Russian Chemical Reviews

March 1988

Translated from *Uspekhi Khimii*, **57**, 337-366 (1988)

U.D.C. 547.8:513.83:519.1

Topological Indices in Organic Chemistry

M.I.Stankevich, I.V.Stankevich, and N.S.Zefirov

The numerical characteristics of molecular structures, called topological indices, which are used for the quantitative description of the structures of organic compounds, are examined. The method of topological indices is analysed as one of the procedures for determining the relation between the structure of chemical compounds and their properties.
The bibliography includes 164 references.

CONTENTS

I.	Introduction	191
II.	Topological indices as quantitative characteristics of molecular structures	193
III.	The use of topological indices for the estimation of physicochemical properties of compounds	199
IV.	Topological indices and biological activity	201
V.	Quantum-chemical methods as a source of topological indices	202

I. INTRODUCTION

The search for regularities in the type of variation of different properties of molecules as a function of their structure belongs to the so called "structure–property" problem which has long attracted the attention of chemists.[1-2] The solution of this problem even for an individual fairly narrow range of compounds is extremely important.[3] The observed regularities can be used for the systematic treatment of molecules of a given class, for the prediction of the properties of hypothetical molecules, and for the search for new compounds with a specified set of properties. Furthermore, they can also prove useful in the development of a theory which would explain the phenomenon under consideration and would provide a key to the understanding of the mechanism of the action of biological active compounds.

In order to find quantitative relations between structure and property, various mathematical modelling methods are used, their choice being determined by the type of chemical compounds, their state of aggregation, and the property being analysed. Procedures for the description of molecular structures occupy a special place among these methods.

The description of molecular structure can be detailed to different extents. Thus the overall formula shows how many and what atoms are contained in the molecule. The structural formula contains information also about the way in which these atoms are linked to one another. Finally, the specification of the coordinates of the atomic nuclei in space makes it possible to construct three-dimensional models of molecules. With the aid of quantum-chemical calculations, it is possible to obtain additional information about the distribution of electron density in the molecule. The physicochemical parameters reflecting the molecular structure indirectly are used in many cases to characterise the molecules.

Modern experimental research methods make it possible to determine fairly reliably the geometry and different physicochemical characteristics of molecules. However, such methods are applicable only to substances which actually exist and not to all substances. In those cases

where experiment is difficult, one can use quantum-chemical calculations of the geometrical and energy characteristics of molecules. Such calculations are still inferior as regards accuracy to experimental data and, moreover, they are associated with considerable computational difficulties. More formal and simpler methods, based on the description of the molecular structure by the structural formula, are therefore also used nowadays to solve the "structure–property" problem. In this topological method for the description of molecular structure, account is taken only of the types of atoms and of the character of their bonding to one another, but the metric characteristics of the molecule are neglected.

The methods based on the analysis of the structural formula in order to discover the relation between the structure and property have long been used in theoretical organic chemistry. For example, the additive schemes employed to estimate the physicochemical properties of hydrocarbons, based on the concept of a numerical characteristic of the property investigated in the form of the sum of the contributions by atoms and bonds, are widely known.[4,5] Further development of the topological approach led to the appearance of various modifications of simple additive schemes associated with allowance for a non-valence interaction, the use of more complex structural fragments, the application of non-linear functional relations between the property and the structural parameters,[6–10] etc. This made it possible to extend significantly the list of properties investigated and also to analyse compounds with a structure more complex than that of hydrocarbons.

