

Polarity-Numbers of Cycle-Containing Structures

István Lukovits^{†,‡} and Wolfgang Linert^{*,‡}

Chemical Research Center, Hungarian Academy of Sciences, H-1525 Budapest, P.O. Box 17, Hungary,
and Technische Universität Vienna, Getreidemarkt 9, A-1060 Vienna, Austria

Received December 29, 1997

A heuristic approach was used to derive a new version of the polarity-number, proposed originally by Wiener for acyclic hydrocarbons. The new versions were used to demonstrate quantitative structure–property relationships in a series of acyclic and cycle-containing hydrocarbons.

INTRODUCTION

Graph invariants are numbers that *do* depend on the structure of the underlying molecule but *do not* depend on the numbering of atoms in this structure. The first application of graph invariants in chemistry was reported by Wiener¹ who explained variation in various physical properties of hydrocarbons in terms of two invariants: (1) the invariant that was later termed “Wiener index” W ; and (2) the polarity-number. The parameter W is the sum of (topological) distances d_{ij} between all pairs of atoms i and j of the given structure:

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

Wiener also proposed to consider “hydrogen suppressed graphs”; that is, the carbon skeleton is needed to calculate W or p only. Since Wiener’s famous work¹ appeared, W became one of the most important graph theoretical invariants in chemistry not only because of its ability to account for steric properties of molecules but also because of its interesting mathematical properties. Today, W is fairly often investigated in the mathematical literature,² and it is also used to solve problems in electrical engineering.³ There are numerous reviews^{4–6} on W . Today, W is usually thought to account for the “compactness” of a molecule. Later, different new graph invariants have been proposed, which are closely related to W , including the detour index, w ,^{7–9} the resistance distance, R ,¹⁰ the “Szeged” index, Sz ,¹¹ the all-path index, P ,¹² and its variant, V . All these variants of W will yield identical numbers for an acyclic structure and (in most cases) different values for a cycle-containing graph. The term “Wiener-type indices” will be used to denote all those invariants being equivalent to W in acyclic structures. In addition it should be mentioned that other variants of W have also been proposed.^{13–16} The performance of several of these indices was tested in structure–property relationship studies.^{17–19}

Most of the proposed graph theoretical invariants were tested for their ability to account for variations in physical properties (for pragmatic reasons, boiling points are most

often considered) of hydrocarbons. Hydrocarbons are ideal for testing steric factors because there are no other physical effects (namely, electron donating or withdrawing, dipole moment, etc.) that might be accounted for by the variation in the observed effects. In this work we also considered hydrocarbons. Usually correlation coefficients between experimental properties and the actual values of the graph theoretical invariants are determined in a series consisting of *acyclic* structures of various sizes. From a chemical point of view, any n -alkane and its cyclic analogue (like n -octane and cyclooctane) *are not* structural isomers because their chemical composition is different. From a graph theoretical point of view, however, because we are neglecting hydrogen atoms, octane and cyclooctane *are* isomers, and it is not clear why attempts to account for variation of properties in such a series of hydrocarbons were rarely considered.²⁰ Therefore, a more rigorous test is to consider series of molecules consisting of both *acyclic and cyclic* structures.

In our view, *homogeneous* series (i.e., consisting of structures with an identical number of carbons) should be used to investigate the performance of a graph invariant in quantitative structure–property studies. If the correlation detected between a property A and a graph invariant X (or sets of invariants) is nearly equal to the correlation demonstrated between A and the number of carbons N , then by using the principle of parsimony (or as it is often referred to, Ockham’s razor), the more simple invariant, in our example N , has to be preferred in a nonhomogeneous series. Invariant₁ is “more simple” than invariant₂ if the number of mathematical operations needed to obtain invariant₁ is less than those needed to obtain invariant₂. It is obvious that in this sense N is more simple than W or p . Therefore, a still more rigorous test series is homogeneous, that is, all structures consist of the same number of atoms (or vertexes in graph theoretical language). Wiener himself investigated homogeneous series,¹ and Balaban also emphasized the importance of this approach.²¹ In this paper we argue that the most severe test concerning the applicability of graph invariants in structure–property relationships is accomplished by considering a series that is homogeneous *and* is made up of cyclic and acyclic structures. The particular series considered in this paper consists of 18 acyclic octanes and

[†] Chemical Research Center.

[‡] Technische Universität Vienna.

11 octanes containing a single cycle.

