GENERALIZED MOLECULAR DESCRIPTORS*

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

Abstract

We review algebraic characterizations of molecular structures and in particular consider different matrices associated with a molecule as a source of novel graph invariants for use in structure–property and structure–activity studies. Such matrices can be classified as structure-explicit, structure-cryptic, and structure-implicit corresponding to a previous classification of molecular descriptors. In order to tame the proliferation of unwarranted topological indices, we propose requirements on indices and on the "source" matrices used for construction of molecular descriptors. Several structure-explicit and structure-implicit matrices are illustrated. A novel bond descriptor P'/P defined by the ratio of the number of paths in a graph G', in which an edge is erased, and in the parent graph G is introduced. The derived bond-additive molecular P'/P index, which correlates well with the octane numbers in octanes, was found to be linearly related to the Wiener numbers.

1. Introduction

To describe a chemical structure, one can list its *properties* or alternatively one can present its *name*, preferably based on structural elements. The former do not necessitate information on atom connectivities or on atomic coordinates, and may be viewed as an *output*, either of experimental measurements or theoretical computations on a molecule. In contrast, the atom–atom connectivities and the atomic coordinates can be viewed as molecular *input.* If codes are used to represent a structure, information is required as to how such codes relate to atomic labels so that the code leads to a reconstruction of the molecular graph. In mathematical terminology, properties are structural *invariants*, which means that they are independent of assumed labels for atoms (vertices in a molecular graph) and independent of pictorial representations of a structure. A short list of molecular invariants for a structure need not be unique because different structures may have some of the same properties, and as such may then lead to the same descriptors. One expects that different structures will differ in at least some of their properties, hence a need for consideration of additional descriptors. As a consequence of the non-uniqueness of descriptors, a list of molecular descriptors in general does not allow one to *reconstruct* a molecule. However, it appears that one

*Dedicated to Dennis H. Rouvray, one of the foremost promoters of chemical graph theory.

can considerably narrow down the póssibilities for candidate structures in some applications and thus succeed in a reconstruction [1]. In contrast, qualified molecular codes will not only permit a reconstruction but often reconstruction is not difficult [2–5].

Both, the codes (i.e. names, labels, indicators) as well as invariants, e.g. the so-called topological indices (which in fact are graph-theoretical indices), continue to proliferate. Hence, is seems the time has come to formulate requirements for these codes and invariants. Read [6] listed a dozen desirable requirements for molecular *codes*, which we reproduce in table 1. The list includes as the most important requirements

Table 1

List of desirable requirements for chemical codes as proposed by Read [6]

1	Codes should be linear strings of symbols
2	Codes should be unique
3	Reconstruction algorithms should be defined
4	Codes should be simple (preferably made by hand)
5	Decoding should be possible, preferably by hand
6	Nonsystematic (trivial) names should not be used
7	Properties should not be used
8	Codes should be brief
9	Codes should be pronounceable
10	Codes should be easily understood
11	Familiar symbols only should be used
12	Coding and decoding should be efficient
13	Similar structures should have codes of similar length (proposed by Randić in ref. [3])

that the code be unique, short and complete, i.e. that they allow reconstruction of a molecule. Such a list can be augmented with additional requirements such as: that codes of similar molecules are of similar length [3]. No extensive recommendations on desirable attributes for topological indices have been proposed in the past. Here, we want to correct the situation by offering a list of properties for molecular descriptors.

2. Desirable requirements on topological indices

Here, under "topological index" we understand a characterization of a molecule or a molecular graph by a *single number*. Clearly, a single number representation of a structure is accompanied by a considerable loss of information. However, what appears remarkable is to see how much of the relevant structural information can be absorbed by a single number "projection" of a structure. Of course, one can view N, the number of atoms in a molecule, and K, the number of Kekulé valence structures, as early topological indices, but they have been used primarily as molecular properties, not descriptors. Indeed, there are too many molecules with either the same *N* or the same *K* for these parameters to be considered as useful "descriptors".

