CROATICA CHEMICA ACTA CCACAA **67** (4) 415–429 (1994)

ISSN 0011–1643
CCA–2195

Original Scientific Paper

Hosoya Matrix – A Source of New Molecular Descriptors*

Milan Randić

Department of Mathematics and Computer Science, Drake University, Des Moines, Iowa 50311, USA

Received April 15, 1993; accepted June 7, 1993

We introduce a novel matrix associated with molecular graphs, the construction of which is related to the construction of the Hosoya Z topological index. The new matrix provides a source of novel invariants that may be of interest in structure-property and structure-activity studies. Construction of the novel matrices is outlined and a few novel invariants extracted. A comparison with closely related topological indices is illustrated on regressions with the boiling points in octanes.

INTRODUCTION

Linear regression analysis remains one of the central tools for data reduction, interpolation, extrapolation and prediction of data despite of the fact that such analysis does not reveal causal relationships between the descriptors and the properties. An important recent development, a procedure for constructing orthogonal molecular descriptors,[1] did not only significantly upgrade this old and veritable methodology but made it the method of choice. The availability of orthogonal (*i.e.*, linearly independent and unrelated) descriptors did not only allow unambiguous interpretation of the results of regression analysis and enable individual evaluation of the descriptors.[2] This is a procedure by which structural information from two and more descriptors are concentrated into a single descriptor. In this way, one can reduce the number of independent variables in multiple regression. With such novel capabilities in regression analysis, attention has again been shifted to the design and construction of novel structural descriptors. The goal is to find descriptors that can produce better regressions with fewer variables, but variable that have a direct structural interpretation. Here, we will outline such a novel route to molecular descriptors that may have a potential in structure-property and structure-activity studies.

---

* Dedicated to the memory of Professor Tibor Škerlak.

CONSTRUCTION OF NOVEL DESCRIPTORS

There are many descriptors (topological indices) outlined in the literature[3] many of which have been available through appropriate software.[4] We can classify these as follows:

Ad hoc Descriptors

These descriptors were initially designed with a particular application in mind but they have later found many additional applications. They include such well known descriptors as the Wiener number W,[5] the Hosoya Z number,[6] and the connectivity index χ,[7] probably the three best known and most commonly used descriptors.

Modifications

Many descriptors allow for some modifications, which will produce additional descriptors. For example, from the connectivity indices, one can construct valence connectivity indices[8] or kappa shape-indices.[9] Many such variations of the existing descriptors are described in the literature, including the valence connectivity indices[10] and the recently introduced optimal weighted paths for molecules containing heteroatoms.[11] Because of their generality, we will mention only two of such modifications. First, it is possible to modify the existing algebraic form used in the definition of an index. For example, consider the algorithm used in the construction of the connectivity index:[7] $(m, n)^p$, where p is –1/2. We can change p to –1/3 (and other values) in searching for alternative descriptors. The case $p = -1/3$ leads to a better linear correlation for the boiling points in alkali (use of $p = -1/2$ shows a slight quadratic departure of the correlation line).[12] The case $p = -1$ reduces the index to that considered by Alterburg,[13] while the case $p = -1$ generates the so called Zagreb index.[14] Second, one can modify »geometrical« considerations such as deleting bonds (or larger fragments) and derive the same descriptors for the fragments so generated, which are subsequently combined into a novel index. In this way, indices P'/P, Z'/Z, W'/W and χ'/χ are derived from indices P, Z, W and χ, respectively.[15,16]

Structurally Related

The path numbers P_i suggested by Platt already in the late 1940's[17] illustrate well that here we deal with a family of structurally closely related descriptors. Additional illustrations include the »higher« connectivity indices,[18] the weighted path numbers[19] and, generally, indices that can be constructed by some iterative procedure or are extracted from matrices by counting paths or using the matrix powers.

Computationally Related

Matrix invariants such as eigenvalues (spectrum) coefficients of the characteristic polynomial, determinant, and other matrix invariants offer a route to several standard molecular descriptors. However, there are several (nonstandard) matrix computational procedures that yield additional graph invariants. Multiplication of a matrix by vectors generates new vectors, which can offer novel invariants, as outlined by Balaban and coworkers.[20] While such well defined computational schemes offer novel possibilities, the derived descriptors may not offer a simple and direct structural interpretation.

Infinite Matrices

Trivially, one can construct higher powers of any matrix but, due to the Cayley-Hamilton theorem,[21] such higher matrices beyond the power n (n being the number of vertices in a graph) do not introduce novel structural information. However, we can modify the procedure and construct invariant from such power matrices that are not subject to the Cayley-Hamilton theorem. One of the simplest of such invariants can be obtained as a bond-additive quantity from a matrix by adding all matrix elements corresponding to adjacent pairs of vertices.[22]