The polarity-number is, according Wiener,¹ equal to "...the number of pairs of carbon atoms which are separated by three carbon-carbon bonds". Wiener's definition will be referred to as Definition 1, and the corresponding polarity-number will be denoted by p_1 . In contrast to W , very little work was devoted to the polarity-number, the other index proposed by Wiener, although the work by Platt should be mentioned here.^{22,23} Today it is usually assumed that the polarity-number accounts for the flexibility of an acyclic structure, because p_1 is equal to the number of bonds around which rotations can take place.

Wiener considered only acyclic hydrocarbons. Definition 1 could clearly be extended to cycle-containing systems, but the value of p_1 is zero in cyclopropane, whereas it is equal to 4 in cyclobutane. For cyclohexane, the value of p_1 is less than that in cyclopentane and cycloheptane ($p_1 = 3$ versus $p_1 = 5$ and $p_1 = 7$, respectively). Because of these inconsistency, W and p_1 may be expected to fail to account for variations in boiling points (bps) in a series composed of acyclic and of cycle-containing structures. Another possibility would be to interpret the polarity-number — and this will be referred to as Definition 2 — as the number of pairs of vertexes between which the (topological) distance is equal to three. This version of the polarity-number will be denoted by p_2 . In acyclic structures, $p_1 = p_2$, but in general cycle-containing systems $p_1 \neq p_2$. It was shown in this paper that p_2 and W were also inadequate to explain the variation of the bps in a homogeneous series composed of acyclic as well as cycle-containing structures.

In this work, we generated a set of "ideal" polarity-numbers p' , explaining (again in combination with W) variations in the bps of acyclic as well as monocyclic hydrocarbons. We tried to construct definitions that were in accordance with Wiener's definition in acyclic structures, are valid for cycle-containing structures, and reproduce p' . These efforts yielded two new kinds of polarity-numbers (namely p_3 and p_4). One of these numbers, p_4 , can be used in connection with W to predict bps of both acyclic as well as monocyclic structures of various sizes. In addition to the Wiener-type indices and polarity-number, two well-known graph theoretical invariants, the branching index proposed by Randić²⁴ and Balaban's J -index,²⁵ have also been considered.

CALCULATIONS

Table 1 lists numerical values of the graph invariants of acyclic octanes and octanes containing a single cycle and Table 2 contains the experimental bps²⁶ of these derivatives. The notation of the compounds is as follows: NX denotes the longest chain containing X atoms (vertexes) and CX denotes a cycle containing X vertexes. The expression "structures containing a single cycle" (SCSC) will denote structures that may contain besides the cycle, acyclic substructures; cyclooctane (C8) and ethyl-cyclobutane (Et-C4) are both SCSCs.