The development of the topological approach is associated to a large extent with the application of the methods of graph theory,[11] since the structural formula of a molecule is equivalent to the corresponding molecular graph. The analysis of the structural formula which must be carried out within the framework of the "structure–property" problem therefore reduces to the treatment by the methods of the theory of graphs of the corresponding abstract mathematical structure—the molecular graph (MG). Weighted MG are usually employed in structural chemistry.[10] The weighting factors for the vertices and edges are chosen in order to take into account the differences between the types of atoms and bonds. The weighted graph is defined by the adjoining matrix $A = (a_{ij})$, $(i,j = 1,\ldots,n)$, where a_{ii} is the weighting factor of the ith vertex, $a_{ij} = a_{ji}$ the weighting factor of the edge (i,j), $a_{ij} = a_{ji} = 0$ if the vertices i and j are not connected by an edge, and n is the number of vertices in the graph. The form of the adjacency matrix depends, generally speaking, on the way in which the vertices of the graph are numbered. However, different numbers characterising its structure and independent of the way in which the vertices are numbered can be attributed to each graph. Such numbers are called the invariants of the graph or structural constants. Examples of the graph invariants are the number of vertices, the number of edges, the number of paths of equal length, the determinant of the adjacency matrix, etc. In organic chemistry, the invariants of weighted molecular graphs are called topological indices (TI). Topological indices have found a wide variety of applications in structural chemistry. In particular, they can be used to code chemical information, in designing a chemical experiment, in the theory of the atomic structure and reactivity of molecules, and for the quantitative description of chemical structures in the analysis of the relation between the structure of a molecule and its properties. In the present review, we shall discuss the last two applications of the TI.

As a rule, in order to establish a relation between the structural parameters $x_i, \ldots, x_k$ (in the given case topological indices) and the numerical characteristics of a property "y", a functional relation of the type $y = a_0 + \sum_i a_i x_i$, where a_i are unknown constants determined by the method of least squares, is sought. The quality of the model constructed is characterised by the regression parameter s (the standard deviation of the regression line; $s \geq 0$) and r (the correlation coefficient; $0 \leq r \leq 1$). In more general cases, non-linear functional relations are used.[12]

Another approach to the "structure–property" problem associated with TI is also possible. Its essential feature consists in the fact that the concept of "distance" between the structures is introduced on the basis of the TI, making it possible to estimate the similarity of the structures and to relate it to changes in the property considered. Furthermore, TI are used as structural descriptors in the solution of the "structure–property" problem by other methods, for example, by the image recognition method.[12]

The approach to the solution of the "structure–property" problem based on TI has a number of advantages, which include, firstly, the simplicity and rapidity of the calculation of TI even on small computers and, secondly, the non-empirical character of the indices and the possibility of obtaining them directly from the structural formula. The latter is especially significant when one is dealing with the estimation of the properties of hypothetical structures for which there are no experimental data whatsoever.

The method of topological indices, based on the formal description of the molecular structure, suffers also from an evident disadvantage: not all the characteristic features of the molecular structure are taken into account. Nevertheless, in many cases TI are satisfactorily correlated with both steric parameters (molecular volume, molecular refraction, the surface area of the molecule) and the electronic parameters (ionisation potentials, electron affinity, polarisability, chemical shifts in NMR spectra, spin density, etc.). Furthermore, in constructing TI one can take into account not only the topology of the molecule but also certain features of its electronic and steric structures, the weighting factors for vertices and edges of the MG being appropriately selected.

One should also note that TI are usually obtained with the aid of formal operations on the MG. It is therefore not always possible to achieve a satisfactory physical interpretation of certain TI and to explain the correlation obtained with their aid even at a qualitative level. In the solution of the specific problems with the aid of the TI method, the problem of the motivated [subjective?] selection of topological indices arises.

However, in certain applied fields, in particular in the search for new biologically active substances,[12,13] a fairly complex situation arises, associated with the necessity to synthesise and investigate a large number of chemical compounds. Therefore any even indirect considerations which enable the investigator to solve *a priori* the question of the structures which can have the required properties or those which need not be considered at all is of great interest. The TI method can be used in the solution of problems of this type as a "coarse filter" for the preliminary sorting of molecules.

Recently there has been considerable growth of interest in topological indices. Numerous publications on this topic have been published in which increasing numbers of new TI have been introduced (see, for example, Bonchev[14]). Naturally, the abundance of TI combined with the lack of a single theory makes it difficult to understand them, promotes the appearance of speculative publications in this field, and even to some extent discredits the TI method.

The main aim of the review is to carry out an approximate classification of the TI described in the chemical literature, to give an account of the general principles of the construction of TI, and to analyse the real possibilities of the TI

method as a procedure for the description of chemical structures in the solution of the problem of the "structure–property" relation. The review also discusses the relation between different indices and their interpretation. The individual Sections are devoted to the discussion of correlations of TI with physicochemical properties and with the biological activity of certain classes of compounds known from the literature.

There exists a series of foreign reviews devoted to the use of TI in structural chemistry.[14–19] However, these publications cover inadequately the different applications of TI and provide as a rule merely a description of the factual data without its analysis.