In table 2, we list several better known topological indices, starting with Hosoya's *Z*, the number which counts nonadjacent pairs of bonds in a molecular structure [7]. This is the first topolgical index proposed in the literature *explicitly* as a single number representation of a chemical structure. Next is listed the connectivity index [8], the index based on discrimination of bonds into (m, n) bond types, where m, n are numbers of the nearest neighbors to the bonded atoms. Apparently, this is the most widely used descriptor in structure–property and structure–activity studies. We continue the list with the Wiener number *W*, which is the sum of all distances for all pairs of atoms in a structure [9] and which according to Platt [10] gives some measure of molecular volume. This particular index was the first nontrivial "topological" (i.e. graph-theoretical) molecular descriptor suggested, but as introduced initially it was not constructed to represent a structure by a single number (as was the case with Hosoya's *Z* number), but rather to be one of the *two* descriptors used in correlating several physico-chemical properties of alkanes and related structures. The other descriptor of Wiener was $P3$, the number of paths of length three. We could include in table 2 (as a single number descriptor) the combinations of path numbers, such

Table 2

Several better known topological indices

Descriptor	Structural interpretation	Author
Z	Count of non-adjacent bonds in a molecular graph	Hosoya
χ	Connectivity index: Sum of weighted bonds. Bond type (m, n) has weight $1/\sqrt{mn}$	Randić
W	Sum of path lengths	Wiener
J	Sum of weighted distances. Pair (i, j) has weight $1/\sqrt{D_{ij}}$	Balaban
P	Path numbers	Platt
ID	Identity numbers: Sum of weighted paths. Weights are given by $1/\sqrt{mn}$	Randić
C	Centric index: Based on the count of steps in pruning terminal bonds	Balaban
k	Shape index: Scaled on the differences in connectivity indices for extreme graphs	Kier
WID	Sum of weighted walks with weights given by $1/\sqrt{mn}$	Trinajstić

as $P2 - P3$ [11] and $P0 + P1 + P3$ [12], which suggests that linear combinations and other functional relationships of indices may be viewed as a single descriptor. Clearly, this opens novel possibilities, and further proliferation of topological indices may be anticipated. Hence, an evaluation of the great multiplicity of possible descriptors is needed, and we now present a list of desirable attributes for topological indices (table 3) to facilitate such evaluations.

Table 3

List of desirable attributes for topological indices

1	Direct structural interpretation
2	Good correlation with at least one property
3	Good discrimination of isomers
4	Locally defined
5	Generalizable to "higher" analogues
6	Linearly independent
7	Simplicity
8	Not based on physicochemical properties
9	Not trivially related to other indices
10	Efficiency of construction
11	Based on familiar structural concepts
12	Show a correct size dependence
13	Gradual change with gradual change in structures

Before discussing the outlined desirable attributes, we would like to suggest that topological *indices* and more general topological *descriptors* be differentiated. A descriptor is a more general quantity used in a characterization of a structure. When a descriptor satisfies the key requirements, in particular when it *alone* can account for at least a single molecular property, we should "promote" it to the status of a topological index or a molecular index.

