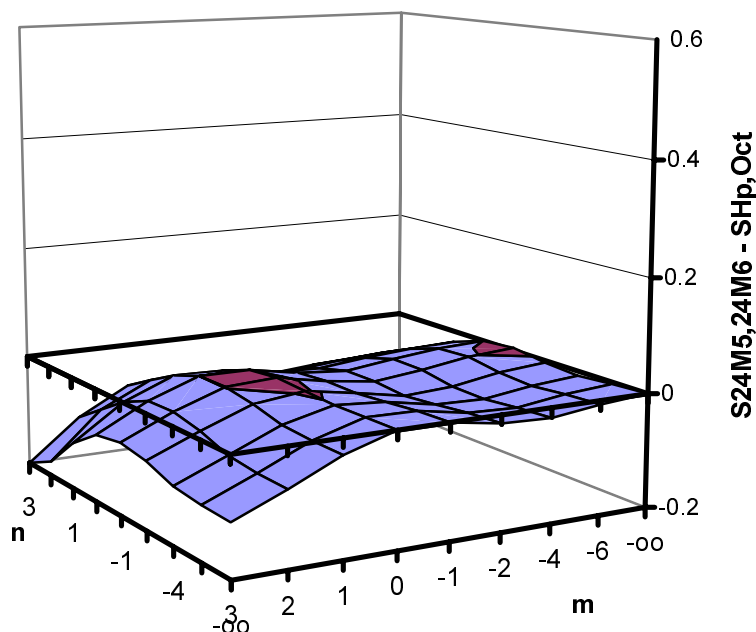
Only the structure 24M, Fig. 9, gives rise to a radically different shape, which indicates that in most cases  $W(m,n)_{24M}$  increases with carbon number slower than  $W(m,n)_n$ .

The highest differences compared to the increase of  $W(m,n)_n$  are at the following combinations of  $m$  and  $n$ , where the departure from the values of  $S_{\text{Hp,Oct}}$  is as follows,

$(m = 3, n = 3):$	$22M > 33M > 23M > 3M > 2M > 223M > 3Et > 24M$
$(m = 2, n = 3):$	$22M > 33M > 23M > 3M > 223M > 2M > 3Et > 24M$
$(m = 3, n = -\infty):$	$2M = 3M = 3Et > 23M = 24M > 22M = 33M > 223M$



**Fig. 8.** The difference of susceptibility for the increase in carbon number of  $W(m,n)$  indices of 22M5 and of  $n$ -heptane.



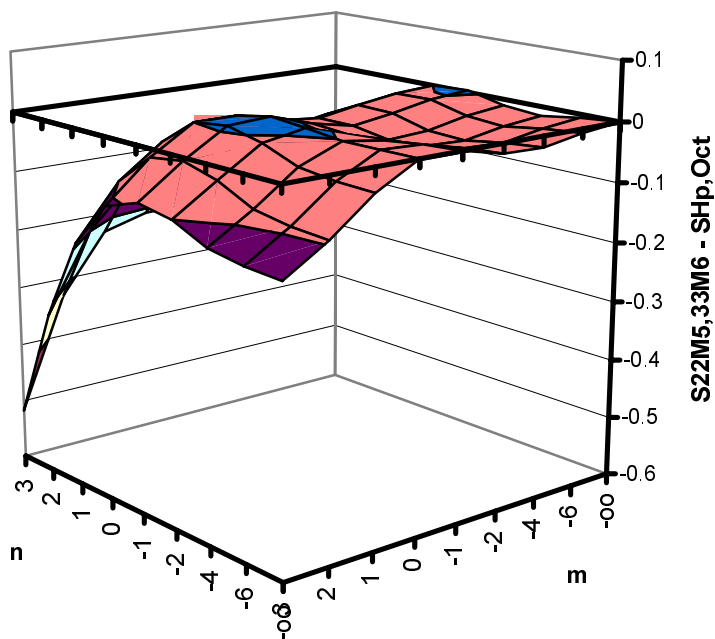
**Fig. 9.** The difference of susceptibility for the increase in carbon number of  $W(m,n)$  indices of 24M5 and of  $n$ -heptane.

As a consequence, the lowest increase with carbon number of all tested  $W(m,n)$  indices is observed at  $W(3,-\infty)$ . The magnitude of increase of this index on increasing carbon number indicates that the more branched the structure, the lower is the increase of  $W(3,-\infty)$ . Besides the number of branches, only the type of branched structure has some influence: tertiary carbons have a lesser influence than the quaternary ones.

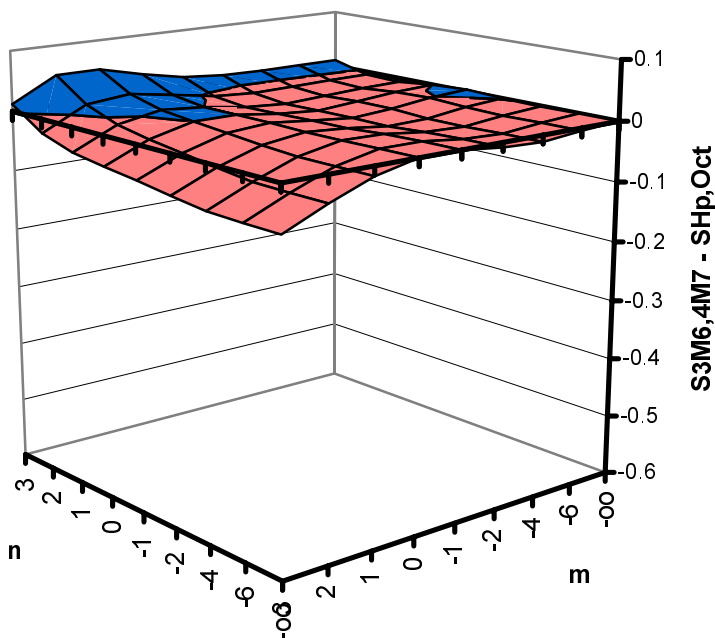
*The branched structure is shifted to the centre of the molecule*

There are also other possibilities to increase the carbon number by 1. For example, the insertion of a methylene group can be performed by shifting the branched structure more to the centre of the molecule:  $2M6 \rightarrow 3M7$ ,  $3M6 \rightarrow 4M7$ ,  $23M5 \rightarrow 34M6$ ,  $22M5 \rightarrow 33M6$ , and  $223M4 \rightarrow 233M5$ .

The influence of increasing the size of molecule in this way is presented in Fig. 10 with the conversion  $22M5 \rightarrow 33M6$  as a general example, and in Fig. 11 with  $3M6 \rightarrow 4M7$  as an exception. For most examples of exponents  $m$  and  $n$  presented in Fig. 10 the increase is less than that of  $W(m,n)_n$ . The magnitude of influence is as follows:  $22M5 \rightarrow 33M6 > 2M6 \rightarrow 3M7 > 223M4 \rightarrow 233M5 \sim 23M5 \rightarrow 34M6$ .



**Fig. 10.** The influence of increase of carbon number by one, i.e. from 7 to 8, on  $W(m,n)$  indices, when the branched structure is shifted simultaneously to the centre of the molecule. Presented is the difference between  $S_{i,i+1}$  of other alkane isomers and  $S_{i,i+1}$  of  $n$ -alkanes.



**Fig. 11.** The influence of increase of carbon number by one, i.e. from 7 to 8, on  $W(m,n)$  indices for the conversion of 3M6 into 4M7. Presented is the difference between  $S_{i,i+1}$  of other alkane isomers and  $S_{i,i+1}$  of  $n$ -alkanes.

The (slightly) positive extremes of influence are observed at  $W(-\infty, -1)$  and  $W(0, -1)$ . The negative extremes of influence are observed at  $W(3, 3)$  and are less expressed at  $W(3, -\infty)$ . Of similar shape and magnitude is also the influence of elongation of the existing branch (in our case, the conversion of methyl into ethyl at  $3M5 \rightarrow 3Et5$ ,  $23M5 \rightarrow 3Et2M5$ , and  $33M5 \rightarrow 3Et3M5$ ). On the other hand, at the conversion of  $3M6 \rightarrow 4M7$ , which is an exception, the positive extremes of influence are observed at  $n = 3$ , especially at  $W(1, 3)$ , but also at  $W(-\infty, -1)$ . The negative extreme of influence is observed at  $W(3, -\infty)$ . All of them are less expressed than in the previous cases.

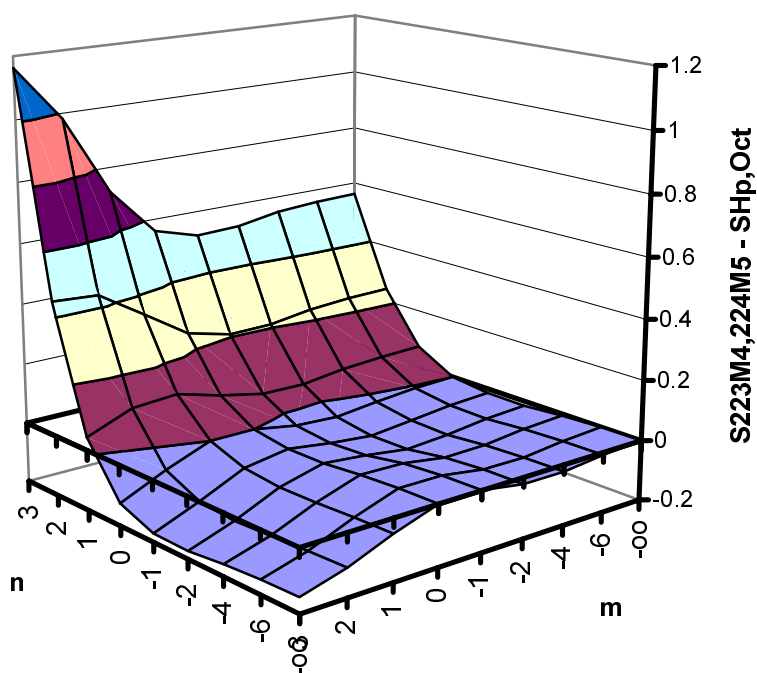
*The branched structure is shifted apart*

A methylene group can be also inserted by shifting the branched structures apart, so that there is simultaneously a shift towards the periphery of the molecule as well as an increase in the distance between the branches. This happens in conversions  $23M5 \rightarrow 24M6$ ,  $24M5 \rightarrow 25M6$ , and  $223M4 \rightarrow 224M5$ . The last of them is presented in Fig. 12. At  $n \geq 0$  the increase of  $W(m, n)$  is higher than that of  $W(m, n)_n$ . The magnitude of this type of influence is as follows:  $223M4 \rightarrow 224M5 \gg 23M5 \rightarrow 24M6 > 24M5 \rightarrow 25M6$ . The extreme of the increase of  $W(m, n)$  with carbon number is in this case at  $W(3, 3)$ , where the additional contribution of the change in the structure  $223M4 \rightarrow 224M5$  is nearly equal to that of the elongation of  $n$ -heptane into  $n$ -octane. At lower values of  $n$ , i.e. when  $n < 0$ , the total increase of the  $W(m, n)$  is lower than that of  $W(m, n)_n$ . The extreme is at  $W(3, -1)$  or  $W(3, -2)$ .

Altogether, the lowest increase of a  $W(m, n)$  index with carbon number is observed for conversions  $3Et5 \rightarrow 3Et6$ ,  $23M5 \rightarrow 23M6$ , and  $33M5 \rightarrow 33M6$  at the index  $W(3, -6)$ , whereas for the conversion  $3M6 \rightarrow 3M7$  it is observed at the index  $W(3, -4)$ , and for the conversion of  $22M5 \rightarrow 22M6$  at the index  $W(3, -2)$ . At these  $W(m, n)$  indices the series of increase of  $W(m, n)$  with carbon number is as follows:

$W(3, -\infty)$ :  $2M = 3M = 3Et > 23M = 24M > 33M = 22M > 223M$   
 $W(3, -6)$ :  $2M > 3M > 3Et > 23M > 24M > 33M > 22M > 223M$   
 $W(3, -4)$ :  $2M > 3Et > 3M > 23M > 24M > 33M > 22M > 223M$   
 $W(3, -2)$ :  $3Et > 2M > 3M > 23M > 24M > 223M > 33M > 22M$





**Fig. 12.** The influence of increase of carbon number by one, i.e. from 7 to 8, on  $W(m,n)$  indices, when the distance between the branches increases. Presented is the difference between  $S_{i,i+1}$  of other alkane isomers and  $S_{i,i+1}$  of  $n$ -alkanes.