Novel Matrices

By associating a matrix with a graph, one can use selected matrix invariants as novel molecular descriptors. This powerful approach has apparently been underutilized. Until quite recently, except for the adjacency matrix $\boldsymbol{A}$, and the distance matrix $\boldsymbol{D}$, there were hardly any other matrices of potential interest in chemistry available. An illustration of such a novel matrix is the χ matrix constructed by Kier and Hall.[23] Matrix elements are simply contributions of each pair of vertices to the corresponding connectivity index, *i.e.*, the contributions are given by the product of weighted vertices (using $1/\sqrt{\mathrm{d}}$, d being the degree of a vertex) involved in the path from i to j. For adjacent vertices, these are the contributions arising in construction of the connectivity index. Although the χ matrix is built from the contributions entering the construction of the connectivity index and higher connectivity indices and offers no novel structural elements, it offers additional matrix invariants as a matrix. One can, for example, consider the spectrum of such matrices and view the index eigenvalue (the largest positive eigenvalue) as a novel index. A novel structural matrix is that of Tratch and collaborators[24] who introduced matrix elements that count paths with an appropriate weight based on the distance between the vertices involved. This author recently initiated an intensive search for additional graph matrices. As a result of these recent efforts, we now have several additional graph matrices. The most recent additions to this growing collection of graph matrices include: the matrix in which uniform electric resistance associated with each edge governs the distance function for the vertices;[25] the matrix (named Wiener matrix because of some relationship to the Wiener number and the procedure used by Wiener to calculate $\boldsymbol{W}$) that enumerates all paths that include a particular pair of vertices;[26] the matrix that enumerates restricted (qualified) random walks over a graph.[27] The elements in this matrix, which incidentally is nonsymmetric, are determined by the probability of a random walk of length D_{ij} between vertices i and j. Finally, the novel matrix introduced here, which we label as $\boldsymbol{Z}$ matrix as it bears a relationship to Hosoya's Z topological index, is the latest addition to this growing class of matrices to serve as invariant generators.

3-Dimensional Matrices

Besides using simple geometrical distances to construct a matrix associated with a graph (structure),[28] we may mention idealized matrices associated with a graph embedded on a regular 2-dimensional or 3-dimensional coordinate grid (such as graphite or diamond lattice).[29] Invariants derived from such topographic matrices differentiate various conformational isomers, such as *cis-trans*, boat-chair, gauche-trans cases. However, one of the drawbacks of geometrical matrices is that they con-

tain no information on bonding. In overcrowded molecules with several close contacts, it may be difficult to distinguish between a caged structure and a structure with fewer direct bonds. This will particularly be the case of idealized structures superimposed on regular grids. However, one can combine graph distances and geometrical distances, and arrive at matrices that discrimate open and closed rings (involving vertices with the same set of coordinates).[30]

HOSOYA *Z* MATRIX

Consider a general problem: How to associate a matrix ***M***(G) with a graph G? To arrive at such a matrix, we have to assign to each pair (*ij*) some number in a unique (well defined mathematical) way. A somewhat simpler task is assigning a number to each pair of adjacent vertices, which would produce a sparse matrix that can be viewed as a adjacency matrix. Many topological indices, *e.g.*, the connectivity index, are bond additive and, hence, permit constructing matrix elements associated with such bond additive contributions for a novel matrix. If we can extend such an algorithm for constructing adjacent elements in a matrix to those nonadjacent, we may arrive at a novel matrix of potential interest in structure-property studies. This has been our strategy in constructing novel matrices, such as the Wiener matrix, the restricted random walk matrix, and in this way we derived the ***Z*** matrix.

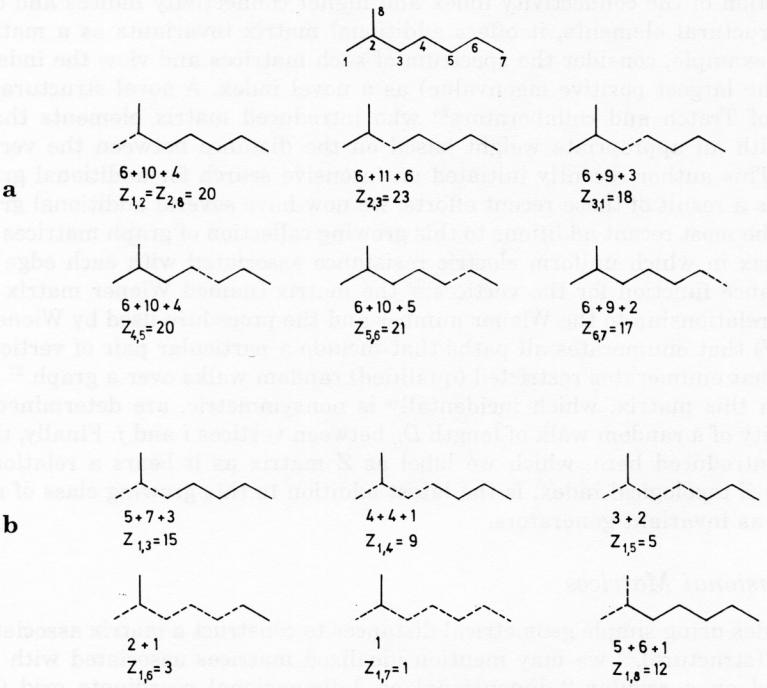

Figure 1. Calculation of the matrix elements in the Hosoya matrix for 2-methylheptane. Part (a) the elements corresponding to bonds (adjacent vertices); Part (b) a few of the elements corresponding to nonadjacent vertices.