II. TOPOLOGICAL INDICES AS QUANTITATIVE CHARACTERISTICS OF MOLECULAR STRUCTURES

This Section gives a brief description of the principal ways of constructing TI used in the analysis of molecular structure.

1. Invariant Sets of Molecular Graphs

We shall denote by G an arbitrary molecular graph. Different sets of numbers (numerical sets) can correspond to the graph G, characterising its structure and being independent of the numbering of the vertices. Each such set $\mathfrak{M}$ will be called an invariant set (IS) of the graph G assuming, to make the treatment concrete, that its elements are arranged in a decreasing sequence. For example, the set of the elements of the adjoining matrix $A=(a_{ij})$ of the graph G as well as the set of degrees of vertices v_i, where

$$v_i=\sum_{j,j\neq i} a_{ij}.$$

are invariant sets. When the weighting factors of the edges are equal to the multiplicities of the corresponding bonds, the quantity v_i is equal to the number of edges incident with the vertex i and related directly to the valence of the ith atom.

Together with the adjacency matrix, other matrices corresponding to the graph G can be used to construct its IS. The distance matrix $D = (d_{ij})$ $(i,j = 1, \ldots, n)$, whose element d_{ij} is equal to the distance between the ith and jth vertices, is most often used in structural chemistry. The quantity d_{ij} is defined as the minimum number of edges which must be traversed when moving along the edges of the graph from the ith vertex to the jth. The set of elements of the matrix D, the set of the distance degrees of the vertices d_i, where

$$d_i=\sum_j d_{ij},$$

and also the spectra of the matrices A and D are IS. Other examples of IS will be examined below.

A whole family of different TI corresponds to each IS $\mathfrak{M}$. We shall describe several ways of constructing TI of this type.

First method. Suppose that the IS $\mathfrak{M}$ of the graph G consists of k elements $m_1, \ldots, m_k$. We shall select an arbitrary symmetrical function $F(x_1,\ldots,x_k)$ of k variables (the values of this function do not change in arbitrary transpositions of the arguments). The quantity $F(m_1,\ldots,m_k)$, regarded as the value of the function F on the graph G, is a TI corresponding to $\mathfrak{M}$. For example, if $\mathfrak{M}$ consists of the element of the adjacency matrix A of the graph G, $k = n^2$ (where n is the number of vertices of the graph G), and

$$F(x_1,\ldots,x_k)=\sum_i x_i,$$

then $F(m_1,\ldots,m_k)$ is equal to the sum of all the elements of the matrix A. If the graph G is simple, i.e. A is a matrix consisting of zeros and ones [a(0,1)-matrix] and $a_{ii}=0$, then $F(m_1,\ldots,m_k) = 2q$, where q is the number of edges of the graph. By altering the form of the function F, it is possible to construct a whole series of other TI also related to the given IS $\mathfrak{M}$. . In yet another example, we may specify a TI generated by the function

$$F_a=\sum_i \delta(x_i),$$

where $\delta(x_i) = 1$ for $x_i = a$ and $\delta(x_i) = 0$ for $x_i \neq a$, a being a fixed number. The value of the function F_a at the point $(m_1,\ldots,m_k)$ is equal to the number of elements of the set $\mathfrak{M}$ identical with a. Thus, if $a = 2$ and the IS consists of the elements of the matrix D, then the value of the function F_a is equal to twice the number of pairs of the vertices of the graph located at a distance of 2.

Second method. We differentiate in the graph G the subgraphs G_j of a specific type with the same numbers of vertices m $(m < n)$. We relate to the subgraph G_j the IS $\mathfrak{M}_j$ consisting of k elements: $m_1^{(j)},\ldots,m_k^{(j)}$. We select a symmetrical function $F(x_1,\ldots,x_k)$ of k variables. For each subgraph G_j, we find the value of $F(m_1^{(j)},\ldots,m_k^{(j)})$ and then we determine the TI equal to the sum of these quantities. In particular, if $F(x_1,\ldots,x_k) \equiv 1$, then the TI determined by the method described is equal to the number of subgraphs of the given type.

Third method. New TI can be obtained from those already known by adding, multiplying, subtracting, and dividing the latter. Finally, it is possible to construct a multi-index—an ordered sequence of simple indices.