We listed as the first criterion that an index has a *structural* interpretation. Only indices which are based on simple structural concepts will help one to interpret convoluted and complex properties in terms of the structure. In addition to being simple, an index has to be *useful* in structure-property correlations. When a descriptor correlates with a *single* molecular property, it indicates the dominant structural component for that molecular property; otherwise, indices are combined and will be viewed as auxiliary molecular descriptors. Accordingly, the Wiener number *W*, as introduced initially by Wiener, does not qualify as a topological index because on its own *W* was not noted to correlate with physico-chemical properties. Indeed, it was introduced as one of two topological descriptors needed to have a successful correlation. Subsequently, however, Rouvray [13] demonstrated that *W* correlates well with some properties of alkanes, by virtue of which we today prequalify *W* as a "legitimate" topological index. Next, topological indices ought to be isomer-sensitive, i.e. they ought to differentiate among isomers so that they can be used in the studies of isomer variations of molecular properties, and in general in the studies of those aspects of molecular properties which are size dependent. Many apparently successful regressions were reported on molecules of *different* size, and not surprisingly, many of these are found to be well represented by n, the number of atoms in a structure, which is a very good descriptor (even if not the only one) of molecular size. The challenges in structure–property studies are variations in properties among molecules of the *same* size [14]. Correlations where the dominant structural feature is size can often be successfully represented by trivial descriptors such as the number of atoms N, or the molecular weight MW. A search for "universal" regressions which incorporate size, shape and functionalities becomes unwarranted if individual structural features can be well investigated separately. There will be cases where the separation of the structural features such as size and shape, or shape and functionality fails, but these have first to be identified [14]. A route to characterization of situations where "separation of variables" fails is to study differences in characterization of molecules using global and local descriptors. Hence our emphasis on capability of topological descriptors to describe molecular *local* features. Finally, because in many situations a single descriptor will not suffice, it is of interest to investigate whether or not a family of structurally related descriptors can account for a property. Thus we require, when possible, that a topological index be generalized into "higher" indices which are to supplement the initial index and offer a more complete basis for a regression.

A recommendation that a novel topological index be "orthogonal" to the existing indices means that such an index leads to a correlation with a property not successfully analyzed with existing descriptors, or that it correlates with a residual in an acceptable regression based on other descriptors. Recently, the concept of orthogonal molecular descriptors has been introduced [16,17] which allows one to test descriptors and establish the degree of a "duplication" involved when several descriptors are combined in a single regression.

3. Systematic generation of graph invariants

Ad hoc descriptors are typically unrelated to each other. Many are introduced in an apparently arbitrary fashion, as illustrated by the classical case of Langevin [18], who assumed that an atomic or molecular magnet carried a permanent moment μ. The prime advantage of *ad hoc* descriptors is precisely that their apparent unrelatedness to the existing structural concepts makes them attractive candidates to "explain" not yet understood (in terms of structural concepts) molecular properties.

Another route to descriptors is to derive them systematically by *generalizing* the existing indices. Such are, for example, "higher" connectivity indices [19] which supplement the connectivity index χ in many regressions, and "higher" path numbers [20]. A way to generalize invariants based on matrix–vector multiplications was outlined by Balaban and coworkers [21]. As vectors, one can select graph-theoretical, quantum chemical or even empirical quantities (such as Van der Waals atomic radii, atomic weights, etc.). The distinction between these three types of vectors, corresponding to structure-explicit, structure-implicit, and structure-cryptic classification of molecular descriptors [22], is worth observing.

4. Matrices as a source of graph invariants

It is natural to extend considerations in deriving graph invariants from the adjacency matrix *A* and the distance matrix *D* to other graph matrices. Some molecular matrices received considerable attention in other areas of chemistry. In the analysis of infrared spectra *F* and *G*, matrices are well known, representing a molecular force field and inverse molecular geometry, respectively [23]. Hamiltonian matrices typically relate the pertinent contributing terms to the total molecular energy via the basis functions adopted. Other less frequently used molecular matrices include, to mention a few: Ugi and Dugundji's *BE* matrices [24], which include count of valence and lone pair electrons in a structure, inverse adjacency matrix (when it exists) [25], bond order and polarizability matrices of MO calculations [26], and Tutte's matrix [27] (which has been referred to also as Kirchhoff's matrix, in view of the fact that its minors enumerate spanning trees, the concept considered by Kirchhoff [28]).