These sequences of increase of  $W(m,n)$  indices on increasing carbon number indicate that as the exponent  $n$  increases from  $-\infty$ , the number of branches as well as their central position have less influence, whereas other structural features have more influence. The other ways of increase of carbon number by one, cf. Fig. 9-12, indicate that the influence is stronger when the branches are more central and especially when they are shifted to the centre, as well as when the branch bearing carbon becomes quaternary. When  $n \geq 0$ , then the increase is greater than in the case of  $n$ -alkanes. The contrary can be observed when the branches are shifted into a more peripheral position.

If we compare the above figures, especially Fig. 10 with Fig. 12, we can see that diverse structural features have different contributions to various  $W(m,n)$  indices. On the other hand, Nikolić et al.<sup>5</sup> have shown that a combination of different indices often has a better correlation with a particular physicochemical property than each index itself. From the above observations we can conclude that if a physicochemical property contains a contribution of particular structural features, which is different from the

contribution contained in each one of the used indices, then a combination of indices can fit the contribution contained in a property better than each index separately.

*The sequences of isomers due to the increase of the size of molecule*

Regarding the influence of increase in carbon number, there is another question: what are the sequences of isomers depending on exponents  $m$  and  $n$  when the carbon number is increased by one? Fig. 13 presents the situation in the case of  $S_{5,6}$ . There are six possible sequences of  $S_{5,6}$  data and all of them are observed among the  $W(m,n)$  indices.

<b>m</b>																		
<b>3</b>	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$C$	$D$			
<b>2</b>	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$C$	$D$	$D$		
<b>1</b>	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$E$	$E$	$E$	$=$	$D$	$D$	$D$	$D$	$D$	$D$	$D$		
$1/2$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$E$	$B$	$B$	$B$	$B$	$A\uparrow$	$A\uparrow$	$A\uparrow$	$A\uparrow$	$D$	$D$	$D$	$D$		
$1/3$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$E$	$B$	$B$	$B$	$B$	$A\uparrow$	$A\uparrow$	$A\uparrow$	$D$	$D$	$D$	$D$	$D$		
$1/4$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$E$	$B$	$B$	$B$	$B$	$B$	$A\uparrow$	$D$	$D$	$D$	$D$	$D$	$D$		
<b>0</b>	$=$	$A\downarrow$	$A\downarrow$	$E$	$B$	$B$	$B$	$B$	$=$	$C$	$C$	$C$	$C$	$C$	$D$	$D$		
$-1/4$	$A\uparrow$	$A\uparrow$	$A\downarrow$	$A\downarrow$	$E$	$E$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$C$	$C$	$C$	$D$	$D$
$-1/3$	$A\uparrow$	$A\uparrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$C$	$C$	$D$	$D$
$-1/2$	$A\uparrow$	$A\uparrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$B$	$D$	$D$
<b>-1</b>	$=$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$B$	$D$	$D$
<b>-2</b>	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$B$	$D$	$D$
<b>-4</b>	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$B$	$D$	$D$
<b>-6</b>	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$E$	$E$	$E$	$E$	$E$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$A\downarrow$	$D$	$B$	$D$	$D$	
<b><math>-\infty</math></b>	$0$	$A\downarrow$	$A\downarrow$	$B$	$A\uparrow$	$A\uparrow$	$B$	$B$	$=$	$B$	$B$	$B$	$B$	$F$	$D$	$D$		
	<b><math>-\infty</math></b>	<b>-6</b>	<b>-4</b>	<b>-2</b>	<b>-1</b>	$-1/2$	$-1/3$	$-1/4$	<b>0</b>	$1/4$	$1/3$	$1/2$	<b>1</b>	<b>2</b>	<b>3</b>			
	<b>n</b>																	

**Fig. 13.** The position of various sequences of  $S_{5,6}$  of  $W(m,n)$  indices in the plane of exponents  $m$  and  $n$ . Sequences:

$A\uparrow$ :  $n < 2M < 22M$      $A\downarrow$ :  $22M < 2M < n$      $B$ :  $2M < n < 22M$      $C$ :  $22M < n < 2M$

$D$ :  $n < 22M < 2M$      $E$ :  $2M < 22M < n$

$0$ : All  $S_{5,6} = 0$ .

$=$ : All  $S_{5,6}$  are equal and  $> 0$ .

$\pm$ : Some of  $S_{5,6}$  are negative and some positive.

By intuition, only the sequences  $A\downarrow$  and  $A\uparrow$  in Fig. 13 seem to be regular. In the sequence  $A\uparrow$ , the  $W(m,n)$  indices increase with carbon number more when the structure is more branched. In the sequence  $A\downarrow$  the reverse is true.

<b>m</b>															
<b>3</b>	°↓	A↓	D↓	D↓	D↓	D↓	D↓	D↓	°↓	A↓	A↓	B↓	B↓	p4	p4
<b>2</b>	°↓	A↓	A↓	D↓	D↓	D↓	D↓	D↓	°↓	B↓	B↓	B↓	p4	p4	p4
<b>1</b>	°↓	A↓	A↓	c6	c6	c6	c6	c6	=	p4	p4	p4	p4	p5	p5
<sup>1</sup> / <sub>2</sub>	°↓	A↓	A↓	c6	c4	c4	c4	c2	°↑	p5	p5	p5	p4	p4	p5
<sup>1</sup> / <sub>3</sub>	°↓	A↓	D↓	c6	c4	c4	<b>c1</b>	c2	°↑	p5	p5	p4	p4	p4	p5
<sup>1</sup> / <sub>4</sub>	°↓	A↓	D↓	c4	c4	c4	c3	<b>c1</b>	°↑	p5	p5	p4	p4	p4	p5
<b>0</b>	=	A↓	D↓	c4	c4	c4	c4	c4	=	p4	p4	p4	p4	p4	p4
<sup>-1</sup> / <sub>4</sub>	°↑	A↑	c7	c6	c6	c6	c6	c6	°↓	p3	<b>p2</b>	p4	p4	p4	p4
<sup>-1</sup> / <sub>3</sub>	°↑	A↑	D↓	c6	c6	c6	c7	c7	°↓	B↓	p3	<b>p2</b>	p4	p4	p4
<sup>-1</sup> / <sub>2</sub>	°↑	A↑	D↓	c7	c6	c7	D↓	D↓	°↓	B↓	B↓	p3	p4	p4	p4
<b>-1</b>	=	F↓	D↓	D↓	D↓	D↓	D↓	D↓	°↓	B↓	B↓	B↓	<b>p1</b>	p4	p4
<b>-2</b>	°↓	D	F↓	D↓	D↓	D↓	D↓	D↓	°↓	B↓	B↓	B↓	<b>p1</b>	p4	p4
<b>-4</b>	°↓	B↓	B↓	B↓	D↓	D↓	D↓	D↓	°↓	B↓	B↓	p6	<b>p1</b>	p4	p5
<b>-6</b>	°↓	B↓	B↓±	B↓	c5	c6	c6	c6	°↓	p6	p3	<b>p2</b>	p4	p4	p5
<b>-∞</b>	0	B↓±	p3±	p4	B↑	c2	c2	<b>c1</b>	=	<b>p2</b>	<b>p2</b>	p4	p4	p4	p5
	<b>-∞</b>	<b>-6</b>	<b>-4</b>	<b>-2</b>	<b>-1</b>	<sup>-1</sup> / <sub>2</sub>	<sup>-1</sup> / <sub>3</sub>	<sup>-1</sup> / <sub>4</sub>	<b>0</b>	<sup>1</sup> / <sub>4</sub>	<sup>1</sup> / <sub>3</sub>	<sup>1</sup> / <sub>2</sub>	<b>1</b>	<b>2</b>	<b>3</b>
	<b>n</b>														

**Fig. 14.** The position of sequences of  $S_{6,7}$  of  $W(m,n)$  indices in the plane of exponents  $m$  and  $n$ .

Sequences:

$A\uparrow$ :  $n < 2M < 3M < 23M < 22M$        $A\downarrow$ :  $22M < 23M < 3M < 2M < n$   
 $B\uparrow$ :  $n < 2M < 3M < 22M < 23M$        $B\downarrow$ :  $23M < 22M < 3M < 2M < n$   
 $D\downarrow$ :  $22M < 23M < 2M < 3M < n$        $F\downarrow$ :  $23M < 22M < 3M < 2M < n$   
 $G$ :  $n < 3M < 2M < 23M < 22M$  - not observed  
**c1**:  $2M < 22M < n < 3M < 23M$       **c2**:  $2M < n < 22M < 3M < 23M$   
**c3**:  $2M < 22M < 3M < n < 23M$       **c4**:  $22M < 2M < 3M < n < 23M$   
**c5**:  $2M < 22M < 3M < 23M < n$       **c6**:  $22M < 2M < 3M < 23M < n$   
**c7**:  $22M < 2M < 23M < 3M < n$   
**p1**:  $23M < 3M < n < 2M < 22M$       **p2**:  $23M < 3M < n < 22M < 2M$   
**p3**:  $23M < 3M < 22M < n < 2M$       **p4**:  $23M < n < 3M < 2M < 22M$   
**p5**:  $n < 23M < 3M < 2M < 22M$       **p6**:  $23M < 3M < 22M < 2M < n$

**0**: All  $S_{6,7} = 0$ .

=: All  $S_{6,7}$  are equal and  $> 0$ .

±: Some of  $S_{6,7}$  are negative and some positive.

°: Degenerated cases:

↑ Increasing number of branches, especially on quaternary carbons, enhances the increase of  $S_{6,7}$  compared to the increase at n-alkane.

↓ Increasing number of branches, especially on quaternary carbons, diminishes the increase of  $S_{6,7}$  compared to the increase at n-alkane.

The  $S_{5,6}$  of the 21 considered physicochemical properties of alkanes belong to the following of the sequences presented in Fig. 13: none belongs to sequences *C* or *E*. To sequence  $A\downarrow$  belong:  $\Delta H_f^\circ$ , **Tc/Pc**, *Vm*, and *Vc*. To sequence  $A\uparrow$  belong: *BP*,  $\Delta H_v$ , **Tc**,

Pc(-),d, dc( $\pm$ ),  $n_D$ , and MR. To sequence C belong: **BP/Tc**,  $\omega$ ,  $V''$ , and  $\alpha_c$ , whereas to sequence B belong:  $Tc^2/Pc$ , A, B, C, and Zc. Regarding this fact and some previous<sup>7</sup> conclusions,  $\Delta H_f^\circ$  and Tc/Pc seem better reference properties than BP/Tc. BP and  $\Delta H_v$  follow them in this respect.

In the case of  $S_{6,7}$ , Fig. 14, a larger number of different sequences is possible than in the case of  $S_{5,6}$ . On the other hand, a lower number of tested combinations of exponents **m** and **n** gives rise to a regular sequence of isomers. They are presented as  $A \uparrow$  and  $A \downarrow$ . When **n** > 0, then in all tested cases the index increases with carbon number more if the branches are placed on the periphery of the molecule than if they are placed near the centre. When **n** < 0, then in the majority of cases the reverse is true. The exceptions are the indices in the area of  $-4 > \mathbf{m} > -\infty$ ,  $-2 > \mathbf{n} > -6$ . The extreme cases are presented in Fig. 14 in bold lowercase letters.

The same sequence as those in Fig. 14 have the following tested physicochemical properties. Sequence  $B \uparrow$ ,  $B \downarrow$ , and D: none; Sequence  $A \uparrow$ :  $\omega$ ; Sequence G: BP.