The TI described above characterise the graph as a whole. Such TI are represented in many cases as the sums of the numerical characteristics of different structural fragments, for example, the vertices or edges of the MG and these characteristics are independent of the numbering of the vertices of the fragment. Some of these invariants, called TI of the local type, are of independent interest and are used in the solution of various problems in theoretical chemistry. For example, the local TI determined as the matrix elements of different functions of the adjacency matrix are used in the theory of the electronic structure of conjugated molecules.

2. Indices Related to the Distance Matrix D

Correct estimates of the dimensions and degree of branching of the molecule are important in the estimation of a number of physicochemical properties of compounds, for example, the boiling point or melting point, because the energy of intermolecular interactions actually depends to a considerable degree on these parameters. In the topological approximation, mainly indices characterising the "size and form" of the graph are used for the solution of this problem. Some of them are directly related to the distance matrix D. This approach was first applied by Wiener,[20] who suggested that two TI, W and p, defined in terms of the distances between the vertices of the corresponding molecular graphs, be used for the comparison of the carbon skeletons of saturated hydrocarbons. The index p is equal to the number of pairs of vertices the distance between which is 3 and W is equal to the sum of the distances between all pairs of vertices of the graph G. The number W, usually called the Wiener index, is expressed in terms of the elements of the distance matrix D by the formula

$$W=\frac{1}{2}\sum_{i,j}d_{ij}.$$

The index W can also be written in the form

$$W=\sum_{j} jg_j,$$

where g_j is the number of pairs of indices separated from one another by the distance j, i.e. the number of elements d_{kl}, $(k<l)$ in the matrix D equal to j.[21] The number p is identical in the given instance with the quantity g_3. The polynomial

$$W(x)=\sum_{j} x^j g_j,$$

introduced by Altenburg,[22] is sometimes used to characterise the distances between the vertices of the MG which can occur in the graph. In terms of the $W(x)$ polynomial, the Wiener index is calculated by the formula

$$W=\frac{dW(x)}{dx}\Big|_{x=1}.$$

The distance matrix is also used to determine the distance polynomial

$$D(x)=(-1)^n \det(D-xE),$$

where E is a unit matrix.[23] This polynomial is the invariant polynomial of the graph.

The indices $J(G)$ and $D^{(k)}$, associated with the matrix D, are described by the formulae†[13]

$$J(G)=\frac{q}{1+c}{}'\sum (d_i d_j)^{-1/2}$$

$$D^{(k)}=\left[\left(\sum_i g_i\cdot i^k\right)\Big/\left(\sum_i g_i\right)\right]^{1/k};\ k=1,2,3,4,$$

where q is the number of edges and c is the cyclomatic number which in the case of simple graphs is equal to $q-n+1$.

In constructing certain TI, the concept of the centre of the graph is used.[11] The centre of the graph is defined as follows. For each vertex of the graph, one calculates initially the quantity $e_i=\max\limits_j d_{ij}$, where i is the number of the vertex. The quantity e_i is called the eccentricity of the ith vertex. Then, among all the numbers e_i, the smallest is selected: $r(G)=\min\limits_i e_i$. This number is called the radius of the graph. By definition, the centre of the graph is understood as the set of vertices whose eccentricity is $r(G)$. For example, the tree has a centre consisting of one or two connected vertices.[11] The centric TI for the tree is determined by means of the following procedure. In the first step, all the first degree vertices (pendant vertices) and the edges incident with them are cut off from the tree. In the second step, all the first degree vertices and the edges incident with them are again cut off from the subgraph obtained in the first step. This procedure is continued until finally only the central vertices of the tree remain. We shall denote by Δ_i the number of vertices eliminated in the ith step and shall put

$$B=\sum_i \Delta_i^2.$$

The quantity B is called the centric Balaban index.[24] The normalised (B') and binormalised (B'') centric indices are also defined in terms of the index D:

$$B'=0.5(B-2n+u),\ B''=2B'((n-2)^2-2+u)^{-1},$$

where $u=0.5\,[1-(-1)^n]$. A definition of the TI B_c, also related to the concept of the centre of the graph, has been proposed.[25] The index B_c is defined in terms of the formula