The all-path index P was obtained with the following formula:¹²

$$P = \sum \sum' |\pi_{ij}| \quad (2)$$

Table 1. Octanes: Numerical Values of Several Graph Theoretical Invariants

| molecule | W | w | S_z | P | V | R | χ | J | p_2 |
|-----------------------|-----|-----|-------|-----|-------|------|--------|-------|-------|
| N8 | 84 | 84 | 84 | 84 | 84.0 | 84.0 | 3.914 | 2.530 | 5 |
| 3Me-N7 | 76 | 76 | 76 | 76 | 76.0 | 76.0 | 3.808 | 2.862 | 6 |
| 3Et-N6 | 72 | 72 | 72 | 72 | 72.0 | 72.0 | 3.846 | 3.074 | 7 |
| 3,4Me,Me-N6 | 68 | 68 | 68 | 68 | 68.0 | 68.0 | 3.719 | 3.292 | 8 |
| 3Et-3Me-N5 | 64 | 64 | 64 | 64 | 64.0 | 64.0 | 3.682 | 3.292 | 9 |
| 4Me-N7 | 75 | 75 | 75 | 75 | 75.0 | 75.0 | 3.808 | 2.920 | 6 |
| 2Me-N7 | 79 | 79 | 79 | 79 | 79.0 | 79.0 | 3.770 | 2.716 | 5 |
| 3Et-2Me-N5 | 67 | 67 | 67 | 67 | 67.0 | 67.0 | 3.719 | 3.355 | 8 |
| 2,3Me,Me-N6 | 70 | 70 | 70 | 70 | 70.0 | 70.0 | 3.681 | 3.171 | 7 |
| 2,4Me,Me-N6 | 71 | 71 | 71 | 71 | 71.0 | 71.0 | 3.664 | 3.099 | 6 |
| 3,3Me,Me-N6 | 67 | 67 | 67 | 67 | 67.0 | 67.0 | 3.621 | 3.373 | 7 |
| 2,5Me,Me-N6 | 74 | 74 | 74 | 74 | 74.0 | 74.0 | 3.626 | 2.928 | 5 |
| 2,3,4Me,Me,Me-N5 | 65 | 65 | 65 | 65 | 65.0 | 65.0 | 3.553 | 3.464 | 8 |
| 2,3,3Me,Me,Me-N5 | 62 | 62 | 62 | 62 | 62.0 | 62.0 | 3.504 | 3.708 | 9 |
| 2,2Me,Me-N6 | 71 | 71 | 71 | 71 | 71.0 | 71.0 | 3.561 | 3.112 | 5 |
| 2,2,3Me,Me,Me-N5 | 63 | 63 | 63 | 63 | 63.0 | 63.0 | 3.481 | 3.623 | 8 |
| 2,2,4Me,Me,Me-N5 | 66 | 66 | 66 | 66 | 66.0 | 66.0 | 3.417 | 3.389 | 5 |
| 2,2,3,3Me,Me,Me,Me-N4 | 58 | 58 | 58 | 58 | 58.0 | 58.0 | 3.250 | 4.020 | 9 |
| C8 | 64 | 160 | 128 | 224 | 112.0 | 42.0 | 4.000 | 2.000 | 8 |
| 1,1Me,Me-C6 | 59 | 119 | 104 | 174 | 91.9 | 43.2 | 3.707 | 2.328 | 7 |
| 1,2Me,Me-C6 | 60 | 124 | 106 | 182 | 94.4 | 44.0 | 3.805 | 2.274 | 8 |
| 1,3Me,Me-C6 | 61 | 123 | 108 | 182 | 94.4 | 44.5 | 3.788 | 2.231 | 7 |
| 1,4Me,Me-C6 | 62 | 122 | 110 | 182 | 94.4 | 44.7 | 3.788 | 2.192 | 7 |
| Et-C6 | 64 | 124 | 109 | 184 | 97.2 | 48.2 | 3.932 | 2.125 | 7 |
| <i>n</i> -Pr-C5 | 67 | 111 | 78 | 168 | 94.1 | 56.0 | 3.932 | 2.058 | 5 |
| <i>i</i> -Pr-C5 | 62 | 106 | 73 | 159 | 89.0 | 51.0 | 3.805 | 2.242 | 6 |
| 1,1,2Me,Me,Me-C5 | 56 | 106 | 67 | 157 | 84.5 | 29.6 | 3.628 | 2.542 | 8 |
| 1,1,3Me,Me,Me-C5 | 58 | 104 | 71 | 157 | 84.5 | 30.4 | 3.601 | 2.435 | 6 |
| Me-C7 | 61 | 142 | 88 | 202 | 102.8 | 43.0 | 3.894 | 2.166 | 9 |

Table 2. Octanes: Boiling Points^a and Construction of an "Ideal" Path Index p' for Cycle-Containing Derivatives