Let us first be reminded of Hosoya's Z topological index and its construction. Z index counts nonadjacent edges in a graph, first one edge at a time, then two edges, *etc.* A Kekulé valence structure of a benzenoid hydrocarbon represents the last member in such a count, the maximal number of disjoint edges (the so called perfect matching). Observe that by definition Z is a global index, *i.e.*, an index (like the Wiener index) that cannot be partitioned into local contributions, such as various bond-additive parts. In contrast, the modified index Z'/Z is based on a bond additive scheme. The index Z'/Z is constructed by considering one bond (a pair of adjacent vertices!) at a time and after this is erased considering the Z topological index for the remaining fragment(s). The top part of Figure 1 illustrates the construction of the Z' (to be later divided by Z) for 2-methylheptane. Under each diagram, we give the number of edges, pairs of nonadjacent edges, and the number of three nonadjacent edges. Such partial contributions can be assigned to each pair of adjacent vertices. Thus, for bonds (1,2), (2,3), (3,4), *etc.*, we obtain respectively, 20, 23, 18, *etc.*. These same contributions can be now viewed as elements of the matrix $\mathbf{Z}$: $Z_{1,2} = 20$, $Z_{2,3} = 23$, $Z_{3,4} = 18$, *etc.* When such contributions arising from adjacent vertices (*i.e.*, bonds) are added, we obtain $Z' = 139$. However, by not adding bond contributions and treating them as matrix elements, we arrive at corresponding sparse matrix elements shown in Table I.

TABLE I

Sparse and complete $\mathbf{Z}$ matrix for 2-methylheptane

2-methylheptane															
Sparse								Complete							
0	20	0	0	0	0	0	0	0	20	15	0	5	3	1	12
20	0	23	0	0	0	0	20	20	0	23	14	8	5	2	20
0	23	0	18	0	0	0	0	15	23	0	18	11	7	3	5
0	0	18	0	20	0	0	0	9	14	18	0	20	13	6	9
0	0	0	20	0	21	0	0	5	8	11	20	0	21	10	5
0	0	0	0	21	0	17	0	3	5	7	13	21	0	17	3
0	0	0	0	0	17	0	0	1	2	3	6	10	17	0	1
0	20	0	0	0	0	0	0	12	20	5	9	5	3	1	0

This constructional procedure can be extended to include also nonadjacent vertices by considering paths (ij), where ij are nonadjacent. By erasing such a path and by enumerating the corresponding topological index for the remaining fragment(s) (*i.e.*, the number of nonadjacent bonds) we can assign a number to any pair of labels (ij). The construction of several elements of the $\mathbf{Z}$ matrix of 2-methylheptane corresponding to nonadjacent vertices is illustrated in the lower part of Figure 1. Here, there is a minor departure from the standard definition of the topological index Z in that we only count disjoint bonds and have not added 1 to the count (but this distinction is not essential and the resulting numbers could be augmented if so desired).

ILLUSTRATIONS

In Table II we list: $\mathbf{Z}$ matrices for normal alkanes having up to ten carbon atoms. The examples show some general features of such matrices. The largest entries ap-

TABLE II

Hosoya matrices for normal alkanes from propane to decane

n-propane

0	1	0
1	0	1
0	1	0

n-butane

0	2	1	0
2	0	3	1
1	3	0	2
0	1	2	0

n-pentane

0	4	2	1	0
4	0	5	3	1
2	5	0	5	2
1	3	5	0	4
0	1	2	4	0

n-hexane

0	7	4	2	1	0
7	0	9	5	3	1
4	9	0	8	5	2
2	5	8	0	9	4
1	3	5	9	0	7
0	1	2	4	7	0

n-heptane

0	12	7	4	2	1	0
12	0	15	9	5	3	1
7	15	0	14	9	5	2
4	9	14	0	14	9	4
2	5	8	14	0	15	7
1	3	5	9	15	0	12
0	1	2	4	7	12	0

n-octane

0	20	12	7	4	2	1	0
20	0	25	15	9	5	3	1
12	25	0	23	14	8	5	2
7	17	23	0	24	14	9	4
4	9	14	24	0	23	15	7
2	5	8	14	23	0	25	12
1	3	5	9	15	25	0	20
0	1	2	4	7	12	20	0

n-nonane

0	33	20	12	7	4	2	1	0
33	0	41	28	15	9	5	3	1
20	41	0	38	23	14	8	5	2
12	28	38	0	39	24	14	9	4
7	15	23	39	0	41	23	15	7
4	9	14	24	41	0	38	28	12
2	5	8	14	23	38	0	41	20
1	3	5	9	15	28	41	0	33
0	1	2	4	7	12	20	33	0

n-decane

0	54	33	20	12	7	4	2	1	0
54	0	67	40	28	15	9	5	3	1
33	67	0	59	41	23	14	8	5	2
20	40	59	0	64	39	24	14	9	4
12	28	41	64	0	71	39	23	15	7
7	15	23	39	71	0	64	41	28	12
4	9	14	24	39	64	0	59	40	20
2	5	8	14	23	41	59	0	67	33
1	3	5	9	15	28	40	67	0	54
0	1	2	4	7	12	20	33	54	0

pear for the adjacent vertices and, generally , the magnitudes of the entries fall off as the distance between the vertices increases. In this respect, the ***Z*** matrix shows some similarity to the Wiener matrix, both of which contrast with the distance matrix ***D*** where the opposite is true, and as the path between vertices increases, the corresponding entries in the distance matrix also increase. This increase of matrix elements with separation in the ***D*** matrix is probably the main reason why such matrices have played a less important role in structure-property studies. It is to be expected that, as the separation between vertices increases, their »interaction« will be less and less important. ***W*** matrix and ***Z*** matrix simulate this behavior and satisfy this general condition and may, therefore, be more important for describing structure-property relationships.