$$B_c=\sum_i r_i^2,$$

where r_i is the distance from the ith vertex to the centre. The centric indices B and B_c can also be defined in a different way in terms of the IS of the graph. Suppose that the set $N=(i_1,\ldots,i_k)$ consists of the numbers of central vertices. For each vertex i, one finds its distance from the centre $r_i=\min d_{ij}$ (j N; $r_i=0$ for the central vertices). The set of such numbers r_i constitutes the IS. Next the symmetrical function

$$F=\sum_{i=1}^{n} x_i^2,$$

defined by this IS, is calculated. As a result, one obtains the topological index B_c. The calculation of the index B reduces the calculation, for this IS, of the number of elements equal to $a=0,1,2,\ldots$.

The invariants of the graph, associated with the distances matrix D, have also been discussed in two other studies.[26,27]

We may note that the TI defined above refer to a simple molecular graph. In order to take into account heteroatoms and multiple bonds, the distance matrix D, constructed taking into account the weighting factors attributed to the vertices and edges of the MG, was introduced.[28] The elements of the matrix D were determined from the formulae

$$d_{ii}=1-\frac{6}{z_i}\ \text{and}\ d_{ij}=\sum\frac{1}{b_{kl}}\frac{36}{z_k\cdot z_l},$$

where z_j is the nuclear charge of the jth atom and b_{kl} the multiplicity of the bond (k,l); $b_{kl}=1$, 2, 3 or 3/2 depending on whether the bond is respectively single, double, triple, or aromatic. The summation is extended over all the bonds (k,l) forming the shortest path between the ith and jth vertices. The question of certain modifications of the distance matrix D in the presence of multiple bonds is also discussed in the above study.[29] In particular, it is suggested that a weighting factor equal to the bond length be attributed to each edge of the corresponding molecular graph.

3. Indices Based on the Invariant Set of Degrees of Vertices

Here we shall discuss certain TI which are defined in terms of the degrees of the MG vertices. Indices of this type can be easily calculated both for simple and weighted graphs. For this reason, they belong to the class of the most popular TI in theoretical chemistry. We shall indicate initially a series of indices (M_1,[30] P,[31] I,[14] and Q[24]) which are defined with the aid of symmetrical functions of degrees of the graph vertices of the type

$$\sum_i x_i^2,\ \prod_i x_i,\ \sum_i x_i$$

$$M_1=\sum_{i=1}^{n} v_i^2,\ P=\prod_{i=1}^{n} v_i,\ I=\sum_{i=1}^{n} v_i=2q,$$

$$Q=0.5\sum_{i=1}^{n}(v_i^2-2v_i)+1=0.5M_1-2q+1.$$

†In the formula for $J(G)$ and henceforth, the prime before the summation symbol implies that the summation is carried out with respect to all the edges (i,j).

We shall now assume that the edges have been selected as the subgraphs of the graph G in order to determine the indices. Using various symmetrical functions of two variables $F(x_1, x_2)$, it is possible to determine a series of TI which depend on the degrees of the vertices of the graph G. For example, the indices f, M_2, and χ have been introduced:[30,32,33]

$$f = {}'\sum(v_i + v_j - 2),$$

$$M_2 = {}'\sum v_i v_j,$$

$$\chi = {}'\sum(v_i v_j)^{-1/2}.$$

The quantity χ is called the molecular connectivity index or the Randić index,[33] M_2 is the Gutman index,[30] and f is the Platt index.[32] An extension of the structural parameter χ has been proposed;[34] the resulting index is called the index of molecular connectivity of order h. This graph invariant ${}^h\chi$ is determined from the formula

$${}^h\chi = \sum(v_{i_1}, v_{i_2}, \ldots, v_{i_{h+1}})^{-1/2} \text{ (along paths of length } h),$$

where $h \geq 1$ is a fixed integer and v_{i_k} is the degree of the vertex i_k along the given path of length h ($k = 1, \ldots, h$). Evidently, ${}^1\chi = \chi$.