| molecule | bp | W | p' | p_2 | p_3 | p_4 |
|-----------------------|-------|-----|------|-------|-------|-------|
| N8 | 125.8 | 84 | 5 | 5 | 5 | 5 |
| 3Me-N7 | 118.8 | 76 | 6 | 6 | 6 | 6 |
| 3Et-N6 | 118.9 | 72 | 7 | 7 | 7 | 7 |
| 3,4Me,Me-N6 | 118.7 | 68 | 8 | 8 | 8 | 8 |
| 3Et-3Me-N5 | 118.2 | 64 | 9 | 9 | 9 | 9 |
| 4Me-N7 | 117.7 | 75 | 6 | 6 | 6 | 6 |
| 2Me-N7 | 117.6 | 79 | 5 | 5 | 5 | 5 |
| 3Et-2Me-N5 | 115.6 | 67 | 8 | 8 | 8 | 8 |
| 2,3Me,Me-N6 | 115.3 | 70 | 7 | 7 | 7 | 7 |
| 2,4Me,Me-N6 | 109.4 | 71 | 6 | 6 | 6 | 6 |
| 3,3Me,Me-N6 | 112.0 | 67 | 7 | 7 | 7 | 7 |
| 2,5Me,Me-N6 | 108.4 | 74 | 5 | 5 | 5 | 5 |
| 2,3,4Me,Me,Me-N5 | 113.4 | 65 | 8 | 8 | 8 | 8 |
| 2,3,3Me,Me,Me-N5 | 114.6 | 62 | 9 | 9 | 9 | 9 |
| 2,2Me,Me-N6 | 107.0 | 71 | 5 | 5 | 5 | 6 |
| 2,2,3Me,Me,Me-N5 | 110.5 | 63 | 8 | 8 | 8 | 8 |
| 2,2,4Me,Me,Me-N5 | 99.3 | 66 | 5 | 5 | 5 | 5 |
| 2,2,3,3Me,Me,Me,Me-N4 | 106.0 | 58 | 9 | 9 | 9 | 9 |
| C8 | 148.5 | 64 | 15 | 8 | 8 | 14 |
| 1,1Me,Me-C6 | 119.5 | 59 | 11 | 7 | 10 | 11 |
| 1,2Me,Me-C6 | 126.6 | 60 | 12 | 8 | 11 | 12 |
| 1,3Me,Me-C6 | 122.3 | 61 | 11 | 7 | 10 | 11 |
| 1,4Me,Me-C6 | 121.8 | 62 | 10 | 7 | 10 | 11 |
| Et-C6 | 131.8 | 64 | 12 | 7 | 10 | 11 |
| <i>n</i> -Pr-C5 | 131.0 | 67 | 11 | 5 | 10 | 10 |
| <i>i</i> -Pr-C5 | 126.4 | 62 | 11 | 6 | 11 | 9 |
| 1,1,2Me,Me,Me-C5 | 113.5 | 56 | 11 | 8 | 13 | 11 |
| 1,1,3Me,Me,Me-C5* | 104.9 | 58 | 8 | 6 | 11 | 9 |
| Me-C7 | 134.0 | 61 | 13 | 9 | 9 | 13 |

^a According to the 76th edition of the *CRC Handbook*; the bp is equal to 115.5 °C according to ref 26.

where the first summation has to be performed for all pairs of vertexes i and j (as in eq 1) and the second summation has to be performed for all paths π_{ij} between vertexes i and j of the structure. The parameter $|\pi_{ij}|$ denotes the length of path π_{ij} . As an example, consider cyclobutane (C4): the

sum of path-lengths between any pair of vertexes is equal to 4. There are six pairs of vertexes, therefore $P = 6 \times 4 = 24$. With increasing number of cycles, the numerical value of P would increase rapidly; therefore, it seemed appropriate to define a second invariant V by using the following equation:

$$V = P/(\text{average number of paths between vertexes}) \quad (3)$$

For C4, $V = 24/2 = 12$. A computer program described earlier¹² was used to obtain P and V for all cycle-containing octanes.

The resistance distance index R , may be obtained by replacing d_{ij} in eq 1 by the resistance between sites i and j , assuming that all bonds are unit resistors. As an example, we illustrate the calculation of R for 1,1Me,Me-C6: $R = 17.5 + 2 \times 5.8333 + 2 \times 6 + 2 = 43.2$, where the first term is the resistance index²⁷ of C6, 5.8333 is the connectivity term of cyclohexane multiplied by the number of vertexes in the side chains (2), the third term is equal to the number of times one "travels" from Me to the cycle, and the last term is the resulting "resistance" between vertexes representing the methyl-groups.

RESULTS AND DISCUSSION

For $n = 18$ acyclic octanes (Tables 1 and 2) regression in terms of W and p_2 (i.e., the polarity number based on definition 2 explains 97% of the sample variance in the bps

$$\begin{aligned} \text{bp} &= 1.52W + 5.54p_2 - 30.08 \\ (0.07) \quad (0.30) \\ r &= 0.985, F = 247 \text{ and } s = 1.1 \quad (4) \end{aligned}$$

where numbers in parentheses denote the standard errors of regression coefficients, F is the result of Fisher's test, s is the standard error of the estimate, and n denotes the number of data points.

It was found that W and p_2 do not explain the variance in a series (Table 2) consisting of 29 acyclic and SCSC octanes:

$$\begin{aligned} \text{bp} &= 0.22W + 2.08p_2 + 89.28 \\ (0.26) \quad (0.66) \\ r &= 0.221, F = 0.7, s = 10.3 \quad (5) \end{aligned}$$

From eq 4 we might expect that the bp of cyclooctane (Table 2) with $W = 64$ and $p = 8$ is ~ 110 °C, whereas in reality it is 148.5 °C. Therefore, it seemed reasonable to examine in connection with p_2 other Wiener-type invariants (w , Sz , P , V , and R), the numerical values of which are equivalent with W if an acyclic structure is considered but are higher than W (except R) in cycle-containing structures. Table 3 lists the correlation coefficients obtained between various pairs of invariants and the bps in the same series. Although the correlation coefficients derived for various pairs are much higher than r in eq 5, the invariants considered were not as effective as W and p_2 in the series of acyclic octanes (eq 4).