The situation is still more complicated in the case of  $S_{7,8}$  and in this case, Fig. 15, only the sequences seeming the most regular are presented. Only a few W(m,n) indices have a regular sequence of  $S_{7,8}$  data considered here. At **n** = 1, also some indices with degenerated  $S_{7,8}$  data can be observed. One of them is the Wiener<sup>15</sup> index  $W \equiv W(0,1)$ .

None of the tested properties gives rise to the sequences of  $S_{7,8}$  presented as "best" by intuition in Fig. 15. The properties  $n_D$ , MR,  $Tc^2/Pc$ ,  $\alpha_c$ ,  $\omega$ , A, B,  $\Delta H_v$ ,  $V_c$ , and BP even have no sequence of  $S_{7,8}$  equal to that of a tested W(m,n) index.

Figs. 13-15 also present another interesting fact. On increasing the size of a molecule of alkanes, the value of the majority of tested W(m,n) indices increases among all isomers. An exception are the W(m,n) indices in the region roughly represented by the following values of exponents **m** and **n**:  $-\infty \leq \mathbf{m} \leq -6$  and  $-\infty < \mathbf{n} \leq -4$ , where the W(m,n) indices of some isomers increase with the size of the molecule whereas the others decrease. The reasons for the decrease are the most easily seen on the increase of carbon number of alkanes from 5 to 6. The increase of the value of the W(m,n) index for *n*-alkane (i.e. when going from *n*-pentane to *n*-hexane) is very small. The decrease of the value of the W(m,n) index on increasing the number of branches is higher in value, so the sum is negative.



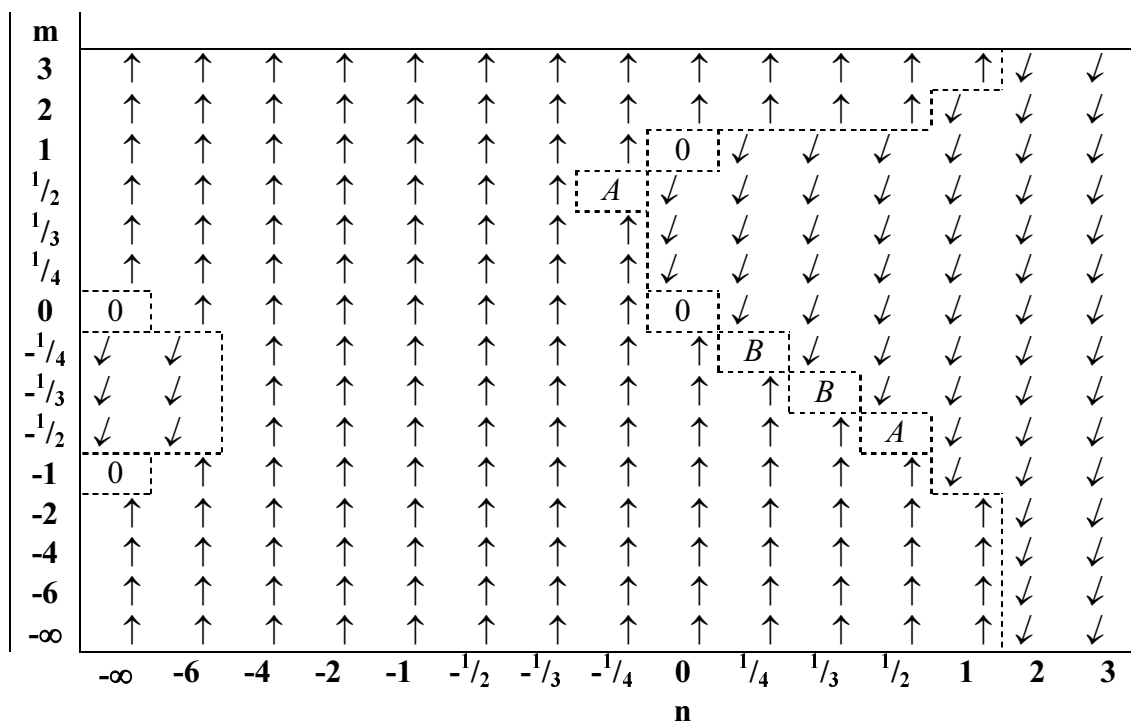


Fig. 16. The dependence of  $W(m,n)$  indices on the branching of pentanes.

↓  $W(m,n)$  decreases with branching

↑  $W(m,n)$  increases with branching

0  $W(m,n)$  presents only the size, not branching

Dotted lines - intermediate region between  $W(m,n)$  indices increasing with branching and those decreasing with branching. In this region some  $W(m,n)$  indices indicate that some isomers seem more branched than the most branched pentane, i.e. **22M3**, or less branched than the least branched pentane, i.e. **Pe**:

A **2M4** > **22M3** > **Pe**

B **22M3** > **Pe** > **2M4**

Three regions are indicated in the plane of exponents  $m$  and  $n$ . Most of the tested indices increase with branching. Most of them are observed at  $n < 0$ . Many of them increase also at  $0 < n < 1$ , especially if  $m > 1$  or  $m < 0$ . At  $n \geq 0$ , especially if  $-1 < m < 2$ , several  $W(m,n)$  indices decrease with branching, including all tested indices having  $n = 3$ . There is also a region of low  $n$ , i.e. when  $n < -4$  and  $-1 < m < 0$ , where the  $W(m,n)$  indices decrease with branching.

Between the regions of  $W(m,n)$  indices increasing with branching and regions of those decreasing with branching, there are intermediate regions, where the indices either do not depend on branching, i.e. the indices  $W(1,0)$ ,  $W(0,0)$ ,  $W(0,-\infty)$ , and  $W(-1,-\infty)$ , or seem to depend on it in an irregular manner. In the latter parts of intermediate regions the  $W(m,n)$  indices indicate that some isomers seem to be more branched than the most

branched alkane of the same carbon number, whereas some other isomers seem to be less branched than the non-branched isomer, i.e. the *n*-alkane. For details cf. Fig. 16-18. The most obvious reason for that seems to be the great difference of influence of the central and/or peripheral position of branches on the  $W(m,n)$  indices. In most cases the centrally positioned branches evoke an impression that such an isomer is more branched than the one having the highest number of branches. The branches positioned on the periphery of the molecule, on the other hand, give the impression that such an isomer is less branched than the corresponding *n*-alkane.

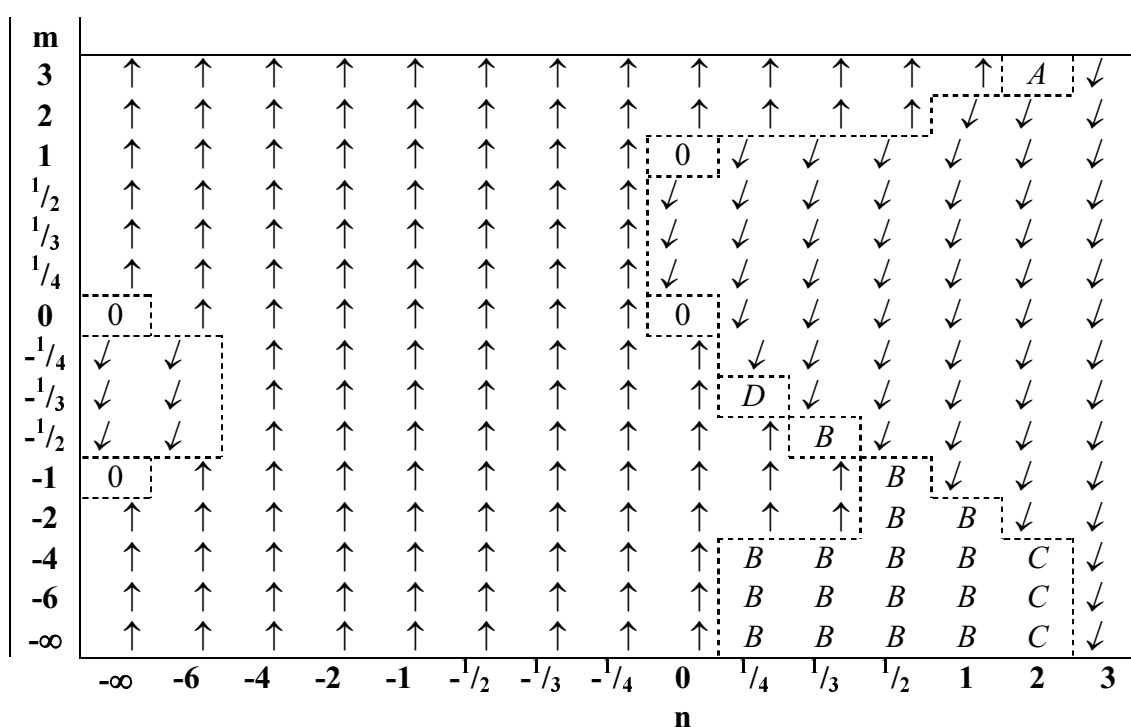


Fig. 17. The dependence of  $W(m,n)$  indices on the branching of hexanes.

↓  $W(m,n)$  decreases with branching

↑  $W(m,n)$  increases with branching

0  $W(m,n)$  presents only the size, not branching

Dotted lines - The intermediate region between the  $W(m,n)$  indices increasing with branching and those decreasing with branching. In this region some  $W(m,n)$  indices indicate that some isomers seem to be more branched than the most branched hexane, i.e. **22M4**, the others seem to be less branched than the least branched hexane, i.e. **Hx**:

A  $23M4 > 3M5 > \mathbf{22M4} > \dots$

B  $23M4 > \mathbf{22M4} > \dots$

C  $\dots > \mathbf{Hx} > 2M5$

D  $23M4 > \mathbf{22M4} > 2M5 > \mathbf{Hx} > 3M5$

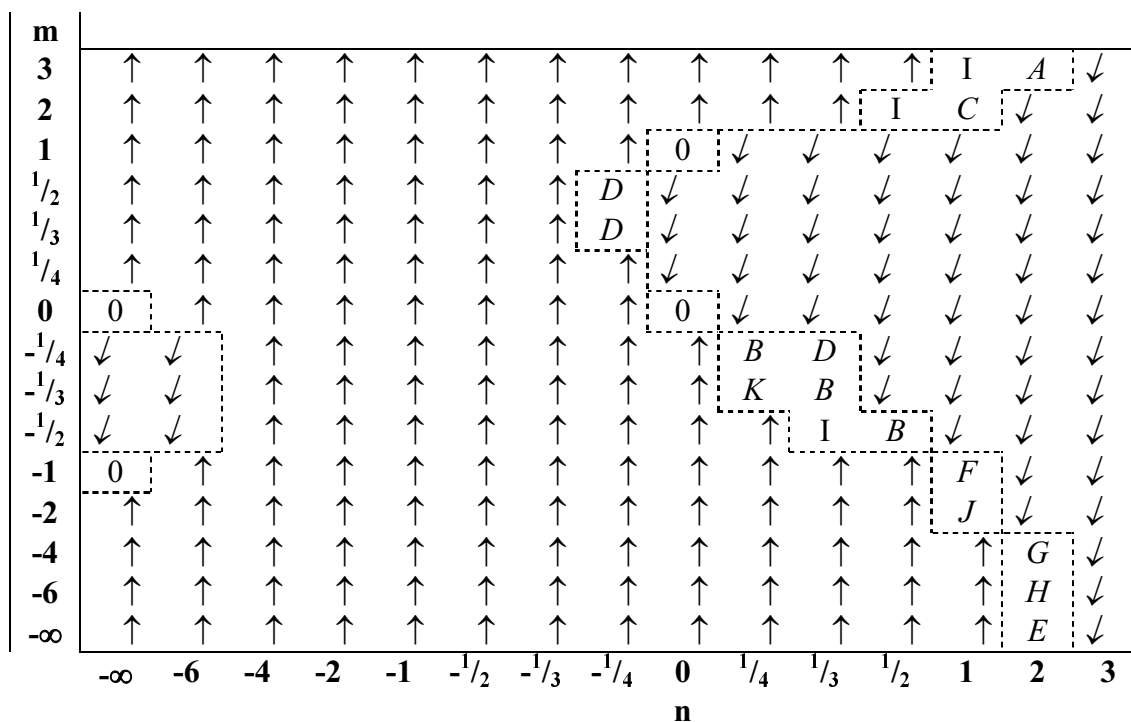


Fig. 18. The dependence of  $W(m,n)$  indices on the branching of heptanes.