In order to curtail proliferation of matrices to be considered as a source of topological indices, it seems prudent to extend the same recommendations listed for the topological indices to the matrices to be considered. Hence, we prefer that matrices, the entries of which have a direct structural interpretation, represent at least one particular molecular property, are sensitive to isomeric variations, can be generalized to "higher" analogues and, if possible, can be easily constructed. One need not insist, of course, that all these requirements be simultaneously satisfied; rather, they ought to be viewed as a guidance, desiderata. Hence, the generalized molecular matrices can be similarly classified into structure-explicit, structure-cryptic and structure-implicit, depending on whether the elements of the matrices are graph (structural) invariants, molecular properties or quantum chemical quantities, respectively. Molecular matrices, in general, can be referred to as a "through bond" and a "through space" type. The former type includes only "interactions" (or information) on adjacent atoms (vertices), and when a non-zero entry occurs in non-adjacent vertices, it is derived from the information on adjacent vertices only, while the latter type corresponds to situations where all interatomic (including all non-adjacent pairs) "interactions" are incorporated in a matrix.

4.1. STRUCTURE-EXPLICIT MATRICES

The adjacency and the distance matrix illustrate the class of structure-explicit matrices. The matrix of the graph-explicit class shown in table 4 is based on a measure of relative "importance" of an edge, a novel graph invariant. The "importance"

Table 4

Matrix based on graphical bond order P'/P and derived path counts illustrated on the molecular graph of 2-methylpentane

Matrix					
0	10/15	0	0	0	0
10/15	0	6/15	0	0	10/15
0	6/15	0	7/15	0	0
0	0	7/15	0	10/15	0
0	0	0	10/15	0	0
0	10/15	0	0	0	0

Atom	Paths			
	$P1$	$P2$	$P3$	$P4$
1	9.6667	0.7111	0.1244	0.0830
2	1.7333	0.1867	0.1244	
3	0.8667	0.8445		
4	1.1333	0.1867	0.2489	
5	0.6667	0.3111	0.1224	0.1659
6	0.6667	0.7111	0.1244	0.0830
Molecule	2.8667	1.4756	0.3733	0.1659
ID (sum)	4.8815			

is here defined as follows: For each edge (bond) we find, separately, the total number of paths in a subgraph obtained by erasure of the bond examined. If the derived subgraph is disjoint, the contributions of each component are added. Each edge is assigned a weight given by the ratio P'/P, where P' is the number (frequency) of paths obtained in subgraph G' in which the edge was deleted, and P is the number of paths in G. This novel local quantity is analogous to more general local descriptors recently considered by Balaban and coworkers [29]. It differs from them in that the molecular descriptor here is based on additivity, rather than on multiplicativity of fragment components. Construction of P'/P is illustrated in fig. 1 on a graph of 2-methylpentane.

The class of structure-explicit matrices includes the topographic matrices [30] in which entries represent the actual (three-dimensional) distances between vertices when the graph is embedded in a two- or three-dimensional grid. Such matrices can be viewed as geometry-based matrices, where bond lengths and bond angles are idealized, restricted by the geometry of the coordinate grid, such as a graphite lattice or a diamond lattice, respectively.

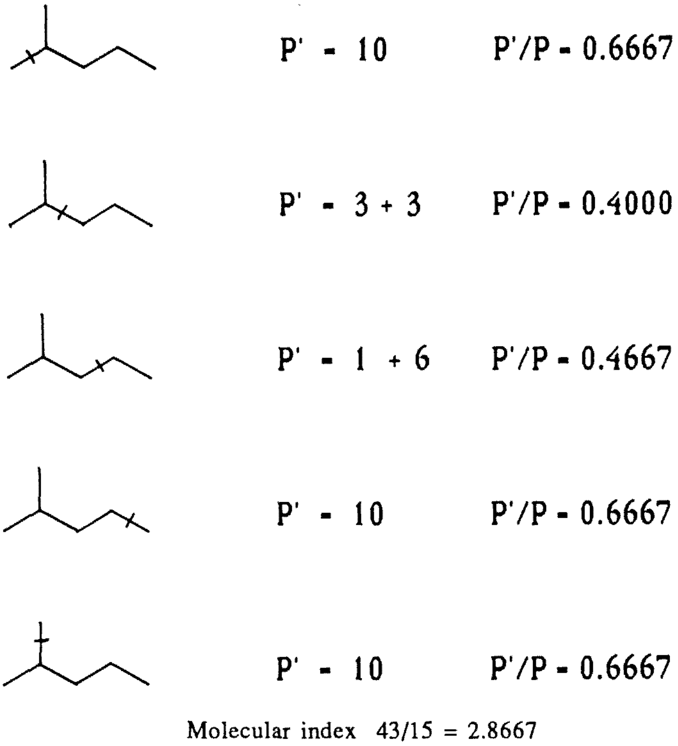

Fig. 1. Illustration of a construction of the graphical bond order P'/P for 2-methylpentane.