Therefore, we assumed that eq 4 will be valid for SCSCs, if p_2 is replaced by another index p' (the ideal polarity-number) being equal to p_2 (or p_1) in acyclic structures, but differing from p_2 in cycle-containing structures. These "ideal" polarity-numbers p' were obtained by using eq 4 and expressing p' in terms of W and bp for each SCSC (the results

Table 3. Correlation Coefficients Derived between Boiling Points of 29 Compounds (Table 1.) and Combinatons of Two Graph Theoretical Indices

| index ₁ | index ₂ | r | index ₁ | index ₂ | r |
|--------------------|--------------------|-------|--------------------|--------------------|-------|
| W | w | 0.817 | Sz | V | 0.797 |
| W | Sz | 0.734 | Sz | R | 0.731 |
| W | P | 0.790 | Sz | p ₂ | 0.755 |
| W | χ | 0.895 | P | χ | 0.886 |
| W | J | 0.695 | P | J | 0.715 |
| W | V | 0.806 | P | V | 0.814 |
| W | R | 0.551 | P | R | 0.841 |
| W | p ₂ | 0.221 | P | p ₂ | 0.709 |
| w | Sz | 0.776 | χ | J | 0.852 |
| w | P | 0.822 | χ | V | 0.880 |
| w | χ | 0.894 | χ | R | 0.877 |
| w | J | 0.763 | χ | p ₂ | 0.918 |
| w | V | 0.796 | J | V | 0.835 |
| w | R | 0.857 | J | R | 0.700 |
| w | p ₂ | 0.774 | J | p ₂ | 0.801 |
| Sz | P | 0.752 | V | R | 0.828 |
| Sz | χ | 0.880 | V | p ₂ | 0.834 |
| Sz | J | 0.755 | p ₂ | R | 0.309 |

were rounded):

$$p' = (\text{bp} - 1.52W + 30.08)/5.54 \quad (6)$$

Table 2 lists the results of this procedure. Observe that for all SCSCs, $p' > p_2$.

Next we tried to find a definition by which p' could be obtained and which is in coherence with Definition 1 in acyclic structures. The variable obtained through this definition (Definition 3) will be denoted by p_3 . Let us denote by P_4 a path of length three. The hydrogen suppressed graph of n -butane is equivalent with P_4 .

Replacing formula "...the number of pairs of carbon atoms which are separated by three carbon-carbon bonds" in Definition 1 by an appropriate expression involving reference to P_4 , we propose Definition 3:

Definition 3. p_3 = number of ways P_4 can be laid upon the hydrogen suppressed graph.

The polarity number increases now monotonically; $p_3 = 3$ in cyclopropane, for which both p_1 and p_2 are equal to zero; $p_3 = 4$ in cyclobutane; $p_3 = 5$ in cyclopentane; and $p_3 = 6$ in cyclohexane, etc. Again $p_1 = p_2 = p_3$ in all acyclic structures. Table 2 lists the values of index p_3 . The parameter p_3 approximates fairly closely p' in the series of SCSCs, except cyclooctane, which instead of being equal to 15 is equal to 8. For methyl-cycloheptane $p_3 = 9$, whereas, $p' = 13$. According to the 74th edition of the *CRC Handbook*,²⁶ the bp of 1,1,3-trimethyl-cyclopentane is 115.5 °C, whereas in the 76th edition of this book, it is 104.9 °C. Therefore p' should be 10, instead of $p' = 8$ in Table 2. The former value of p' , and therefore the higher value of its bp seems to be more realistic. Table 2 lists the values of p_3 . To account for the variation in bps of simple cycles, p_3 should be multiplied by a factor equal to two, whereas for other SCSCs, the contribution of the N -cycle would be equal to N .