↓  $W(m,n)$  decreases with branching

↑  $W(m,n)$  increases with branching

0  $W(m,n)$  presents only the size, not branching

Dotted lines - The intermediate region between  $W(m,n)$  indices increasing with branching and those decreasing with branching. In this region some  $W(m,n)$  indices indicate that some isomers seem to be more branched than the most branched heptane, i.e. **223M4**. Some other  $W(m,n)$  indices indicate that some isomers seem to be less branched than the least branched heptane, i.e. **Hp**. A number of  $W(m,n)$  indices indicate that some isomers seem to be more branched than the most branched heptane, i.e. **223M4**, whereas some other isomers seem to be less branched than the least branched heptane, i.e. **Hp**:

A  $3Et5 > 33M5 > \mathbf{223M4} > \dots$

B  $33M5 > 3Et5 > \mathbf{223M4} > \dots$

C  $33M5 > 3Et5 = \mathbf{223M4} > \dots$

D  $33M5 > \mathbf{223M4} > \dots$

E  $33M5 > 3Et5 = \mathbf{223M4} > \dots > \mathbf{Hp} > 2M6 > 24M5$

F  $33M5 > 3Et5 = \mathbf{223M4} > \dots > 2M6 = \mathbf{Hp}$

G  $33M5 > \mathbf{223M4} > \dots > \mathbf{Hp} > 24M5 > 2M6$

H  $33M5 > \mathbf{223M4} > \dots > \mathbf{Hp} > 2M6 > 24M5$

I  $\dots > \mathbf{Hp} > 3Et5$

J  $24M5 > \mathbf{223M4} > \dots$

K  $24M5 > \mathbf{223M4} > \dots > \mathbf{Hp} > 3M6 > 33M5 > 3Et5$

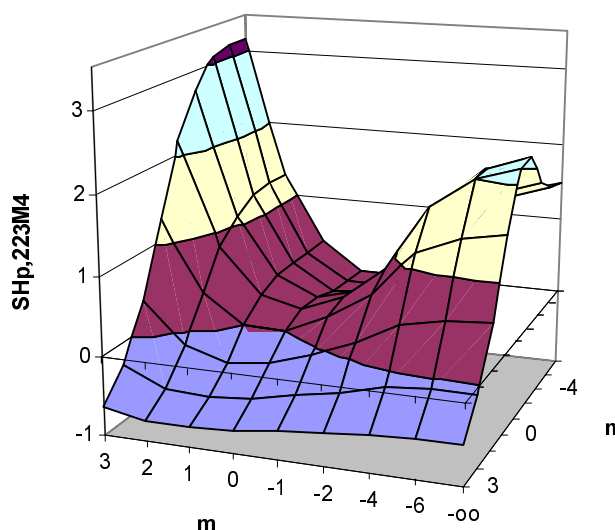
In a lesser number of cases the reverse seems to be true. An extreme case among the tested indices is  $W(-1/3, 1/4)$ , where only 24M5 seems to be more branched than



223M4, whereas 3M6, 33M5, and 3Et5 seem to be less branched than *n*-heptane, Fig. 18, case *K*. We understand these facts not as a proof that a molecule having centrally positioned branches is more branched than one with a higher number of branches, but as an indication that at some combinations of exponents the position of branches has a higher influence on the corresponding  $W(m,n)$  indices than the number of branches itself.

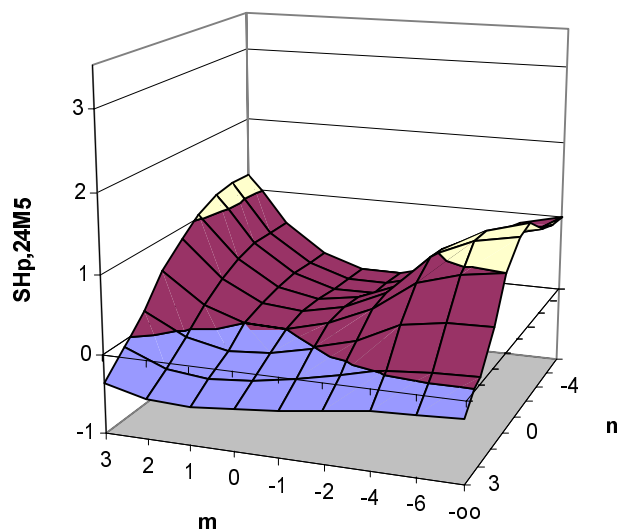
#### *The influence of structural features*

To give a visual impression of how the  $W(m,n)$  indices depend on the exponents  $m$  and  $n$  as well as on some structural features, Fig. 19-25 present their susceptibilities for branching,  $S_{n,i}$ , or their differences.

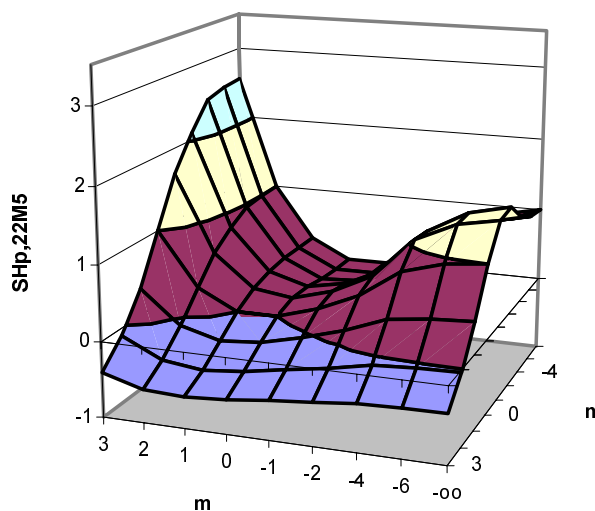


**Fig. 19.** The dependence of susceptibility for branching of the  $W(m,n)$  indices of 223M4 on exponents  $m$  and  $n$ .

Fig. 19 illustrates how  $S_{n,i}$  of  $W(m,n)$  indices depends on the exponents  $m$  and  $n$  in a saddle-like manner. At positive  $n$  and depending somewhat on  $m$ , the indices tend to decrease with branching. At negative  $n$  and depending on  $m$ , they mostly increase with branching. The increase is the lowest at  $m \approx 0$  and it is the higher the more  $m$  is different from 0. Fig. 20 shows the dependence on branching when there are two branches placed on the tertiary carbons present in the structure. Fig. 21 shows the dependence on branching when there are two branches placed on a quaternary carbon present in the structure.

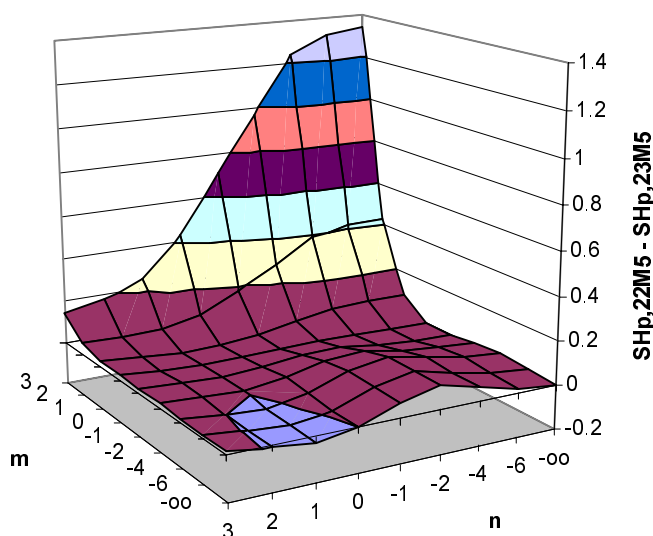


**Fig. 20.** The dependence of susceptibility for branching of the  $W(m,n)$  indices of 24M5 containing two tertiary but no quaternary carbon, on exponents  $m$  and  $n$ .



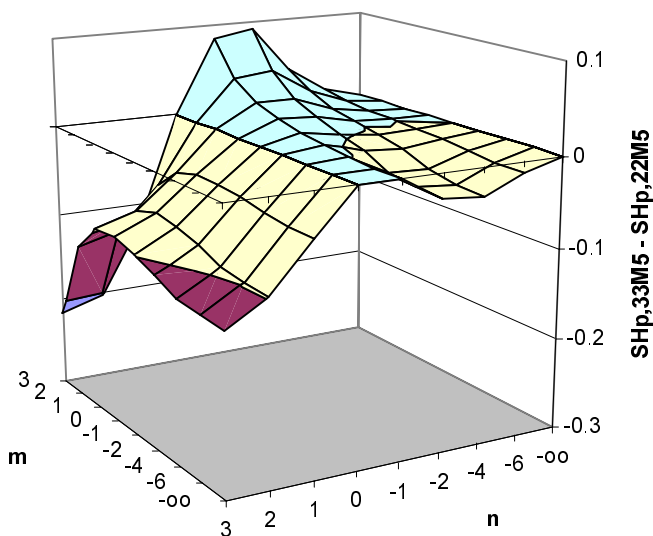
**Fig. 21.** The dependence of susceptibility for branching of the  $W(m,n)$  indices of 22M5 containing one quaternary but no tertiary carbon, on exponents  $m$  and  $n$ .

Comparing Fig. 19-21 we can see that a higher number of branches gives rise to higher susceptibility for branching, as well as that the contribution of a tertiary structure differs from that of a quaternary structure.

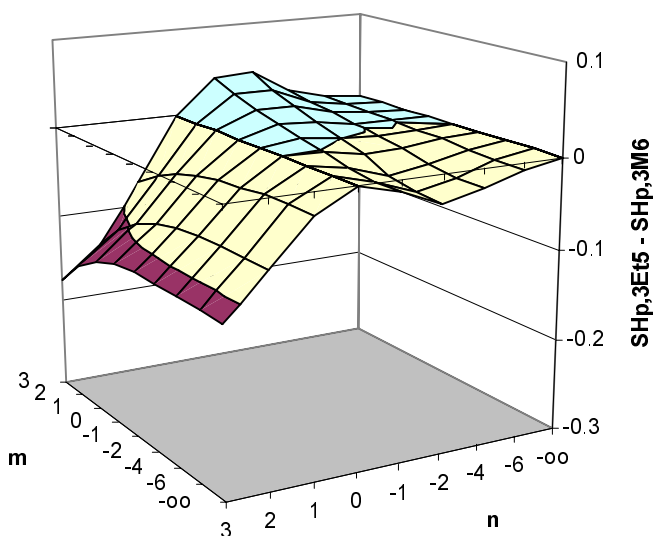


**Fig. 22.** The difference of the dependence of susceptibility for branching of the  $W(m,n)$  indices on exponents  $m$  and  $n$  where one of the two alkanes with the same carbon number and the same number of branches contains a quaternary carbon and the other alkane two tertiary ones.