The index ${}^h\chi$ was further modified[35,36] by differentiating in the graph G tree-like subgraphs G_j of a special type, each containing h edges. A certain number δ_i (for example, $\delta_i = \nu_i$) is ascribed to each vertex i of the graph G. Then the quantity

$$F(\delta_{j_1}, \ldots, \delta_{j_{h+1}}) = \prod_{k=1}^{h+1} \delta_{j_k}^{-1/2},$$

is calculated for each subgraph G_j with the vertices $j_1, \ldots, j_{h+1}$. The numbers $F(\delta_{j_1}, \ldots, \delta_{j_{h+1}})$ are then summed over all the subgraphs G_j. In order to take into account multiple bonds and heteroatoms, it has been suggested[36,37] that the quantities

$$\delta_i = z_i^v - r_i \quad \text{or} \quad \delta_i = (z_i^v - r_i)/(z_i - z_i^v - 1),$$

where z_i and z_i^v are respectively the total number of electrons and the number of valence electrons in the ith atom and r_i is the number of hydrogen atoms linked to the ith atom, be used as the numbers δ_i. The connectivity indices obtained in this way may be called the valence connectivity indices and are denoted by ${}^h\chi^v$.

However, cases are possible where atoms of one kind, forming bonds of different types with their neighbours, cannot be distinguished with the aid of the number δ_i. For this reason, a "perturbation" connectivity index has been introduced.[9] Instead of δ_i, in each construction use was made of the numbers δ_i^p, which are determined from the equation $\Gamma\,\overline{\delta^p} = \overline{\delta}$, where $\overline{\delta^p}$ and $\overline{\delta}$ are column vectors with the components δ_i^p and δ_i ($i = 1, \ldots, n$), respectively and $\Gamma = (\gamma_{kl})$, where $kl = 1, \ldots, n$, is a square matrix. The matrix elements γ_{kl} satisfy the conditions $\gamma_{kk} = 1$ and $\gamma_{kl} = 0$, provided that the vertices k and l are not connected. In the remaining cases, the γ_{kl} are selected empirically. Examples of correlations of different physicochemical properties with the "perturbation" connectivity index for a special selection of the elements of the matrix Γ (for example, for $\gamma_{kl} = \gamma_{lk} = 0,1$) have been published[9] and it has been noted that the quality of the correlations is higher than in the case where "non-perturbation" quantities δ_i are used.

The index χ has also been modified in a study[38] whose authors introduced the concept of the lth order connectivity K_l, where $l = 2,3$, using the following formula to find this quantity:

$$K_l = {}'\sum((v_i a_i v_j a_j) b_{ij})^{-1/2},$$

where ${}^l\nu_i$ is the number of second-order ($l = 2$) or third-order ($l = 3$) neighbours of the vertex i, the number a_i characterises the type of the ith atom, and b_{ij} is the (i,j) bond order. The indices ${}^1\chi$, ${}^2\chi$, f, and M_1, determined for weighted graphs, have been discussed.[29]

In the case of weighted graphs, the problem of the choice of the weighting factors for the vertices and edges arises. Some of the ways of introducing the weighting factors were discussed above. One of the possible variants of the distribution of the weighting factors among the vertices of the graph, associated with the method of potentials,[40] has been proposed.[39] A number u_i, called the potential defined by the formula

$$u_i = \sum_{j=1}^{n} h_{ij} h_{jj},$$

where $i = 1, \ldots, n$ and the h_{ij} are the elements of a non-degenerate matrix $H = (h_{ij})$, which depend in a special way on the elements of the adjoint matrix, is attributed to each vertex. The topological index is defined by the expression

$$\prod_{i=1}^{n} u_i \text{ [39].}$$

Certain TI which depend on the degrees of the vertices, can be represented as functions of the TI associated with more complex fragments. For example, $f = 2N_2$, where f is the Platt index and N_2 the number of paths on the MG whose length is 2 (the quantity N_2 has been introduced as the TI in the study of Gordon and Scantlebury[41]). On the other hand, if the degrees of the graph vertices do not exceed 4, then the index N_2 is linearly related to the index Q and the index M_1: $Q = N_2 - n + 2 = 3 - 2n + 0.5M_1$. For acyclic graphs,

$$\sum(v_i - 1)v_i \text{ [16].}$$

The index N_2 has been used to construct a new topological index ${}^2\kappa$ on the basis of the formula

$${}^2\kappa = \frac{2N_{2,\max} N_{2,\min}}{N_2^2},$$

where $N_{2,\max} = 0.5(n-1)(n-2)$ and $N_{2,\min} = n - 2$ are the values of N_2 for a simple chain and a star respectively and n is the number of vertices of the graph.[42]