The estimation of the ideal polarity number was also carried out by using instead of eq 6 an analogous equation derived from eq 9 (vide infra). It can be shown that $p' = 10$ for cyclohexane, methyl-cyclohexane, and propyl-cyclohexane, $p' = 11$ for 1,1-dimethyl-cyclohexane, i-propyl-

cyclohexane, and ethyl-cyclohexane, whereas $p' = 12$ for 1,2-dimethyl-cyclohexane. These results forced us to devise a fourth definition of the polarity number:

Definition 4. p_4 = Number of ways P_4 can be laid upon the acyclic edges of a graph (including those cases, when the second edge of P_4 coincides with a cyclic edge) + 1.8x (number of ways P_4 can be laid upon the cyclic part of SCSC). All edges, which do not belong to a cycle, are acyclic edges, and the product $1.8N_c$ (where N_c denotes the number of vertexes in the cycle) has to be rounded to yield an integer.

Examples of Definition 4 for cyclopropane, cyclobutane, cyclopentane, cyclohexane, cycloheptane and cyclooctane are $p_4 = 5, 7, 9, 11, 13$, and 14 , respectively. For methyl-cyclobutane, there is a single edge, which belongs to the acyclic part of the molecule; therefore $p_4 = 7$. For 1,2-dimethyl-cyclopropane; $p_4 = 5 + 1 = 6$, because there is a single path of length three that connects the two (acyclic) methyl groups. This situation is similar to the case of acyclic molecules; that is, the polarity number of methane, ethane, and propane is zero and therefore the contribution of methyl, ethyl, (but not of propyl), isobutyl etc. is zero. For acyclic structures, $p_1 = p_2 = p_3 = p_4$, by definition.

With W and p_4 we obtained for our standard series of 29 octanes (Tables 1 and 2)

$$\text{bp} = 1.61W + 5.75p_4 - 37.22 \quad (0.13) \quad (0.33) \quad r = 0.959, F = 148, s = 3.0 \quad (8)$$

where the multiple correlation coefficient r was *definitely* higher than any of those listed in Table 3.

To extend our method to any SCSC, we considered an equation derived between bps and W and p_1 of 40 acyclic hydrocarbons ranging from methane to octanes²⁵

$$\text{bp} = 179W/N^2 + 6.9p_1 - 124 \quad (9) \quad (0.9) \quad r = 0.986, F = 658, s = 10.5 \quad (9)$$

Replacing p_1 by p_4 in eq 9 we obtain:

$$\text{bp} = 179W/N^2 + 6.9p_4 - 124 \quad (10)$$

Although eq 10 remains valid for acyclic structures, it can be used to estimate the bps of all cycle-containing structures ($n = 42$) listed in Tables 2 and 4. The estimated bps, bp_{estd} (integer values were considered, only) and the experimental bps, bp_{exp} , are depicted in Figure 1 and are related by eq 11:

$$\text{bp}_{\text{exp}} = 0.96\text{bp}_{\text{estd}} + 4.38 \quad (0.02) \quad r = 0.990, F = 2050, s = 7.5 \quad (11)$$

Note that the regression coefficients in eq 9 are related to acyclic structures, and the SCSCs were *not* included in the training set used to derive eq 10. Therefore, the high correlation coefficient (eq 11) strongly supports the usefulness of Definition 4. Equation 11 yields correct estimates of bps of SCSCs (and of course the accounts for the bps of acyclic hydrocarbons, too).

Table 4. Boiling Points ($^{\circ}\text{C}$), Wiener-Numbers (W), and Polarity-Numbers (p_4) of 31 Monocyclic Hydrocarbons