CONSTRUCTION OF *Z* MATRICES

Enumeration of graph invariants, even for relatively small graphs, may be time consuming and error-prone. Hence, it is desirable to delegate such construction to computers, providing a practical algorithm can be designed. To arrive at some useful computational approach to such matrices, one should examine the properties of such matrices in more detail. Regularities and patterns observed will serve as a check on numerical work and may lead to recursive relations or an algorithmic solution. So far, however, too few regularities have been observed to arrive at either recursions

or an algorithm. In the case of n-alkanes, we observe that for the terminal vertices the difference between successive elements (in the first and the last rows or the first and the last columns) gives the Fibonacci sequence: 1, 1, 2, 3, 5, 8, 13,... One can also observe that for n-alkanes, except for the »leading« entry in each row (corresponding to adjacent vertices), the »tail« part is the same as in a molecule having one carbon atom less.

This is, however, too little (and too special a case) to start designing $\boldsymbol{Z}$ matrices for molecules of interest. Hence, in order to construct such matrices for all the isomers of octanes (shown in the Appendix), we had to resort to the »brute force«. How-

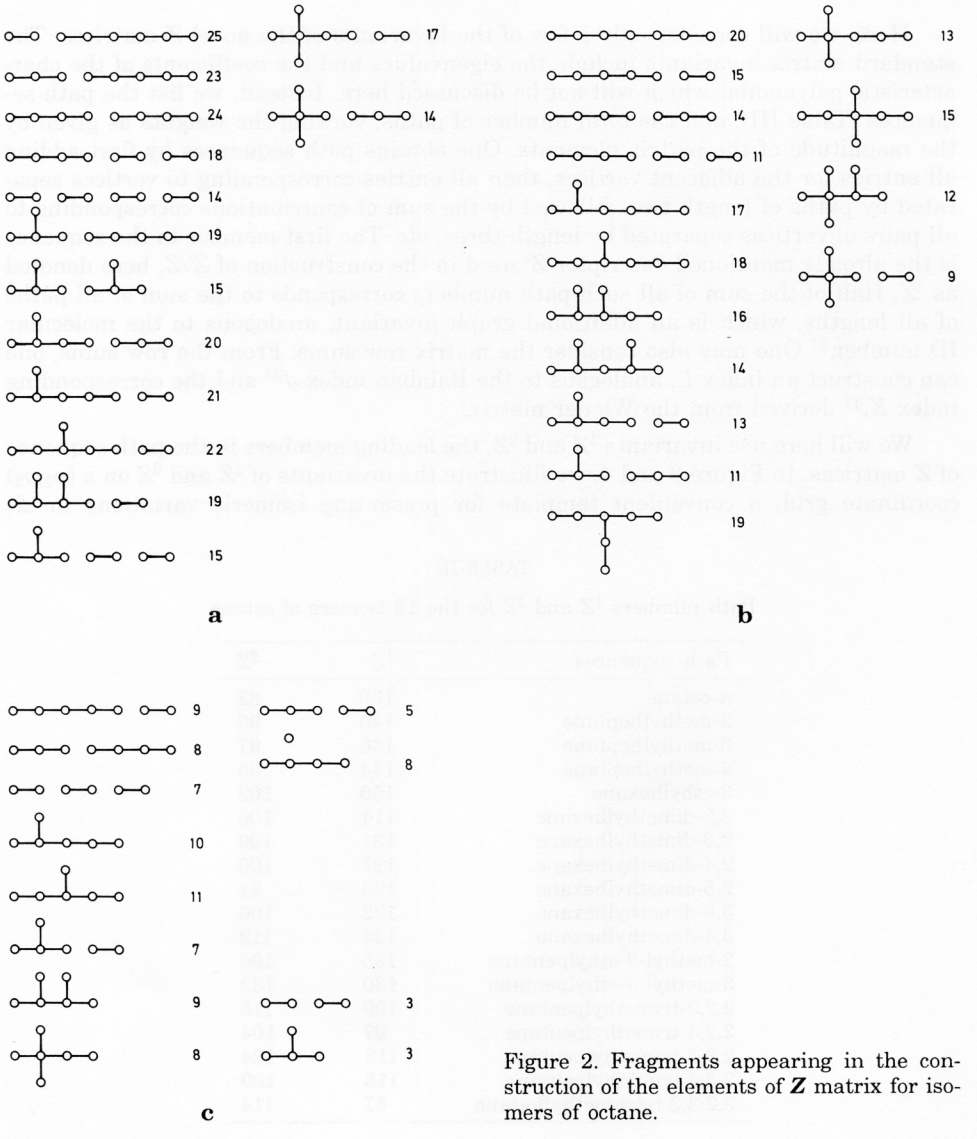

Figure 2. Fragments appearing in the construction of the elements of $\boldsymbol{Z}$ matrix for isomers of octane.

ever, instead of constructing some 500 subgraphs to build all eighteen 8×8 (symmetrical) matrices, it suffices to consider some 50 subgraphs (Figure 2), since many of the subgraphs arising in the construction appear again and again in different isomers. Once the count of nonadjacent edges in these 50 subgraphs (many of which are disjoint) is made, the information is combined in different ways in different isomers. In the Appendix, we list the 18 $\boldsymbol{Z}$ matrices for all isomers of octane, as they may serve as a source for constructing additional matrix invariants (not considered in this report).