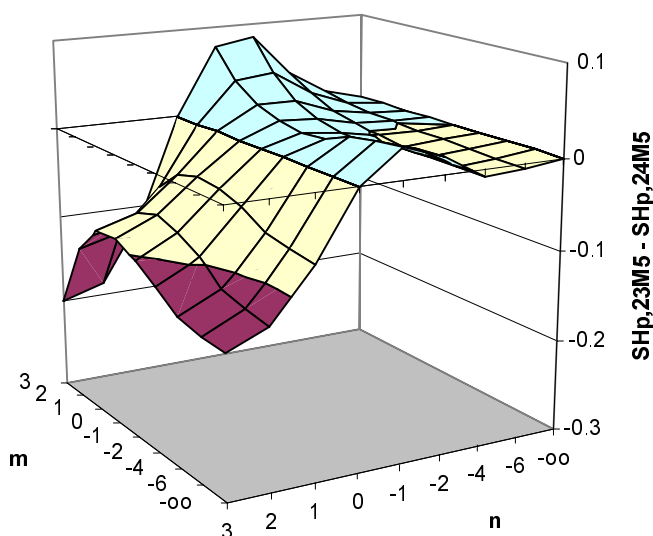
Fig. 22 shows that the difference in contribution between a quaternary and a tertiary structure is noticeable at around  $n = -2$  if  $m < 0$  and very expressed at high  $m$  and low  $n$ . In a small region roughly represented by  $-2 < n < 0$  and  $m < -2$ , a branch placed on a tertiary carbon has a slightly greater contribution than a branch placed on a quaternary carbon. Fig. 23 shows that the central position of branches contributes to the decrease of  $W(m,n)$  at  $n > 0$  more than a peripheral position. A similar effect can be observed at  $m < 0$  and  $-\infty < n < -1$ . At  $n < 0$  and high  $m$  an additional increase is observed. Fig. 24 shows a similar but not equal effect if as a branch there is an ethyl group instead of a methyl one. This is not surprising since among alkanes with carbon number 7 or 8 an ethyl group can only be placed centrally. The real effect of this structural detail may be seen with higher alkanes. In Fig. 25 the differences in influence of a centrally placed branch vs. the one being placed on the periphery of the molecule, as well as the influence of the increase in the distance between the branches are superimposed. Compared to Fig. 23, there is a clear difference around  $n = -1$  at  $m < 0$ .



**Fig. 23.** The difference of dependence of the susceptibility for branching of the  $W(m,n)$  indices on exponents  $m$  and  $n$  where one of the two alkanes with the same carbon number contains a quaternary carbon in the centre of the molecule and the other one on the periphery of the molecule.



**Fig. 24.** The difference of dependence of the susceptibility for branching of the  $W(m,n)$  indices on exponents  $m$  and  $n$  where one of the two alkanes with the same carbon number and the same number of branches contains an ethyl group and the other a methyl one.



**Fig. 25.** The difference of dependence of the susceptibility for branching of the  $W(m,n)$  indices on exponents  $m$  and  $n$  where in two alkanes with the same carbon number and the same number of branches there is a different distance between the branches.

#### *The information about branching and the complexity in the $W(m,n)$ indices*

The complexity of molecules<sup>21,22</sup> increases with their size. The smallest hydrocarbon, methane, is the least complex, whereas the higher the hydrocarbon the more complex it is, being branched or not. This agrees with the *Methane-based* definition of branching<sup>6</sup>, which seems to hold true also if the complexity is in question. On the other hand, a sufficient size of a hydrocarbon is a necessary condition for branching to be possible.

If we look at the  $W(m,n)$  indices, we can see that certain groups of tested  $W(m,n)$  indices show a clear influence of some structural features on their values. We can see a group of  $W(m,n)$  indices indicating only the size of molecules; an index indicating only the most fundamental contribution to branching, i.e. the number of branches; another index indicating the contribution of the size and of the number of branches; a group of indices indicating the contribution of the size together with the contribution of the number of branches and the type of branched structure, etc.

### The size of the molecule

The size of molecules as the only information is presented by the indices  $W(1,0)$ ,  $W(0,0)$ ,  $W(0,-\infty)$ , and  $W(-1,-\infty)$ , which do not depend on branching:

$$W(0,-\infty) = N_v - 1$$

$$W(-1,-\infty) = N_v / 2$$

$$W(0,0) = (N_v - 1) \times N_v / 2 = W(-1,-\infty) \times W(0,-\infty)$$

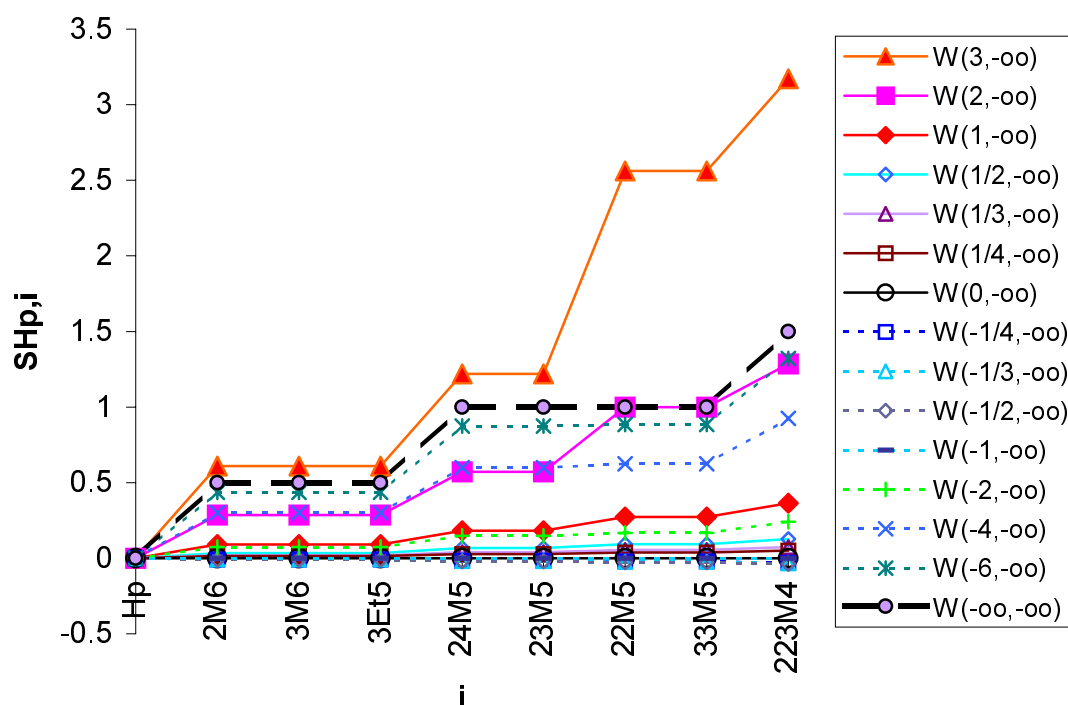
$$W(1,0) = (N_v - 1)^2 = (W(0,-\infty))^2$$

where  $N_v$  is the number of vertices in the graph of the alkane or the carbon number of the alkane. The influence of all other structural features on these indices is null. The simplest, i.e. linear measures of size are the indices  $W(0,-\infty)$ , and  $W(-1,-\infty)$ . The indices  $W(1,0)$ ,  $W(0,0)$ ,  $W(0,-\infty)$ , and  $W(-1,-\infty)$  are, on the other hand, the simplest highly degenerated indices of  $BI_M$  type. They obey the *Methane based* definition of branching, presenting the influence of the size of molecules and nothing else.

### Number of branches

Among the other  $W(m,n)$  indices it is the easiest to understand  $W(-\infty,-\infty)$ , cf. Fig. 26.

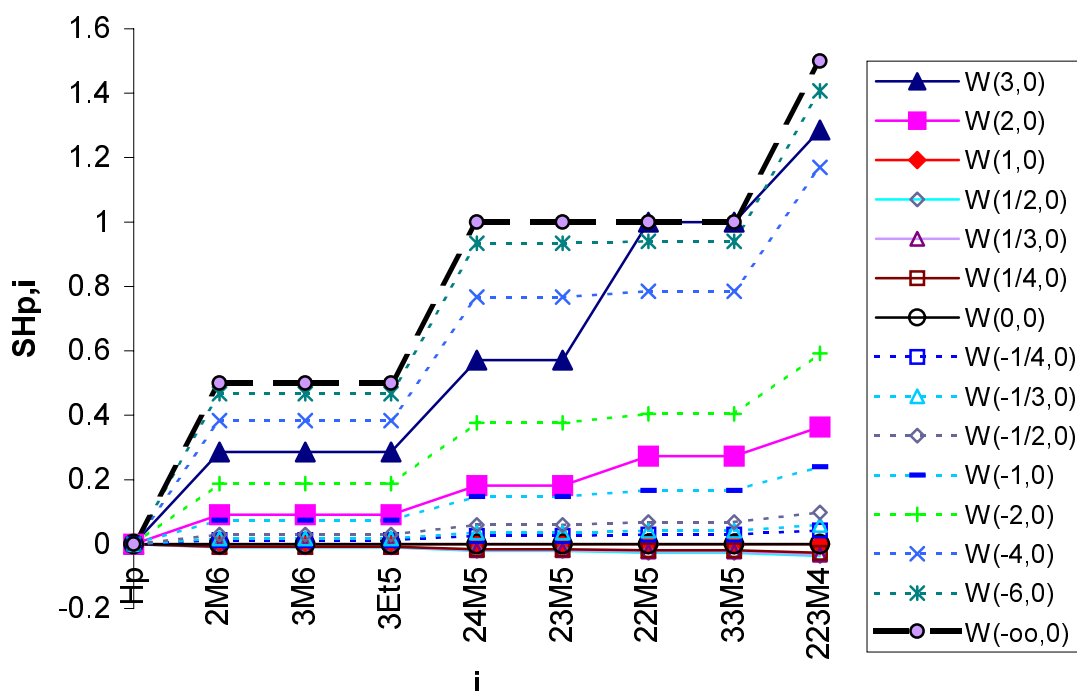
$W(-\infty,-\infty) = N_p / 2 = 1 + N_{br} / 2$ , where  $N_p$  is the number of vertices of degree one.



**Fig. 26.** The dependence of indices  $W(m,-\infty)$  of heptans on branching.

In hydrocarbons it is the number of primary carbons, and  $N_{br}$  is the number of branches. This result indicates again that  $W(m,n)$  should rather be defined as  $W(m,n) = \sum v_j^m \times d_{ij}^n$  and not as  $W(m,n) = \frac{1}{2} \sum v_j^m \times d_{ij}^n$ . A similar influence of branching as the index  $W(-\infty,-\infty)$  presents also the index  $W(-\infty,0)$ , Fig. 27, although it also depends on the size of molecule.

$W(-\infty,-\infty)$  seems to be a true branching index. At first glance it does not seem to follow the *Methane based* but the *n-Alkane based* definition of branching. Thus, it seems to belong to the  $BI_A$  type of indices and not to the  $BI_M$  type of indices. But if we recall that the information about the size of molecules among the  $W(m,-\infty)$  indices is  $2^m N_v - (2^{m+1} - 1)$ , i.e.  $2^{-\infty} N_v - (2^{-\infty} - 1)$  at the  $W(-\infty,-\infty)$  index, then we can conclude that the  $W(-\infty,-\infty)$  index belongs to the  $BI_M$  as well as to the  $BI_A$  type of indices. The index  $W(-\infty,-\infty)$  is in fact a link between the  $BI_M$  and  $BI_A$  types of indices.



**Fig. 27.** The dependence of indices  $W(m,0)$  of heptanes on branching.

$W(-\infty,-\infty)$  seems to be one of the most simple, primitive, and degenerated branching indices presenting only the most important contribution to branching or to the

complexity<sup>21,22</sup> of branched alkanes. The message of  $W(-\infty, -\infty)$  is to be understood as follows: The simplest one, the non-branched alkane is methane; other alkanes are more complex; they contain methyl groups. The more branches (methyl groups) they contain, the more complex they are.

$W(-\infty, -\infty)$  is derived from the matrix having the most reduced content of information. It seems that only the most important part of information about branching and complexity of molecules is contained in it: *The presence of vertices of degree one has the highest contribution to the complexity of molecules.* Or in other words: *The most complex are those molecules containing the vertices of degree one.* And: *Molecules containing **no** vertices of degree one are **less complex** than those containing them.* As well as: *The more vertices of degree one are contained in a structure, the more complex it is.* This seems to be information of fundamental importance. It is contained in most other indices, too, but in several instances it is obscured by other information contained in them.