| molecule | bp | W | p_4 | N |
|-------------------|-------|-----|-------|-----|
| C3 | -32.7 | 3 | 5 | 3 |
| C4 | 12.0 | 8 | 7 | 4 |
| Me-C3 | 4.5 | 8 | 5 | 4 |
| C5 | 49.2 | 15 | 9 | 5 |
| 1,1Me,Me-C3 | 20.6 | 15 | 5 | 5 |
| 1,2Me,Me-C3 | 33.0 | 16 | 6 | 5 |
| Et-C3 | 34.5 | 17 | 5 | 5 |
| Me-C4 | 36.3 | 16 | 7 | 5 |
| C6 | 80.7 | 27 | 11 | 6 |
| Et-C4 | 70.7 | 29 | 7 | 6 |
| Me-C5 | 71.8 | 26 | 9 | 6 |
| 1,1,2Me,Me,Me-C3 | 52.6 | 26 | 7 | 6 |
| C7 | 118.5 | 42 | 13 | 7 |
| Me-C6 | 100.9 | 105 | 11 | 7 |
| 1,1Me,Me-C5 | 87.5 | 39 | 9 | 7 |
| 1,2Me,Me-C5 | 95.5 | 40 | 10 | 7 |
| Et-C5 | 103.5 | 43 | 9 | 7 |
| Bu-C5 | 156.7 | 99 | 11 | 9 |
| i-Pr-C6 | 154.5 | 88 | 11 | 9 |
| 1Me-4Et-C6 | 150.0 | 90 | 11 | 9 |
| n-Pr-C6 | 156.4 | 94 | 12 | 9 |
| 1,1,3-Me,Me,Me-C6 | 138.9 | 82 | 11 | 9 |
| 1,3,5-Me,Me,Me-C6 | 138.5 | 84 | 11 | 9 |
| 1,2,4-Me,Me,Me-C6 | 144.8 | 84 | 12 | 9 |
| 1,2Et,Et-C5 | 153.6 | 87 | 10 | 9 |
| 1Me-2Pr-C5 | 152.6 | 90 | 11 | 9 |
| 1-i-Pr-4-Me-C6 | 170.8 | 120 | 11 | 10 |
| n-Bu-C6 | 181.0 | 133 | 13 | 10 |
| i-Bu-C6 | 171.3 | 126 | 13 | 10 |
| s-Bu-C6 | 179.3 | 121 | 13 | 10 |
| t-Bu-C6 | 171.5 | 114 | 11 | 10 |

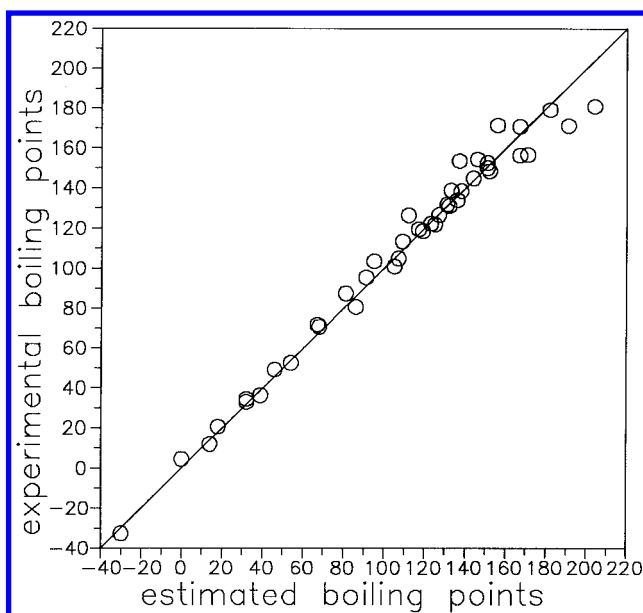


Figure 1. Experimental versus estimated boiling points ($^{\circ}\text{C}$) of cycle-containing hydrocarbons.

In concluding this section we may claim that the extension of Wiener's formula (eqs 4 and 9) to cycle-containing structures is problematic: simple definitions of the polarity number (i.e., Definitions 1, 2, and 3) yield inadequate results, whereas it is difficult to interpret a complicated definition yielding correct results. The present results are valid for acyclic structures and SCSCs only. It can be assumed that an even more complex definition will be needed when systems containing condensed cycles have to be considered. Interpretation of other possible "idealized" polarity-numbers, that might be derived by using other Wiener-type indices, remains to be done.

ACKNOWLEDGMENT

The authors are indebted to Prof. Ivan Gutman (Kragujevac) and to Dr. Christian Rücker (Freiburg) for helpful suggestions. This work was supported by Foundation "Aktion Österreich-Ungarn", project no. 29ö1 and the "Fonds zur Förderung der wissenschaftlichen Forschung in Österreich" (Projekt 11218-CHE). I.L. thanks for the kind hospitality at the Technical University of Vienna.