NOVEL INVARIANTS

Here, we will consider only a few of the invariants of the novel $\boldsymbol{Z}$ matrices. The standard matrix invariants include the eigenvalues and the coefficients of the characteristic polynomial which will not be discussed here. Instead, we list the path sequences (Table III), and the total number of paths, viewing the weights as given by the magnitude of the matrix elements. One obtains path sequences by first adding all entries for the adjacent vertices, then all entries corresponding to vertices separated by paths of length two, followed by the sum of contributions corresponding to all pairs of vertices separated by length three, etc. The first member in the sequence is the already mentioned descriptor Z' used in the construction of Z'/Z, here denoted as ${}^1Z$. Half of the sum of all such path numbers corresponds to the sum of all paths of all lengths, which is an additional graph invariant, analogous to the molecular ID number.[31] One may also consider the matrix row sums. From the row sums, one can construct an index L, analogous to the Balaban index J[32] and the corresponding index K,[33] derived from the Wiener matrix.

We will here use invariants ${}^1Z$ and ${}^2Z$, the leading members in the path sequence of $\boldsymbol{Z}$ matrices. In Figure 3 and 4, we illustrate the invariants of ${}^1Z$ and ${}^2Z$ on a (p_2,p_3) coordinate grid, a convenient template for presenting isomeric variations in al-

TABLE III

Path numbers ${}^1Z$ and ${}^2Z$ for the 18 isomers of octane

Path sequences	${}^1Z$	${}^2Z$
n-octane	160	82
2-methylheptane	140	90
3-methylheptane	146	97
4-methylheptane	144	95
3-ethylhexane	150	102
2,2-dimethylhexane	114	106
2,3-dimethylhexane	131	109
2,4-dimethylhexane	127	100
2,5-dimethylhexane	123	94
3,3-dimethylhexane	122	106
3,4-dimethylhexane	134	112
2-methyl-3-ethylpentane	135	106
3-methyl-3-ethylpentane	130	132
2,2,3-trimethylpentane	109	118
2,2,4-trimethylpentane	97	104
2,3,3-trimethylpentane	113	124
2,3,4-trimethylpentane	118	109
2,2,3,3-tetramethylbutane	87	114

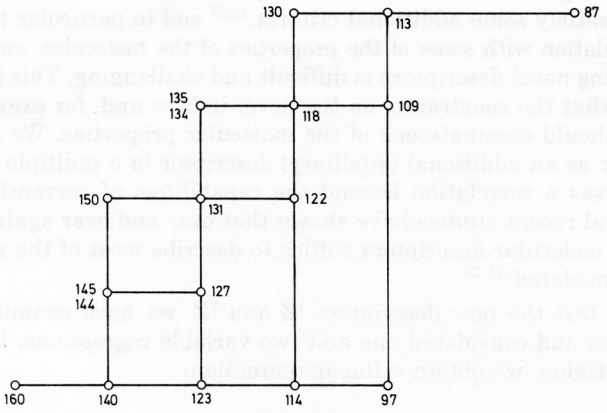

Figure 3. Regular behavior of ^1Z on the (p_2,p_3) coordinate grid.

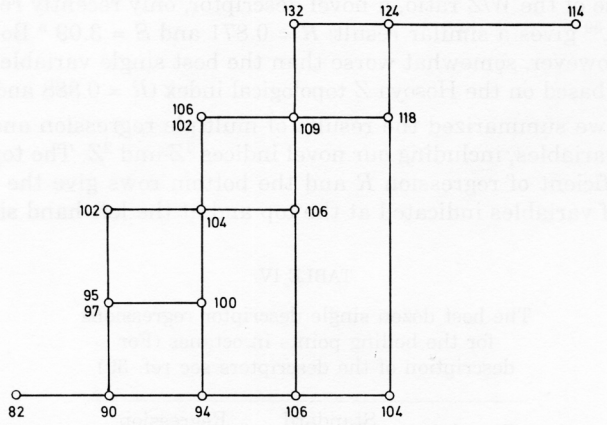

Figure 4. Regular behavior of ^2Z on the (p_2,p_3) coordinate grid.

kanes.[34,35] As we can see from Figure 2 and 4, both novel descriptors show quite regular variations in the (p_2,p_3) coordinate system, as we move along either of the two axes. This immediately suggests that the novel invariants ^1Z and ^2Z may produce good regressions for selected physicochemical properties of alkanes, as the latter also show regular variations along the ^1Z and ^2Z axes of the (p_2,p_3) coordinate system, also referred to as the periodic table of isomers.[35]

APPLICATIONS

To suggest a novel invariant, even one with a simple structural interpretation, is not necessarily a difficult task, which is part of the problem of noncontrolled pro-

liferation of topological indices. However, if one makes the restriction that the novel indices should satisfy some additional criteria,[15,36] and in particular that they should show good correlation with some of the properties of the molecules, one finds that the task of introducing novel descriptors is difficult and challenging. This is so even when we relax somewhat the constraints on the novel indices and, for example, do not insist that they should dominate one of the molecular properties. We may consider a novel descriptor as an additional (auxiliary) descriptor in a multiple regression, but one that improves a correlation beyond the capabilities of currently available descriptors! Several recent studies have shown that over and over again the same limited number of molecular descriptors suffice to describe most of the molecular properties so far considered.[37-39]