In view of  $W(-\infty, -\infty)$  the size of molecule seems, presented as the only information in indices  $W(0, -\infty)$ ,  $W(-1, -\infty)$ ,  $W(0, 0)$ , and  $W(1, 0)$ , to be of lower importance for the complexity of molecules than the number of branches. The size of molecules is one of and often the most important factor influencing the physicochemical properties of alkanes. It is especially important among the properties dependent on the mobility of molecules. If we consider branching, it cannot be expressed until the size of a molecule exceeds the critical value. This may be the reason why the size seems to and indeed does contribute to these properties much more than branching. The branching and complexity can only occur when the size of molecules is large enough. As far as the index  $W(-\infty, -\infty)$  is concerned, it seems to neglect this fact or to consider it as granted.

While the index  $W(-\infty, -\infty)$  represents a link between the  $BI_M$  and  $BI_A$  types of indices, the index  $W(-\infty, 0)$  follows the *Methane based* definition of branching and clearly belongs to the  $BI_M$  type of indices. The index  $W(-\infty, 0)$  reflects a situation which is more usual to us: it presents information about the size of the molecule as well as about the number of branches:  $W(-\infty, 0) = N_v * (1 + N_{br}/2) - 1$ .



*The type of branches*

The other indices derived from the matrix having the non-diagonal elements  $v^m d^{-\infty}$ , i.e. the indices  $W(m, -\infty)$ , contain information about the size of molecule, the number of branches, as well as the type of the branched structure, i.e. whether the branch bearing carbons are tertiary or quaternary:

$$W(-\infty, -\infty) = N_p/2 = 1 + N_{br}/2 = 0 * N_v + 1 + (N_t + N_q)/2$$

$$W(0, -\infty) = (N_v - 1) + 0N_t + 0N_q$$

$$W(1, -\infty) = (2N_v - 3) + 1N_t + 1.5N_q$$

$$W(2, -\infty) = (4N_v - 7) + 6N_t + 10.5N_q$$

$$W(3, -\infty) = (8N_v - 15) + 25N_t + 52.5N_q$$

where  $N_t$  is the number of branches on vertices of degree three and  $N_q$  is the number of branches on vertices of degree four. The contribution of a branch in a quaternary structure indicates again that the division by 2 when calculating the  $W(m, n)$  indices gives rise to less simple data and should therefore not be performed. The information about the size of the molecule is in  $W(m, -\infty)$  indices equal to  $2^m N_v - (2^{m+1} - 1)$ . The information about branching is simply added to it. The contribution of a branch in a quaternary structure is higher than that of a branch in a tertiary structure and their ratio increases with  $m$ . The ratio of these contributions can be described well for  $-6 < m < 3$  with the following polynomial, Eq. 5, whereas each one of them can not be described well with a polynomial of this type up to the exponent on  $m$  of 6:

$$q/t = 0.0001m^4 + 0.0029m^3 + 0.0279m^2 + 0.1446m + 1.3245 \quad (5)$$

$$r^2 = 1$$

This polynomial is only a "good fit" one, since at  $m = -\infty$ ,  $q/t = 1$ .

Eq. 6 describes the relation between the contribution of a branch in a quaternary structure ( $q$ ) and a branch in a tertiary structure ( $t$ ). This is only a "good fit" polynomial, too.

$$q = 0.0185t^2 + 1.6408t - 0.086 \quad (6)$$

$$r^2 = 0.9999$$

The information contained in the majority of  $W(m, -\infty)$  indices, in addition to that contained in  $W(-\infty, -\infty)$ ,  $W(0, -\infty)$ , and  $W(-1, -\infty)$ , is that *a quaternary structure is more complex than two tertiary ones*. For this reason, the majority of  $W(m, -\infty)$  indices seem to be true branching indices, too. They are quite simple and degenerated branching

indices presenting besides the information on the size of the molecule and on the most important contribution to the complexity of branched alkanes, i.e. the number of branches, also the information that a quaternary structure is to be considered more branched than two tertiary ones. The same information is contained in the majority of indices  $W(m,0)$ , except in  $W(1,0)$ ,  $W(0,0)$ , and  $W(-\infty,0)$  considered above.

Among other  $W(m,0)$  indices, the ratio of contribution of a branch in a quaternary structure and of contribution of a branch in a tertiary structure, can be fitted by the following "good fit" polynomial, Eq. 7:

$$\frac{q/t}{r^2} = 1 = 0.0001\mathbf{m}^4 + 0.0023\mathbf{m}^3 + 0.0203\mathbf{m}^2 + 0.0975\mathbf{m} + 1.2045 \quad (7)$$

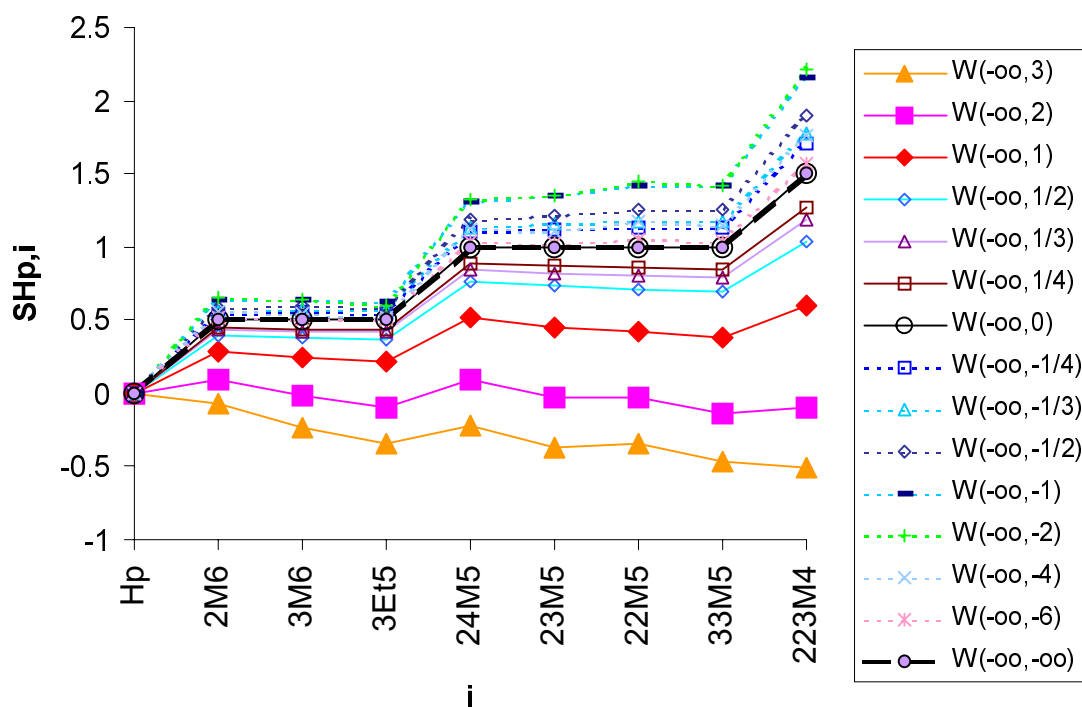
A closer inspection of indices  $W(1,0)$ ,  $W(0,0)$ ,  $W(0,1)$ , and  $W(1,1)$  indicates that the degree of vertex itself, i.e.  $v^1$ , does not contribute any information about branching to a  $W(m,n)$  index. Only when  $\mathbf{m} \neq 0$  and  $\mathbf{m} \neq 1$ , then  $v^m$  contributes the information that  $q > t$ , i.e. a quaternary structure (containing two branches) is more branched than two tertiary ones. But this holds true only if  $\mathbf{n} = 0$  or  $1$ . At other values of  $\mathbf{n}$  the contribution of  $v^1$  is slightly different than that of  $v^0$ ,  $0.99 < r^2 < 1$ . In other words, a graph containing one vertex of degree four is more branched than a graph containing two vertices of degree three. Or, at an equal number of branches in an alkane, those placed on quaternary carbons contribute more to branching than those placed on tertiary carbons. The higher the  $\mathbf{m}$ , the more expressed this contribution is.

#### *Position of branches*

Other groups of  $W(m,n)$  indices that might be easy to understand due to the reduced amount of information are the groups of  $W(-\infty,n)$  and  $W(0,n)$  indices. Comparing Fig. 28 and 29 with Fig. 26 and 27 we can see that while the index  $W(-\infty,-\infty)$  presents only the influence of the number of branches and the indices  $W(m,-\infty)$  and  $W(m,0)$  present also the influence of the type of branched structure, the indices  $W(-\infty,n)$  and  $W(0,n)$  present a different kind of information.

To see what this information is, let us look at the sequence of increasing branching of heptanes presented by the index  $W(-\infty,1)$ :  $\text{Hp} < 3\text{Et5} < 3\text{M6} < 2\text{M6} < 33\text{M5} < 22\text{M5} < 23\text{M5} < 24\text{M5} < 223\text{M4}$ . Obviously, the index  $W(-\infty,1)$  suggests that the structure

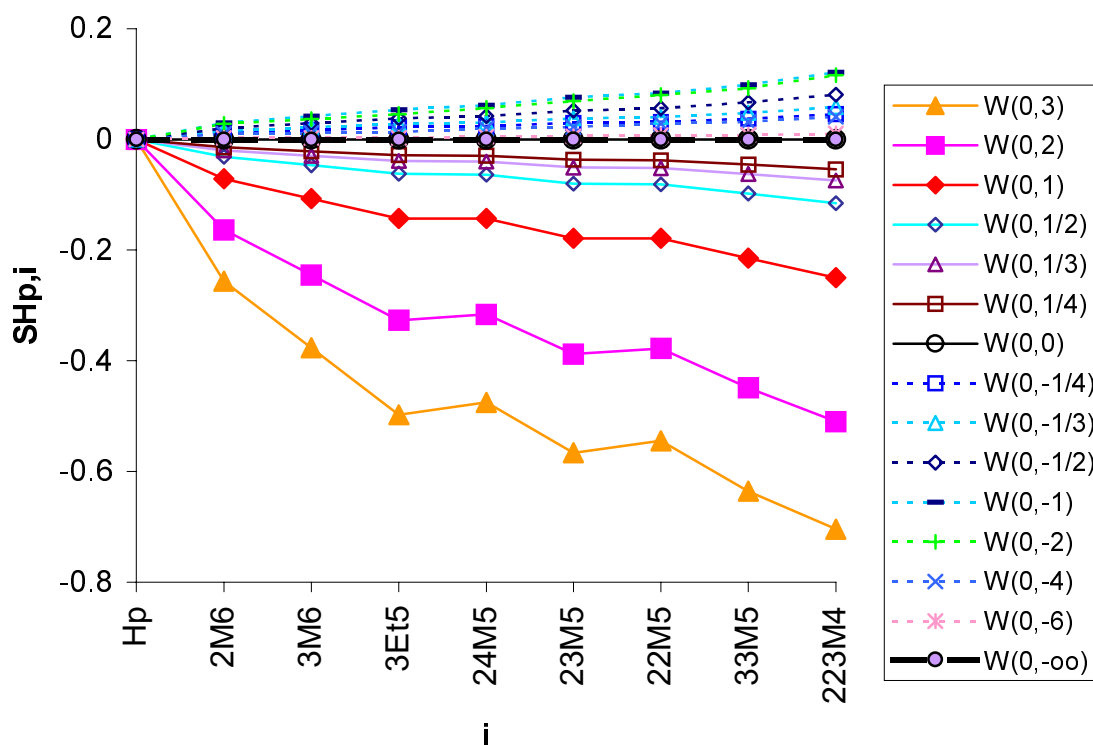
having a higher number of branches is more branched than that having a lower number of branches. Among the structures having an equal number of branches, those having more centrally positioned branches seem to be less branched than those having the branches placed more to the periphery of the molecule. The structures containing quaternary carbons seem to be presented by this index as less branched than those containing tertiary structures. The question is whether the structures containing quaternary carbons are really presented by the  $W(-\infty, n)$  and  $W(0, n)$  indices as less branched than those containing tertiary structures.