REFERENCES AND NOTES

- (1) Wiener, H. Correlation of heats of isomerization and differences in heats of vaporization of isomers among the paraffin hydrocarbons. *J. Am. Chem. Soc.* **1947**, 69, 2636–2638.
- (2) Doyle, J. K.; Graver, J. E. Mean distance in a graph. *Discrete Math.* **1977**, 17, 147–154.
- (3) Doyle, P. G.; Snell, J. L. *Random walks and electric networks*, Math. Association: Washington, 1984.
- (4) Gutman, I.; Yeh, Y. N.; Lee, S. L.; Luo, Y. L.. Some recent results in the theory of the Wiener number. *Indian J. Chem.* **1993**, 32A, 651–661.
- (5) Nikolic, S.; Trinajstić, N.; Mihalic, Z. The Wiener index: Development and applications. *Croat. Chem. Acta* **1995**, 68, 105–129.
- (6) Zhu, H. Y.; Klein, D. J.; Lukovits, I. Extensions of the Wiener number. *J. Chem. Inf. Comput. Sci.* **1996**, 36, 420–428.
- (7) Amic, D.; Trinajstić, N. On the detour matrix. *Croat. Chem. Acta* **1995**, 68, 53–62.
- (8) John, P. Ueber die Berechnung der Wiener Index fuer ausgewählte delta-dimensionale Gitterstrukturen. *Commun. Math. Chem.* **1995**, 32, 207–219.
- (9) Ivanciuc, O.; Balaban, A. Design of topological indices. 8. Path matrixes and derived molecular graph invariants. *Commun. Math. Chem.* **1994**, 30, 141–152.
- (10) Klein, D. J.; Randić, M. Resistance distance. *J. Math. Chem.* **1993**, 12, 81–95.
- (11) Gutman, I. A formula for the Wiener number of trees and its extension to graphs containing cycles. *Graph Theory Notes, New York* **1994**, 27, 9–15; Khadikar, P. V.; Deshpande, N. V.; Kale, P. P.; Dobrynin, A.; Gutman, I.; Dömötör, Gy. The Szeged Index and an Analogy with the Wiener Index. *J. Chem. Inf. Comput. Sci.* **1995**, 35, 547–550.
- (12) Lukovits, I. An all-path version of the Wiener index. *J. Chem. Inf. Comput. Sci.* **1998**, 38, 125–129.
- (13) Diudea, M. V. Cluj-matrix invariants. *J. Chem. Inf. Comput. Chem.* **1997**, 37, 300–305.
- (14) Tratch, S. S.; Stankevich, M. I.; Zefirov, N. S. Combinatorial models and algorithms in chemistry. The expanded Wiener number – a novel topological index. *J. Comput. Chem.* **1990**, 11, 899–908.
- (15) Randić, M. Novel molecular descriptor for structure–property studies. *Chem. Phys. Lett.* **1993**, 211, 478–483.
- (16) Mihalic, Z.; Nikolic, S.; Trinajstić, N. Comparative study of molecular descriptors derived from the distance matrix. *J. Chem. Inf. Comput. Sci.* **1992**, 32, 28–37.
- (17) Lukovits, I. The detour index. *Croat. Chem. Acta* **1995**, 69, 873–882.
- (18) Trinajstić, N.; Nikolic, S.; Lucic, B.; Amic, D.; Mihalic, Z. The detour matrix in chemistry. *J. Chem. Inf. Comput. Sci.* **1997**, 37, 631–638.
- (19) Das, A.; Dömötör, G.; Gutman, I.; Joshi, S.; Karmarkar, S.; Khaddar, D.; Khaddar, T.; Khadikar, P. V.; Popovic, L.; Sapre, N. S.; Sapre, N.; Shirhatti, A. A comparative study of the Wiener, Schultz and Szeged indices of cycloalkanes. *J. Serb. Chem. Soc.* **1997**, 62, 235–239.
- (20) Kirby, E. J. Sensitivity of topological indices to methyl group branching in octanes and azulenes, or what does a topological index index? *J. Chem. Inf. Comput. Sci.* **1994**, 34, 1030–1035.
- (21) Balaban, A. T. Using real numbers as vertex invariants for third-generation topological indices. *J. Chem. Inf. Comput. Sci.* **1992**, 32, 23–28.
- (22) Platt, J. R. Influence of neighbor bonds on additive bond properties in paraffins. *J. Chem. Phys.* **1947**, 15, 419–420.
- (23) Platt, J. R. Prediction of isomeric differences in paraffin properties. *J. Phys. Chem.* **1952**, 56, 328–336.
- (24) Randić, M. On characterization of molecular branching. *J. Am. Chem. Soc.* **1975**, 97, 6609–6615.
- (25) Balaban, A. T. Highly discriminating distance-based topological index. *Chem. Phys. Lett.* **1982**, 89, 399–404.
- (26) *CRC-Handbook of Chemistry and Physics*, 74th ed.; CRC: Boca Raton, FL, 1994.
- (27) Klein, D. J.; Lukovits, I.; Gutman, I. On the definition of the hyper-Wiener index for cycle-containing structures. *J. Chem. Inf. Comput. Sci.* **1995**, 35, 50–52.

CI970122J