In order to test the new descriptors ^1Z and ^2Z. we have examined the boiling points in octanes and considered one and two variable regressions. Using ^1Z as the independent variable, we obtain a linear regression:

$$BP = 0.2866\ ^1Z + 77.416$$

with the coefficient of regression $R = 0.871$ and the standard error $S = 2.10$ °C. This is comparable with a dozen best results, as it can be seen from Table IV. Regression based on the use of the W/Z ratio, a novel descriptor, only recently recognized as potentially useful,[39] gives a similar result: $R = 0.871$ and $S = 3.09$ ° Both of these correlations are, however, somewhat worse than the best single variable regression (for octane isomers) based on the Hosoya Z topological index ($R = 0.888$ and $S = 2.90$ °C).[39]

In Table V we summarized the results of multiple regression analysis based on the use of two variables, including our novel indices ^1Z and ^2Z. The top rows in Table V give the coefficient of regression R and the bottom rows give the standard error S for the pair of variables indicated at the top and at the left-hand side of the table.

TABLE IV

The best dozen single descriptor regressions for the boiling points in octanes (For description of the descriptors see ref. 39)

Descriptor	Standard error	Regression coefficient
Z	2.90 °C	0.888
$1/^2\chi$	2.91 °C	0.887
$^1\chi/^2\chi$	2.93 °C	0.886
$^2\chi$	2.98 °C	0.882
W/Z	3.09 °C	0.872
$^1\chi-^2\chi$	3.09 °C	0.872
^1Z	**3.10 °C**	**0.870**
$^2\chi-^3\chi$	3.11 °C	0.870
χ^{-1}	3.25 °C	0.857
AZV	3.53 °C	0.823
$^1\chi$	3.60 °C	0.821
$\chi(W)$	3.70 °C	0.809
$1/^1\chi$	3.78 °C	0.801

We see that by using variables ^1Z and ^2Z, we achieve one of the better results (R = 0.912 and S = 2.659 °C):

$$BP = -0.3428\ {}^1Z + 0.1631\ {}^2Z + 53.074$$

This is almost the same as the two parameter regression based on the Wiener number W and the topological index Z (R = 0.913 and S = 2.658 °C):

$$BP = 0.3071\ W + 1.5607\ Z + 94.149$$

and slightly worse than the best combination that uses Z and ^2Z!

TABLE V

The best two variable regressions for the boiling points in octanes. Top row gives the standard error (°C) and the bottom row the coefficient of regression

	Z	W/Z	^1Z	^2Z
W	2.658 0.913	2.745 0.906	2.719 0.908	4.535 0.718
Z	– –	2.665 0.912	2.835 0.900	2.642 0.914
W/Z		– –	2.703 0.909	2.860 0.898
^1Z			– –	2.659 0.912

TABLE VI

Experimental boiling points, calculated boiling points and the difference for the best two parameter regression (using Z and ^2Z)

Isomer	BP(exp)	BP(calc)	Diff.
n-octane	125.67	121.53	+4.14
2-methylheptane	117.65	115.63	+2.02
3-methylheptane	118.93	119.25	–0.32
4-methylheptane	117.71	117.62	+0.09
3-ethylhexane	118.53	121.25	–2.72
2,2-dimethylhexane	106.84	109.32	–2.48
2,3-dimethylhexane	115.61	115.21	+0.40
2,4-dimethylhexane	109.43	112.72	–3.29
2,5-dimethylhexane	109.10	110.60	–1.50
3,3-dimethylhexane	111.97	112.08	–0.11
3,4-dimethylhexane	117.73	118.34	–0.61
2-methyl-3-ethylpentane	115.65	116.22	–0.57
3-methyl-3-ethylpentane	118.26	119.43	–1.17
2,2,3-trimethylpentane	109.84	109.43	+0.41
2,2,4-trimethylpentane	99.24	103.56	–4.32
2,3,3-trimethylpentane	114.76	111.55	+3.21
2,3,4-trimethylpentane	113.47	111.07	+2.40
2,2,3,3-tetramethylbutane	106.47	102.04	+4.43

$$BP = 1.3796\ Z + 0.1235\ {}^2Z + 64.505$$

with $R = 0.914$ and $S = 2.643$ °C. In Table VI, we list the computed boiling points and the residuals for this best two variable regression. If we try regressions using three descriptors at a time, we do not improve the standard error, which increases in the range of 2.72–2.84 for various combinations of three descriptors (the best combination involves W, Z and ${}^2Z$ with $R = 0.915$ and $S = 2.719$ °C).

CONCLUDING REMARKS

The novel graph invariants ${}^1Z$ and ${}^2Z$ have passed the critical test mentioned before and they have been found useful descriptors already when considering the boiling points in octanes. Hence, the $\mathbf{Z}$ matrix, the matrix elements of which are structurally related to the topological index Z of Hosoya and the descriptor Z'/Z, appears to be a promising source of novel graph invariants. Because of the close relationship to Hosoya's Z topological index, we propose to refer to this matrix as Hosoya's matrix.