**Fig. 28.** The susceptibility for branching of the  $W(-\infty, n)$  indices of heptanes.

If we take a closer look also at the indices  $W(0, 1)$  ( $\equiv W$ ) and  $W(1, 1)$ , presenting the series of heptanes of decreasing values of the index which indicates the series of increasing branching as  $Hp > 2M6 > 3M6 > 3Et5 = 24M5 > 23M5 = 22M5 > 33M5 > 223M4$ , we can see among the heptanes having one branch that the structure having the branch closer to the centre of the molecule is presented as more branched (more complex). This is contrary to what is presented by the index  $W(-\infty, 1)$ . The same

conclusion is reached if two branches are present. In this case a quaternary structure seems at first sight to be presented as more branched than two tertiary ones. To understand the situation let us make the following comparisons. In the present cases  $W(0 \text{ or } 1,1)_{23M5} = W(0 \text{ or } 1,1)_{22M5}$  although, as presented by the indices  $W(m,-\infty)$  and  $W(m,0)$ , a quaternary structure should be felt more branched than two tertiary ones. If we compare the values of indices within the pairs  $W(0 \text{ or } 1,1)_{24M5}$  and  $W(0 \text{ or } 1,1)_{23M5}$ ,  $W(0 \text{ or } 1,1)_{23M5}$  and  $W(0 \text{ or } 1,1)_{22M5}$ ,  $W(0 \text{ or } 1,1)_{24M5}$  and  $W(0 \text{ or } 1,1)_{22M5}$ , as well as  $W(0 \text{ or } 1,1)_{22M5}$  and  $W(0 \text{ or } 1,1)_{33M5}$ , we can see that the two structures in a pair differ in the position of branches.



**Fig. 29.** The susceptibility for branching of the  $W(0,n)$  indices of heptanes.

Within the former three pairs, there is also a difference in the distance between the branches, whereas in the last pair there is the same distance between the branches, which is equal null. Within the pairs  $W(0 \text{ or } 1,1)_{24M5}$  and  $W(0 \text{ or } 1,1)_{23M5}$  as well as  $W(0 \text{ or } 1,1)_{23M5}$  and  $W(0 \text{ or } 1,1)_{33M5}$ , there is a the difference in the position of branches

and in their distance in the same direction. The values of their indices are different. Within the pairs  $W(0 \text{ or } 1,1)_{24M5}$  and  $W(0 \text{ or } 1,1)_{22M5}$  as well as  $W(0 \text{ or } 1,1)_{22M5}$ , and  $W(0 \text{ or } 1,1)_{33M5}$ , the difference is either in the position of branches or in their distance. The values of these pairs of indices are different, too. If in the pair  $W(0 \text{ or } 1,1)_{23M5}$ ,  $W(0 \text{ or } 1,1)_{22M5}$ , on the other hand, the position of branches changes in one direction, then the distance between them changes in the opposite direction. As a result, they are equal.

From these observations it follows that in the  $W(-\infty, n)$  and  $W(0, n)$  indices where  $d^n$  contributes essential information, the information about the size of molecules is presented. There is a distinction between the branches positioned in or near the centre of the molecule and those positioned on the periphery of the molecule. There is also a distinction regarding the distance between the branches. When a structure having a central position of branches is presented as more branched (more complex) than one having a peripheral position, then also a smaller distance between the branches is presented as causing the structure to seem more branched, and vice versa. Thus, the distance between branches distinguishes a quaternary structure from tertiary ones and not the difference in the type of branched structure as it was observed above considering the influence of  $v^m$ .

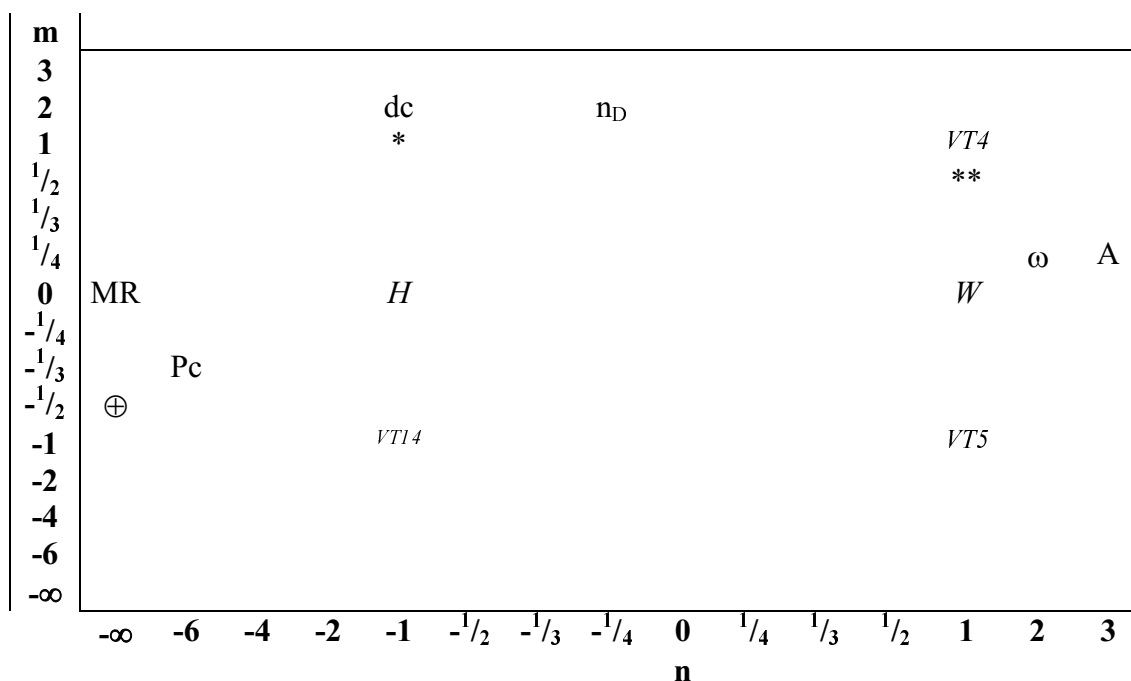
In conclusion, we can state that in  $W(m, n)$  indices derived from the matrix having the non-diagonal elements  $v^m \times d^n$ , the  $v^m$  (when  $m \neq 0$  and  $m \neq 1$ ) contribute the information about the influence of the number of branches, as well as the information that a structure having the branches on quaternary carbons (i.e. on vertices of degree four) is more branched than that having an equal number of branches on tertiary carbons (i.e. on vertices of degree three), but no information about the influence of the position of branches. The  $d^n$  (when  $n \neq -\infty$  and  $n \neq 0$ ), on the other hand, contributes information about the influence of the number of branches, of the position of branches, as well as the influence of the distance between the branches, and nothing else.

The contributions of  $v^m$  and  $d^n$  can thus be synergistic. They can also lead to an apparent over-expression of the importance of a particular structural feature. As an example we can present the index  $W(-1/3, 1/4)$ , where the position of branches seems to have the highest influence.

### Correlation of physicochemical properties with $W(m,n)$ indices

In order to see the suitability of  $W(m,n)$  indices as indices of properties, they were correlated with 21 previously tested physicochemical properties.<sup>6,7</sup>

The position in the plane of exponents  $m$  and  $n$ , where the correlation between a particular physicochemical property and a  $W(m,n)$  index is the highest among the tested ones, is presented in Fig. 30 for data from propane through octanes. For comparison, the positions of some previously known indices are included in Fig. 30.



**Fig. 30.** The positions of indices  $W$ ,  $H$ , and  $VTI_i$  (in italics), as well as of 21 physicochemical properties determined by the highest correlation coefficient (data in parentheses, see below) for data of alkanes from propane through octanes, in the plane of exponents  $m$  and  $n$ .

Single standing ones ( $r$ ):  $dc$  (0.774),  $n_D$  (0.884),  $\omega$  (0.948),  $A$  (0.571),  $MR$  (**0.998**),  $Pc$  (0.952)

\*:  $VTI_{13}$ ,  $d$  (0.915),

\*\* :  $BP/Tc$  (0.951),  $C$  (-0.940),  $\alpha_c$  (0.932),  $Zc$  (-0.662),

⊕:  $\Delta H_f^\circ g$  (**0.993**),  $Tc^2/Pc$  (**0.995**),  $Tc/Pc$  (**0.991**),  $BP$  (**0.991**),  $Tc$  (0.983),  $\Delta H_v$  (0.983),  $B$  (0.974),  $V_c$  (0.989),  $V_m$  (0.980),  $V''$  (0.845)

Several properties (i.e. 16 out of 21 tested) presented in Fig. 30 have the correlation coefficient with a  $W(m,n)$  index  $r > 0.90$ , five of them even  $r \geq 0.99$ . To a substantial extent this is due to the mutual dependence on carbon number. If the influence of the size of molecule (i.e. the carbon number) is excluded, the correlations

are worse. Eight tested physicochemical properties have  $|r| > 0.90$  and only 2 of them have  $|r| > 0.99$ . These results relate to the tested exponents. It could be reasonably expected that fine-tuning of exponents might give rise to better correlations, but this is beyond the scope of the present paper. The possible primary references<sup>7</sup> have their best correlation coefficients  $|r| > 0.90$ , i.e.  $\omega \sim Tc/Pc > Tc^2/Pc > C > \Delta Hf^\circ g$ , followed by Pc and  $\Delta Hv$ . This is an additional reason why they can be considered as suitable reference properties.

Comparing Fig. 3 with Fig. 30 we can see that the indices H and W belong to the medium degenerated ones, whereas the other indices presented in Fig. 30 are either little or not degenerated at all.

### *The $W(m,n)$ indices that might be good branching indices.*

Which topological index might be a good branching index remains an open question. There is no clear-cut criterion to evaluate it. Several attempts have been made to use one or another physicochemical property as a standard, but none of the tested physicochemical properties can be considered as the single best primary reference for branching.<sup>6,7</sup> Therefore, intuition<sup>11</sup> remains an important tool to assess the candidates for good branching indices.

Combining intuition with the results obtained by studying the dependence of physicochemical properties on different contributions to branching,<sup>6</sup> a good candidate for a branching index would depend first of all on the number of branches, followed by the type of branched structure, i.e. whether the branch bearing carbon is tertiary or quaternary, by the position of branches, etc. As a result, a regular sequence of isomers on increasing branching would be obtained. Some regular sequences are suggested in Fig. 31, where the position of the  $W(m,n)$  indices having these sequences is presented in the plane of exponents  $m$  and  $n$ ; these indices are candidates for good branching indices. Most candidates for good branching indices have  $n < 0$ .

Of the 21 physicochemical properties taken into account, among heptanes  $\omega$  has the sequence A, whereas  $Tc^2/Pc$  is close to it, having  $23M5 > 24M5$ . The sequence C is observed at  $\Delta Hv$ . Of the other physicochemical properties taken into account,  $V_m$ ,  $V''$ ,  $d$ ,  $dc$ , Pc, Vc, Tc/Pc, and MR indicate that the structures having centrally positioned

branches are more branched than those having branches on the periphery of molecules. A similar indication is presented by the indices  $W(3,3)$ ,  $W(3,2)$ ,  $W(-2 \text{ to } -\infty, 2)$ ,  $W(2,1)$ ,  $W(-1,1)$ , etc. The physicochemical properties BP, Tc,  $n_D$ , and B indicate the contrary.