4.2. STRUCTURE-CRYPTIC MATRICES

When known experimental atomic and bond properties are used directly in a construction of matrices, we obtain structure-cryptic matrices with entries representing some molecular property. Alternatively, if a molecular property is *partitioned* into atomic and bond contributions, one can use such data for a construction of a property-related structure-cryptic matrix. Matrices, the entries of which record molecular properties, local as well as global, were considered already fifty years ago as potentially interesting objects in chemistry by Balandin [31].

4.3. STRUCTURE-IMPLICIT MATRICES

As an illustration of structure-implicit matrices, we consider the bond order matrix (table 5) and the bond overlap matrix (table 6). In the first case, the elements of the matrix are bond orders such as, for example, defined by Coulson [32] or Pauling [33] for MO and VB calculations, respectively. The matrix in table 5 represents bond orders of benzene based on Coulson's MO bond orders. Observe that entries in a bond order matrix may be negative, hence the "count" of paths can

Table 5

Bond overlap matrix and derived path counts illustrated on the benzene graph

Bond order matrix

1.0000	0.6667	0	– 0.3333	0	0.6667
0.6667	1.0000	0.6667	0	– 0.3333	0
0	0.6667	1.0000	0.6667	0	– 0.3333
– 0.3333	0	0.6667	1.0000	0.6667	0
0	– 0.3333	0	0.6667	1.0000	0.6667
0.6667	0	– 0.3333	0	0.6667	1.0000

Atomic paths

*P*1	*P*2	*P*3	*P*4	*P*5
1.000000	0.000133	– 0.44438	0.296326	0.296346

Molecular paths

3.0000	0.000400	– 1.33333	0.88898	0.88904

Total number of paths: 3.44558

Table 6

Maximum overlap matrix for norbornane and derived path counts

Bond overlap matrix

0	*A*	0	0	0	*A*	*B*	
A	0	*C*	0	0	0	0	
0	*C*	0	*A*	0	0	0	*A* = 0.6431
0	0	*A*	0	*A*	0	*B*	*B* = 0.6389
0	0	0	*A*	0	*C*	0	*C* = 0.6445
A	0	0	0	*C*	0	0	
B	0	0	*B*	0	0	0	

Paths for nonequivalent vertices

Vertex	*P*1	*P*2	*P*3	*P*4	*P*5	*P*6
a	1.9521	1.2371	1.0581	1.0218	0.2210	
b	1.2876	1.2389	1.0604	0.8500	0.5478	0.1407
c	1.2778	1.6435	1.0592	0.6812	0.4381	0.2823

Molecular paths

5.1392	4.5368	3.708	3.0625	1.5357	0.4226

Total number of paths: 18.4053

become a negative quantity. The bond overlap matrix is illustrated for the case of norbornane, a bicyclic C_7H_{12} hydrocarbon. The entries A, B, C in the bond overlap matrix (table 6) represent carbon–carbon bond overlaps computed using the maximum overlap method [34,35].

5. Illustration

Balaban and Motoc [36] investigated correlations with octane numbers of numerous topological indices in alkanes having from four to eight carbon atoms. The following is the summary of the statistics of single variable regressions for a few of the descriptors considered by Balaban and Motoc [36]:

Wiener W number	$R = 0.702$	$S = 7.17$
Centric index	$R = 0.940$	$S = 8.44$
Paths of length 2	$R = 0.882$	$S = 11.62$
Connectivity χ	$R = 0.778$	$S = 15.50$
Hosoya's Z index	$R = 0.957$	$S = 17.52$

The centric index, introduced by Balaban [37] as a measure of centrality for acyclic graphs, is defined via the number of steps needed to prune terminal edges of a graph. Apparently, from the above it appears that the centric index is the best among about a dozen descriptors tested. Notice that the highest correlation coefficient (R), achieved by Z, shows the worst standard error (S), while the poorest correlation coefficient, the Wiener number, shows the smallest standard error! The role of the individual descriptor thus remains obscure, in particular as seen above, it is not clear how well the individual descriptors characterize molecular size and how well they vary with variations in shape.