APPENDIX

The Hosoya matrices for the 18 isomers of octane (only the upper triangular part is shown since the matrices are symmetrical).

n–octane

0	20	12	7	4	2	1	0
	0	25	15	9	5	3	1
		0	23	14	8	5	2
			0	24	14	9	4
				0	23	15	7
					0	25	12
						0	20
							0

2-methylheptane

0	20	15	9	5	3	1	12
	0	23	14	8	5	2	20
		0	19	11	7	3	15
			0	20	13	6	9
				0	21	10	5
					0	17	3
						0	1
							0

3-methylheptane

0	17	12	9	5	3	1	7
	0	25	18	11	7	3	15
		0	24	14	9	4	20
			0	20	13	6	14
				0	22	11	8
					0	18	5
						0	2
							0

4-methylheptane

0	18	10	7	5	3	1	4
	0	21	15	11	7	3	9
		9	23	17	11	5	4
			0	23	15	7	20
				0	21	10	14
					0	18	9
						0	4
							0

3-ethylhexane

0	18	12	8	5	2	9	4
	0	25	17	11	5	18	9
		0	23	15	7	25	12
			0	22	11	17	8
				0	19	11	5
					0	5	2
						0	18
							0

2,3-dimethylhexane

0	17	14	8	5	2	12	12
	0	19	11	7	3	17	17
		0	14	9	4	14	14
			0	17	8	5	5
				0	13	5	5
					0	2	2
						0	12
							0

2,3-dimethylhexane

0	18	15	11	7	3	10	9
	0	23	17	11	5	18	14
		0	20	13	6	15	17
			0	19	9	11	11
				0	16	7	7
					0	3	3
						0	9
							0

2,4-dimethylhexane

0	18	13	9	7	3	11	5
	0	20	14	11	5	18	8
		0	19	15	7	13	11
			0	21	10	9	17
				0	14	7	13
					0	3	6
						0	5
							0

2,5-dimethylhexane

0	17	13	11	8	3	10	3
	0	20	11	8	5	17	5
		0	15	11	7	13	7
			0	20	13	11	13
				0	17	8	10
					0	3	10
						0	3
							0

3,3-dimethylhexane

0	13	10	8	5	2	7	7
	0	21	17	11	5	15	15
		0	20	13	6	18	18
			0	17	8	14	14
				0	15	9	9
					0	4	4
						0	12
							0

3,4-dimethylhexane

0	16	11	9	7	3	6	5
	0	21	18	14	7	13	11
		0	24	18	9	18	14
			0	21	11	14	18
				0	16	11	13
					0	5	6
						0	8
							0

2-methyl-3-ethylpentane

0	19	15	11	5	11	11	5
	0	23	17	8	19	17	8
		0	21	10	15	21	10
			0	16	11	15	7
				0	5	7	3
					0	11	5
						0	16
							0

3-methyl-3-ethylpentane

0	15	11	9	4	7	9	4
	0	22	18	9	15	18	9
		0	22	11	19	22	11
			0	15	15	18	9
				0	7	9	4
					0	15	7
						0	15
							0

2,2,3-trimethylpentane

0	16	14	11	5	11	11	8
	0	19	15	7	16	16	11
		0	17	8	14	14	13
			0	12	11	11	9
				0	5	5	4
					0	11	8
						0	8
							0

2,2,4-trimethlypentane

0	14	11	8	5	10	10	5
	0	15	11	7	14	14	7
		0	14	9	11	11	9
			0	13	8	8	13
				0	5	5	8
					0	14	5
						0	5
							0

2,2,3-trimethylpentane

0	15	13	11	5	8	9	9
	0	20	17	8	15	14	14
		0	19	9	13	16	16
			0	12	11	13	13
				0	5	6	6
					0	9	9
						0	10
							0

2,3,4-trimethylpentane

0	16	13	11	7	9	7	7
	0	20	17	11	16	16	11
		0	20	13	13	14	13
			0	16	11	11	16
				0	7	7	9
					0	7	7
						0	7
							0

2,2,3,3-trimethylbutane

0	12	11	8	8	8	8	8
	0	15	11	12	12	11	11
		0	12	11	11	12	12
			0	8	8	8	8
				0	8	8	8
					0	8	8
						0	8
							0