The MG fragments corresponding to paths whose lengths are greater than 2 have also found applications in the construction of TI.[8,43-45] For example, it has been suggested that the sequence of numbers $N_1, \ldots, N_k$, where N_i is the number of paths of lengths i on the MG, be used to code chemical structures.[43-45] The "distance" d between the structures can be determined in terms of the numbers N_i by means of one of the formulae

$$d^2 = \sum_i (N_i - N_i')^2 \text{ and } d = \sum_i (N_i - N_i') \log_2 \frac{N_i}{N_i'},$$

which makes it possible to classify these structures in terms of the extent to which they approach a certain standard structure. In many cases the sequence of structures obtained in this way is consistent with the character of the changes in the properties.[43,45]

4. Indices which Depend on the Spectroscopic Characteristics of the Graph

A series of TI used in the analysis of molecular structures are related directly to the eingenvalues λ_j, where $j = 1, \ldots, n$, of the adjacency matrix A of the molecular graph.

The numbers $\lambda_1, \ldots, \lambda_n$ form the spectrum of this graph and are identical with the set of zeros of the characteristic polynomial:[23]

$$R(G, x) = (-1)^n \det(A - xE) = x^n + a_1 x^{n-1} + \ldots + a_n .$$

The spectrum of the graph is an IS and the coefficients of the polynomial $R(G,x)$ are symmetrical functions of the eigenvalues. Therefore any coefficient of the polynomial $R(G,x)$ can be used as the TI. In particular, the sum of the absolute coefficients of the polynomial

$$\sum_i |a_i| .$$

is also a TI. For acyclic graphs, this quantity is identical with the Hosoya index Z:[46]

$$Z = \sum_k p(G,k) ,$$

where $p(G,k)$ is the number of ways of selecting k edges of the graph in such a way that no two of them have common vertices. The Hosoya indices used for the estimation of the physicochemical properties of hydrocarbons have been tabulated.[47,48]

Individual eigenvalues, the sums of non-negative eigenvalues, etc. are also used as the TI.[44,49,50] For example, the maximum eigenvalue is used as the characteristic of the degree of branching of the molecule.[49] The relation between the coefficient of the polynomial $R(G,x)$ and other invariants of the MG has been discussed.[23] The spectroscopic characteristics of MG have found applications in the theory of the electronic structures of conjugated systems. These questions are considered in greater detail in Section IV.

5. Information Topological Indices

One of the ways of comparing molecular structures in terms of the topological approximation consists in the calculation for each of them the number of structural fragments of a specific type and subsequent comparison of these quantities. Since the structural fragments of one type can differ in terms of some feature, they can be classified further with different degrees of detail. We shall assume that a certain criterion has been selected for the classification, which leads to the subdivision of the set of the chosen structural fragments X, consisting of n elements, into k equivalence classes X_i, and

$$n = \sum_i n_i ,$$

where n_i is the number of elements in the set X_i. Definite information about the molecular structure, which can be characterised quantitatively, is associated with each selection of the numbers $(n_i, \ldots, n_k)$. The quantity n_i/n can be regarded as the probability of finding the fragment from the class X_i in the structural formula of the molecule considered. Therefore, for a quantitative assessment of the information about the molecular structure contained in the set of numbers $(n_1/n, \ldots, n_k/n)$, use is made of the entropy of the distribution of the probabilities $\bar{I}$ and the associated quantity $I = n\bar{I}$, which are calculated from the formulae

$$\bar{I} = -\sum_i \frac{n_i}{n} \log_2 \frac{n_i}{n} , \quad I = \bar{I}n = n \log_2 n - \sum_i n_i \log_2 n_i .$$