In order to see to what extent the results of Balaban and Motoc are dominated by the size dependence and to what extent they are due to isomeric variations of molecular shapes, we restricted attention to octane isomers only. The revised regression analyses for the same selected descriptors, limiting the sample to octane isomers, are summarized in table 7.

Overall, one still obtains similar regression statistics, with the exception of a dramatic improvement in R for the Wiener number, and a visibly reduced correlation coefficient for Z. The connectivity index and the Hosoya index in the case of isomeric variations of octane numbers account for less than 50% of the variance in the data. Not surprisingly, these indices do not appear successful here because they reproduce the *ordering* of other properties, such as heats of formation and boiling points, which as is known do not correlate with octane numbers.

We augmented the analysis by adding two novel descriptors A'/A and P'/P. A construction of the P'/P descriptors is illustrated in fig. 1 for 2-methylpentane. The A'/A index is similarly constructed with the difference that here also the path

Table 7

Regression equations, the correlation coefficients (R) and the standard errors (S) for correlations between selected topological descriptors and octane numbers in octane isomers

Descriptor[a]	Regression equation		R	S
	Variable	Constant		
W	- 5.15	431.83	0.954	7.41
C	12.57	21.93	0.934	8.81
$P2$	21.07	- 102.80	0.850	13.02
χ	- 140.29	584.50	0.688	17.93
Z	- 4.09	178.15	0.609	19.59
A'/A	141.99	- 635.93	0.181	14.20
P'/P	144.31	- 578.06	0.954	7.41

[a] W: Wiener number. C: Centric index by Balaban. $P2$: Paths of length 2. χ: Connectivity index. Z: Topological index by Hosoya. A'/A: Path numbers including atomic contributions. P'/P: Path numbers.

Table 8

P'/P values for octane isomers, experimental and calculated

Octane	P'	P'/P	Octane number	
			Exp.	Calc.
n-octane	128	4.0000	–	0.09
2-M	117	4.1786	23.8	25.48
3-M	120	4.2857	35.0	40.71
4-M	121	4.3214	39.0	45.78
2, 5-MM	122	4.3571	55.7	50.86
3-E	124	4.4286	52.4	61.02
2, 4-MM	125	4.4643	69.9	66.10
2, 2-MM	125	4.4643	77.4	66.10
2, 3-MM	126	4.5000	78.9	71.17
3, 4-MM	128	4.5714	81.7	81.32
3, 3-MM	129	4.6071	83.4	86.40
2-M, 3-E	129	4.6071	88.1	86.40
2, 3, 3-MMM	130	4.6429	99.4	91.49
2, 2, 4-MMM	130	4.6429	100.0	91.49
2, 3, 4-MMM	131	4.6786	95.9	96.56
3-M, 3-E	132	4.7143	88.7	101.64
2, 2, 3-MMM	133	4.7500	99.9	106.71
2, 2, 3, 3-MMMM	138	4.9286	–	132.10

of length zero, i.e. atomic contributions, is included in the count. In table 8 are listed, for octane isomers, their octane numbers and the molecular P'/P indices, obtained as outlined by adding all P'/P for individual bonds. In addition, table 8 gives the computed octane numbers as derived from the regression equation. The best regressions of table 7 are based on P'/P and on the Wiener number. Observe that both regressions have somewhat better R value and smaller S value than the correlation of Balaban and Motoc using the centric index. The corresponding R and S values for P'/P and W appear to be the same, which raises a suspicion that the two quantities may be related or are highly correlated. To test this, we derived a regression of the novel P'/P descriptors against the Wiener numbers. One obtains:

$$P'/P = -0.035685\,W + 6.997889, \quad \text{with } R = 1.0000 \text{ and } S = 0.00019.$$