REFERENCES

1. M. Randić, *New J. Chem.* **15** (1991) 517; *J. Chem. Inf. Comput. Sci.* **31** (1991) 311.
2. M. Randić, *Acta Pharm.* (submitted).
3. M. Randić and N. Trinajstić, *J. Mol. Struct. (Theochem.)* (in press); N. Trinajstić, *Chemical Graph Theory*, 2-nd ed., CRC Press, Boca Raton, FL (1992) Chapter 10, pp 225; N. Trinajstić, M. Randić, and D. J. Klein, *Acta Pharm. Jugosl.* **36** (1986) 267; A. T. Balaban, I. Motoc, D. Bonchev, and O. Mekenyan, *Top. Curr. Chem,.* **114** (1983) 21.
4. POLLY: S. Basak (National Resources Research Institute, University of Minnesota – Duluth, Mn); L. H. Hall (Dept. of Chemistry, East Nazarine College, Quincy, Ma); ALL PATH: M. Randić, G. M. Brissey, R. Spencer and C. L. Wilkins, *Comput. Chem.* **3** (1979) 5, or: P. C. Jurs, *Computer Software Applications in Chemistry,* Wiley, New York 1986.
5. H. Wiener, *J. Amer. Chem. Soc.* **69** (1947) 17.
6. H. Hosoya, *Bull. Chem. Soc. Japan* **44** (1971) 2332.
7. M. Randić, *J. Amer. Chem. Soc.* **97** (1975) 6609.
8. L. B. Kier and L. H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Academic Press, New York 1976.
9. L. B. Kier, Indices of molecular shape from chemical graphs, in: *Computational Chemical Graph Theory* (D. H. Rouvray, ed.) Nova Sci. Publ. Commack, NY 1990.
10. L. B. Kier and H. L. Hall, *J. Pharm. Sci.* **65** (1976) 1806.
11. M. Randić, *Chemometrics & Intel. Lab. Sys.* **10** (1991) 213; M. Randić, *J. Comput. Chem.* **n12** (1991) 970;Randić and J. Cz. Dobrowolski, *J. Math. Chem.* (submitted).
12. M. Randić, P. J. Hansen, and P. C. Jurs, *J. Chem. Inf. Comput. Sci.* **28** (1988) 60.
13. K. Altenburg, *Kollood Z.* **178** (1961) 112; K. Altenburg, *Z. Phys. Chem.* (Leipzig) **261** (1980) 389.
14. I. Gutman, B. Ruščić, N. Trinajstić, and C. F. Wilcox, Jr., *J. Chem. Phys.* **62** (1975) 3399.
15. M. Randić, *J. Math. Chem.* **7** (1991) 155.
16. M. Randić, *Croat. Chem. Acta* **64** (1991) 43.
17. J. R. Platt, *J. Chem. Phys.* **15** (1947) 419.
18. L. B. Kier, W. J. Murrey, M. Randić, and L. H. Hall, *J. Pharm. Sci.* **65** (1976) 1806.
19. M. Randić, *Int. J. Quant. Chem., Quant. Biol. Symp.* **11** (1984) 137.
20. P. A. Filip, T. S. Balaban, and A. T. Balaban, *J. Math. Chem.* **1** (1987) 61.
21. F. Harary, *Graph Theory,* Addison-Wesley, Reading, Ma 1969.
22. M. Randić, *J. Chem. Inf. Comput. Sci.* **32** (1992) 686.
23. L. H. Hall, Computational; aspects of molecular connectivity and its role in structure-property modeling, in: *Computational Chemical Graph Theory*, D. H. Rouvray Ed., Nova Sci. Publ., Commack, NY 1990.
24. S. S. Tratch, M. I. Stankevich, and N. S. Zefirov, *J. Comput. Chem.* **11** (1990) 899.
25. D. J. Klein and M. Randić, *J. Math. Chem.* **12** (1993) 81.
26. M. Randić, *Chem. Phys. Lett.* (in print); M. Randić, Xiaofeng Guo, T. Oxley, and P. Krishnapriyan, *J. Math. Chem.* (submitted); L. W. Naylor and M. Randić, *Comput. Chem.* (submitted).
27. M. Randić, *Chem. Phys. Lett.* (submitted).
28. K. Balasubramanian, *ibid.* **169** (1990) 224; N. Trinajstić, *Chemical Graph Theory*, (second ed.), CRC Press, Boca Raton, FL., (1992.) pp 262.
29. M. Randić, *Int. J. Quant. Chem.: Quant, Biol. Symp.* **15** (1988) 201; M. Randić, B. Jerman-Blažič, and N. Trinajstić, *Comput. Chem.* **14** (1990) 237.
30. M. Randić, A. F. Kleiner, and L. DeAlba, *J. Chem. Inf. Comput. Sci.* (Proc. of the 5-th Intl. Conf. on Mathematical Chemistry, Kansas City (May 1993)
31. M. Randić, *J. Chem. Inf. Comput, Sci.,* **24** (1984) 164.
32. A. T. Balaban, *Chem. Phys. Lett.* **89** (1982) 399.
33. Xiaofeng Guo and M. Randić, *J. Math. Chem.* (to be submitted).
34. M. Randić and C. L. Wilkins, *J. Phys. Chem.* **83** (1979) 1525.
35. M. Randić, *J. Chem. Educ.* **69** (1992) 713.
36. A. T. Balaban, *J. Mol. Struct. (Theochem.)* **165** (1988) 243.
37. S. C. Basak, V. R. Magnuson, G. J. Niemi, R. R. Regal, and G. D. Veith, *Math. Modelling* **8** (1986) 300.

38. A. R. Katritzky and E. V. Gordeeva, *J. Chem. Inf. Comput. Sci.* (submitted)
39. M. Randić, *Croat. Chem. Acta* **66** (1993) 289.

SAŽETAK

Hosoyina matrica – izvor novih molekulnih deskriptora

Milan Randić

Uvedena je nova matrica za opis molekulnih grafova. Konstrukcija ove matrice povezana je s računanjem Hosoyina Z indeksa. Nova matrica, nazvana Hosoyinom matricom zbog veze sa njegovim topologijskim indeksom, služi kao izvor novih graf-teorijskih invarijanti, koje mogu biti važne pri proučavanju odnosa strukture i svojstava molekula. Pokazano je kako se Hosoyina matrica konstruira i kako se iz nje mogu izvesti novi topologijski indeksi. Usporedba sa srodnim topologijskim indeksima ilustrirana je na primjeru vrelišta oktana.