<b>m</b>																
<b>3</b>	°	A	A	A	A	A	A	A	°	C	C	C				c
<b>2</b>	°	A	A	A	A	A	A	A	°	C	C	p	c			c
<b>1</b>	°	A	A	A	A	A	A	A	=	B	B	B	B			p
$1/2$	°	A	A	A	A			c	°	B	B	B	B			
$1/3$	°	A	A	A	A			c	°	B	B	B	B			
$1/4$	°	A	A	A	A			c	°	B	B	B	B			
<b>0</b>	=	A	A	A	A	A	A	A	=	B	B	B	B			
$-1/4$	°	B	A	A	A	A	A	A	°	c	c					
$-1/3$	°	B	A	A	A	A	A	A	°	<b>p</b>	c					
$-1/2$	°	B	C	A	A	A	A	A	°		p	c				
<b>-1</b>	=	C	C	A	A	A	A	A	°						c	
<b>-2</b>	°	C	C			A	A	A	°	C				c	c	
<b>-4</b>	°	C	C					A	°							c
<b>-6</b>	°	C	C						°							c
<b><math>-\infty</math></b>	•	C	C						=							c
		<b><math>-\infty</math></b>	<b>-6</b>	<b>-4</b>	<b>-2</b>	<b>-1</b>	$-1/2$	$-1/3$	$-1/4$	<b>0</b>	$1/4$	$1/3$	$1/2$	<b>1</b>	<b>2</b>	<b>3</b>
		<b>n</b>														

**Fig. 31.** The position of  $W(m,n)$  indices of heptanes considered as the best candidates for good branching indices in the plane of exponents **m** and **n**. The sequences of isomers:

A: Hp > 2M6 > 3M6 > 3Et5 > 24M5 > 23M5 > 22M5 > 33M5 > 223M4

B: 223M4 > 33M5 > 22M5 > 23M5 > 24M5 > 3Et5 > 3M6 > 2M6 > Hp

C: Hp > 3Et5 > 3M6 > 2M6 > 23M5 > 24M5 > 33M5 > 22M5 > 223M4

c: The  $W(m,n)$  index indicates that branches positioned at the centre of the molecule give rise to higher branching than those positioned at the periphery.

p: The  $W(m,n)$  index indicates that branches positioned at the periphery of the molecule give rise to higher branching than those positioned at the centre.

°: Position of branches and the distance between them has no influence.

•: The index indicates only the influence of the number of branches.

Among the tested indices,  $W(-1/3, 1/4)$  indicates the most strongly that the structures having peripherally positioned branches are to be considered the most branched, more than 223M4, whereas those having centrally positioned branches are less branched than *n*-heptane. This index obviously gives more importance to the position of branches than to their number, cf. Fig. 18, sequence *K*.



To propose which  $W(m,n)$  index might be a good branching index, we should compare the results presented in Fig. 31 with those presented in Fig. 13-16, as well as Fig. 3. This comparison suggests that the indices marked in Fig. 15 as "best" could be candidates for good branching indices. The comparison of Fig. 13-15 also suggests an extrapolation to the assumption that there might be no ideal branching index among the  $W(m,n)$  indices. The reason for this assumption is the fact that the higher the carbon number of the considered alkanes, the lesser number of tested  $W(m,n)$  indices has the sequence of isomers proposed by intuition. The higher the carbon number, the lower  $n$  is to be used to fulfil this condition. With increasing carbon number of alkanes, the useful values of exponent  $n$  decrease towards  $-\infty$ . At  $n = -\infty$ , there do exist true branching indices but they are too degenerated to be considered as "good" ones. Thus, it is probably not possible to find an ideal, generally useful branching index. But there seems to be a sufficient number of possibilities to find several  $W(m,n)$  indices that are "good enough" branching indices to be used in the range of carbon numbers in question. We should not forget that most of the compounds used in everyday life, if polymers are neglected, have the molar weight of up to a few thousand.

If, having in mind these statements, we look at the most popular topological indices we see that one of them, the Randić<sup>14</sup> index  $\chi$  falls into the region of the best candidates for good branching indices. It is not a  $W(m,n)$  index but a (let us say)  $W(a,a,c)$  index. In fact, it is a "reciprocal geometric mean of the degree of adjacent vertices" index, derived from the matrix having the non-diagonal elements  $(v_i \times v_j)^{-1/2} \times d_{ij}^{-c}$ . It could be extended by inclusion of exponents other than  $-\infty$ , e.g.  $-\infty < c < -2$ , preferably  $-8 \leq c \leq -4$ , for example,  $c = -6$ , i.e. in the region around the exponents of the Lennard-Jones potential.

### Conclusions

Several topological indices are derived from matrices, e.g. the adjacency matrix, the distance matrix, the reciprocal distance matrix, which are special cases of a more general  $v^m d^n$  matrix, which is again a special case of the general vertex-property, vertex-distance,  $f(a,b, \dots, v, \dots) \times f(d)$  matrix. Nine of the 225 Wiener-like indices,  $W(m,n)$ ,

derived from the  $v^m d^n$  matrix and tested here correlate well with a particular or several of the 21 tested physicochemical properties of alkanes, i.e.  $W(-2, -\infty)$ ,  $W(-1, -4)$ ,  $W(-1/2, -\infty)$ ,  $W(-1/4, -6)$ ,  $W(1/4, 2)$ ,  $W(1/3, 3)$ ,  $W(1/2, 1)$ ,  $W(1, -1)$ , and  $W(2, -1/4)$ .

To a substantial extent this is due to the mutual dependence on the size of the molecule. The values of  $W(m, n)$  indices increase with the size of the molecule. The increase is in general non-linear. Other structural features, e.g. the number of branches, the type of branched structure, the position of branches, and the distance between the branches, can have a substantial influence on the value of  $W(m, n)$  indices when the size of the molecule increases. Depending on the value of exponents  $m$  and  $n$ , these structural features can enhance or diminish the increase of  $W(m, n)$  indices due to the increase of the size of the molecule.

In the case of alkanes with the same carbon number, the structural features can have a crucial influence on the values of  $W(m, n)$  indices as well. The exponents  $m$  and  $n$  significantly influence the values of  $W(m, n)$  indices. The combination of exponents decides whether a  $W(m, n)$  index increases or decreases with branching, as well as to what extent. The dependence is in general saddle-like, i.e. small around  $m = 0$  and  $n = 0$ , and high at much higher and much lower values of  $m$  and/or  $n$ .

The indices  $W(1, 0)$ ,  $W(0, 0)$ ,  $W(0, -\infty)$ , and  $W(-1, -\infty)$  present only the information about the size of molecules.  $W(0, -\infty)$  and  $W(-1, -\infty)$  are linear measures of the size. The index  $W(-\infty, -\infty)$  is a simple, primitive and degenerated, but a true branching index linking the  $BI_M$  indices obeying the *Methane based* definition of branching with the  $BI_A$  indices obeying the *n-Alkane-based* definition of branching. The information contained in it seems to be: the presence and the number of methyl groups (vertices of degree one) is of the highest importance in the branching (complexity) of molecules. The index  $W(-\infty, 0)$  is a simple, primitive, and degenerated, but a true  $BI_M$  type branching index obeying the *Methane-based* definition of branching. It contains information about the size of the molecule, as well as information contained in the index  $W(-\infty, -\infty)$ . Other  $W(m, -\infty)$  indices, except  $W(0, -\infty)$  and  $W(-1, -\infty)$ , as well as the  $W(m, 0)$  indices, except  $W(1, 0)$  and  $W(0, 0)$ , suggest that a branch on a quaternary carbon (i.e. on a vertex of degree four) contributes more to branching (complexity) than a branch on a tertiary carbon (i.e. on a vertex of degree three). The indices  $W(-\infty, n)$  and  $W(0, n)$  present,

besides the information about the size of the molecule, also information about the position of branches and the distance between them. Other  $W(m,n)$  indices present different, in some cases quite extreme combinations of all these parts of information. An extreme case is  $W(-1/3, 1/4)$ , where the contribution of the position of branches is greater than that of the number of branches. Appropriate combinations of different indices seem to enable the formation of an appropriate mixture of particular parts of information presented above that fits a physicochemical property better than that contained in a single index. Among the  $W(m,n)$  indices there may be no ideal and general branching index. In spite of that, there may be several "good enough" branching indices for compounds of molar weight up to a few thousand, using  $n < -2$  and  $m > -1$ .

### References

1. M. Randić, C. Basak, M. Pompe, M. Novič, *Acta Chim. Slov.* **2001**, *48*, 169-180.
2. O. Ivanciuc, *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 1412-1422.
3. E. Estrada, *Chem. Phys. Lett.* **2001**, *336*, 248-252.
4. M. Randić, X. Guo, P. Calkins, *Acta Chim. Slov.* **2000**, *47*, 489-506.
5. S. Nikolić, N. Trinajstić, M. Randić, *Chem. Phys. Lett.* **2001**, *333*, 319-321.
6. A. Perdih, M. Perdih, *Acta Chim. Slov.* **2000**, *47*, 231-259.
7. A. Perdih, *Acta Chim. Slov.* **2000**, *47*, 293-316.
8. A. Perdih, *Acta Chim. Slov.* **2000**, *47*, 435-452.
9. D.R. Lide, *CRC Handbook of Chemistry and Physics*, 76<sup>th</sup> Ed., CRC Press, Boca Raton 1995-96.
10. J.A. Dean, *Lange's Handbook of Chemistry*. McGraw-Hill, New York 1985.
11. D. Bonchev, N. Trinajstić, *J. Chem. Phys.* **1977**, *67*, 4517-4533.
12. S.L. Taraviras, O. Ivanciuc, D.C. Cabrol-Bass, *J. Chem. Inform. Comput. Sci.* **2000**, *40*, 1128-1146.
13. O. Ivanciuc, *Rev. Roum. Chim.* **1989**, *34*, 1361-1368.
14. M. Randić, *J. Am. Chem. Soc.* **1975**, *97*, 6609-6615.
15. M.V. Diudea, I. Gutman, *Croat. Chim. Acta* **1998**, *71*, 21-51.
16. H. Wiener, *J. Am. Chem. Soc.* **1947**, *69*, 17-20.
17. H. Hosoya, *Bull. Chem. Soc. Japan* **1971**, *44*, 2332-2339.
18. D. Plavšić, S. Nikolić, N. Trinajstić, Z. Mihalić, *J. Math. Chem.* **1993**, *12*, 235-250.
19. O. Ivanciuc, T.S. Balaban, A.T. Balaban, *J. Math. Chem.* **1993**, *12*, 309-318.
20. H.P. Schultz, T.P. Schultz, *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 107-112.
21. S. Nikolić, N. Trinajstić, I.M. Tolić, *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 920-924.
22. D. Bonchev, *J. Chem. Inf. Comput. Sci.* **2000**, *40*, 934-941.

### Povzetek

Matrika z elementi  $v^m d^n$  (matrika  $v^m d^n$ ) je korak k poenotenju matrik, iz katerih izvajajo topološke indekse, saj vključuje matriko sosednosti, matriko razdalj, matriko recipročnih razdalj itd. Z uporabo matrike  $v^m d^n$  je bilo izračunanih 225 Wienerjevemu podobnih indeksov  $W(m,n)$ . Vsi so tipa  $BI_M$ , ker ustrezajo metanskemu pravilu. Indeks  $W(-\infty, -\infty)$

povezuje indekse tipa  $BI_M$  z indeksi tipa  $BI_A$ , ki ustrezajo  $n$ -alkanskemu pravilu. To je enostaven in degeneriran, vendar pravi indeks razvejanja, ki pove, da je prisotnost in število metilnih skupin (točk valence ena) najpomembnejši podatek v zvezi z razvejanostjo (kompleksnostjo) molekul. Indeks  $W(-\infty, 0)$  vsebuje tudi podatek o velikosti molekule. Večina indeksov  $W(m, -\infty)$  in  $W(m, 0)$  kaže, da veja na kvarternem ogljiku (točki valence štiri) prispeva k stopnji razvejanosti molekule (kompleksnosti) več kot veja na terciarnem ogljiku (točki valence tri). Indeksi  $W(1, 0)$ ,  $W(0, 0)$ ,  $W(0, -\infty)$  in  $W(-1, -\infty)$  dajejo le podatke o velikosti molekul. Indeksi  $W(-\infty, n)$  in  $W(0, n)$  vsebujejo podatke o velikosti molekul, o številu vej, o legi vej in o razdalji med njimi. Drugi indeksi  $W(m, n)$  predstavljajo različne, nekateri prav posebne kombinacije vseh teh posameznih podatkov.