This can be recognized and rewritten as:

$$P'/P = -(1/28)W + 7 \quad \text{or} \quad W = 196 - P',$$

since for octane, $P = 28$. Hence, the descriptor P'/P, although novel and defined without references to the Wiener number, is nevertheless simply, although not trivially, related to the Wiener number W, at least in the case of acyclic graphs!

It is interesting to see how two different concepts, the Wiener number, which is a global descriptor paralleling molecular volume according to Platt [10], and the bond additive local descriptor P'/P, which is sensitive to the bond environment, are linearly related. Even with this hindsight, the close relationship between the two descriptors is not obvious. Recollect that according to Wiener [9]: "The path number W is defined as the sum of the distances between any two carbon atoms in the molecule, in terms of carbon–carbon bonds. Brief method of calculation: *Multiply* the number of carbon atoms on one side of any bond by those on the other side; W is the sum of these values for all bonds." In contrast, a construction of P' can be summarized briefly as: *Add* the number of paths on one side of any bond to those on the other side; P' is the sum of these values for all bonds. Although the total number of paths (in trees) is simply related to the number of vertices, the two algorithms involve different operations, multiplication versus addition, yet P' remains linearly related to W.

6. Concluding remarks

We are optimistic of the future expansion of the regression analysis in structure–property and particularly structure–activity studies. Our optimism is mainly based on the potentiality of the recently introduced methodology of orthogonalized molecular descriptors. Orthogonal descriptors, in addition to their proven numerical stability [38], allow one to evaluate the statistical significance of each individual descriptor and discard those that are of marginal importance. Even though the approach based on orthogonal descriptors bears a subtle resemblance to the Principal Components Analysis (PCA), it is conceptually and computationally different and deserves its own label. Hence, we propose to refer to it as the Dominant Components Analysis (DCA). The parallelism in the naming of the procedure with the well-known PCA is, of course, deliberate and very appropriate in view of the fact that DCA can identify *dominant* components in a multivariate regression. Moreover, due to the constancy of the coefficients of the regression equation, the dominant structural factors in a regression can be interpreted. Hence, we may be at the beginning of a new direction in applications of multivariate regression analysis in structure–property and structure–activity studied. Our optimism is further enhanced by the following:

(1) Continuing use of algebraic descriptors and other mathematical invariants in regression analysis, thus avoiding meta-structural characterizations of molecules via their own properties, which would leave the nature of such descriptions structure-cryptic.

(2) Augmenting the existing descriptors matrices describing other than adjacency relationships in graphs as a source for invariants. It is advantageous to use structure-explicit and structure-implicit matrices as a source for matrices based on molecular properties when an understanding of structure–property or structure–activity relationships is intended. Descriptors based on matrices of molecular properties (disregarding that these may be subject to limitations due to experimental errors) also have their use. As a matter of fact, at least for a discussion of isomeric variations, as has been found recently, properties are by far less intercorrelated among themselves than was hitherto believed to be the case [14]. Hence, descriptors based on properties are likely to cover adequately the "structure-space" of molecules investigated.

Acknowledgement

I would like to thank Professor D.J. Klein (Texas A&M at Galveston) for numerous comments and suggestions that resulted in an improved presentation of the material.

References

[1] I.I. Baskin, E.V. Gordeeva, R.O. Devdariani, N.S. Zefirov, V.A. Palyulin and M.I. Stankevich, Doklady Akad. Nauk. SSSR 307(1989)613.

[2] J.V. Knop, W.R. Müller, Z. Jeričević and N. Trinajstić, J. Chem. Inf. Comput. Sci. 21(1981)91.

[3] M. Randić, J. Chem. Inf. Comput. Sci. 26(1986)136.

[4] M. Randić, J. Chem. Inf. Comput. Sci. 17(1977)171.

[5] R.C. Read, in: *Graph Theory and Computing*, ed. R.C. Read (Academic Press, New York, 1972), p. 153.

[6] R.C. Read, J. Chem. Inf. Comput. Sci. 23(1983)135.

[7] H. Hosoya, Bull. Chem. Soc. Japan 44(1971)2332.

[8] M. Randić, J. Amer. Chem. Soc. 97(1975)6609.

[9] H. Wiener, J. Amer. Chem. Soc. 69(1947)17.

[10] J.R. Platt, J. Phys. Chem. 56(1952)328.

[11] M. Randić and N. Trinajstić, Theor. Chim. Acta 73(1988)233;
Y. Miyashita, T. Okuyama, H. Ohsako and S. Sasaki, J. Amer. Chem. Soc. 111(1989)3469.

[12] Y. Miyashita, H. Ohsako, T. Okuyama, S. Sasaki and M. Randić, Magn. Res. Chem., submitted.

[13] D.H. Rouvray, in: *Mathematics and Computational Concepts in Chemistry*, ed. N. Trinajstić (Ellis Horwood, Chichester, 1986), ch. 25.

[14] M. Randić and P.G. Seybold, J. Amer. Chem. Soc., submitted.

[15] M. Randić, J. Chem. Inf. Comput. Sci., in press.

[16] M. Randić, New J. Chem., in press.

[17] M. Randić, Croat. Chem. Acta, in press.

[18] J.H. Van Vleck, Science 201(1978)113.

[19] L.B. Kier, W.J. Murray, M. Randić and K.H. Hall, J. Pharm. Sci. 65(1976)1226;
L.B. Kier and L.H. Hall, *Molecular Connectivity in Chemistry and Drug Research* (Academic Press, New York, 1976).

[20] M. Randić, Int. J. Quant. Chem.: Quant. Biol. Symp. 11(1984)137.

[21] P.A. Filip, T.-S. Balaban and A.T. Balaban, J. Math. Chem. 1(1987)61.

[22] N. Trinajstić, M. Randić and D.J. Klein, Acta Pharm. Jugosl. 36(1986)267.

[23] E.B. Wilson, J.C. Decius and P.C. Cross, *Molecular Vibrations* (McGraw–Hill, New York, 1955), ch. 8, pp. 169–231.

[24] J. Dugundji and I. Ugi, Topics Curr. Chem. 39(1973)19.

[25] E. Heilbronner, Helv. Chim. Acta 45(1962)1722.

[26] C.A. Coulson and A. Streitwieser, *Dictionary of Pi-Electron Calculations* (Pergamon, Oxford, 1965).

[27] W.T. Tutte, *Graph Theory*, Encyclopedia of Mathematics, Vol. 21 (Addison–Wesley, Menlo Park, CA, 1984), section 6.4, p. 138. In mathematical literature, it is also referred to as a Laplace matrix.

[28] G. Kirchhoff (translation), IRE Trans. Circuit Theory 5(1958)4.

[29] O. Mekenyan, D. Bonchev and A.T. Balaban, J. Math. Chem. 2(1988)347.

[30] M. Randić, *MATH/CHEM/COMP 1987*, ed. R.C. Lacher, Studies in Phys. Theor. Chem. 54(1988)101;
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.

[31] A.A. Balandin, Uspekhi Khem. 9(1940)390.

[32] C.A. Coulson, *Valence* (Oxford University Press, 1952).

[33] L. Pauling, *The Nature of the Chemical Bond* (Cornell University Press, Ithaca, 1960).

[34] M. Randić, Int. J. Quant. Chem. 8(1974)643.

[35] M. Randić and S. Borčić, J. Chem. Soc. (1967)596.

[36] A.T. Balaban and I. Motoc, MATCH 5(1979)197.

[37] A.T. Balaban, Theor. Chim. Acta 53(1979)355.

[38] A.F. Kleiner and M. Randić, work in progress.