The quantities $\bar{I} \geq 0$ and $I \geq 0$ can be used to estimate the degree of inhomogeneity of the set of the selected structural fragments of the molecule. This method of interpretation is based on the following properties of the entropy: for the most inhomogeneous subdivision (i.e. when $n_i = 1/n$, $i = 1, \ldots, n$), the quantity $\bar{I}$ assumes its maximum value, which is $\log_2 n$, while for the most homogeneous subdivision (i.e. when one of the numbers n_i is 1 and all the others become zero), $\bar{I}$ attains its minimum value, which is zero. The function $\bar{I}$ is sometimes referred to as the average information content of the given subdivision and I is the information content of the subdivision. The use of such terminology is associated with the fact that in information theory formulae of a similar type are used to characterise the amount of information contained in a message of a special type. The specification of the number n_i/n can be interpreted as a message about the composition of the set X. Furthermore, the formula for the index $\bar{I}$ is analogous to the expression for the thermodynamic entropy.[51]

At the present time, a series of TI, referred to as the information TI, are used in structural chemistry. All these indices are constructed in accordance with the principle described above: initially a certain set X, consisting of the structural fragments of the molecular graph, is selected and then it is subdivided into equivalent classes X_i and the quantities I and $\bar{I}$, which depend on the chosen set of structural fragments and the criterion constituting the basis of the subdivision of the set X into classes, are calculated. We may note that the methods used to construct quantities of the type I and $\bar{I}$ fit within the general scheme for the construction of TI described in subsection 1. In this case, the set $= (n_1/n, \ldots, n_k/n)$ is an IS and the symmetrical function F has the form

$$F(x_1, \ldots, x_k) = -\sum_i x_i \log_2 x_i .$$

We shall deal brirefly with certain TI of the information type used in the analysis of molecular structures.

(a) Allowance for the atomic composition in the wide variety of types of chemical bonds. The information indices were first used in theoretical chemistry in a study[52] where the assembly of all the atoms of the molecule was considered as a set X and the assembly of atoms of a single type i were treated as the subsets X_i. The degree of inhomogeneity of the elemental composition was characterised by the information-theoretic index I_{at}. Subsequently the index I_{at} was modified;[53] apart from the elemental composition, account was taken also of the valence of the atoms. The inhomogeneity of the molecular structure was then estimated with the aid of the Morowitz formula

$$I_{MOR} = I_{at} + \sum_i n_i \log_2 t_i ,$$

where the numbers t_i are calculated from the valences of atoms of the ith kind. The index I_B, characterising the variety of chemical bonds, was determined analogously:[54]

$$I_B = B \log_2 B - \sum_i B_i \log_2 B_i ,$$

where B_i is the number of bonds of type i (for example, single, double, and triple) and B is the total number of bonds.

(b) Allowance for the symmetry of the molecular graph. It has been suggested[55-57] that the vertices of the graph be subdivided into equivalence classes in such a way that the vertices of different classes cannot be interconverted whatever the automorphism of the graph. For example, it is possible to combine in a single class vertices having the same degrees, the same number of second neighbours, etc. The "finest" subdivision is generated by the orbits of the group of automorphisms of the graph. The information index I_{orb} corresponds to such subdivision into classes. In order to characterise the molecular structure, the sum of the automatic and orbital indices has been introduced:[55]

$$\bar{I}_{\text{RASH}} = \bar{I}_{\text{at}} + \bar{I}_{\text{orb}}.$$

(c) Chromatic information index. The graph G is said to be k-colourable if its vertices can be coloured by k colours in such a way that each vertex has neighbours of another colour only.[11] The vertices of a graph coloured in this way can be divided into non-intersecting classes. The elements of each class have the same colour and are not connected. The smallest number $k = k(G)$ for which k-colouring of the graph is possible is called the chromatic number and the corresponding subdivision into classes is called the chromatic resolution of the vertices.

Suppose that V is the chromatic resolution of the vertices of the graph and $n_i(V)$ the number of vertices in the ith class $[i = 1, \ldots, k(G)]$. The chromatic information index has been introduced:[58]

$$\bar{I}_{\text{chr}} = \min_{V} \left\{ -\sum_{i=1}^{k(G)} \frac{n_i(V)}{n} \log_2 \frac{n_i(V)}{n} \right\}$$

and $\bar{I}_{\text{chr}}$ and $\bar{I}_{\text{orb}}$ have been compared. In particular, it has been shown that, for monocyclic structures of any size, $\bar{I}_{\text{orb}} = 0$, since all the vertices belong to the same orbit of the group of automorphisms of the graph, but $\bar{I}_{\text{chr}} \neq 0$ for such graphs. However, for certain molecular graphs, $\bar{I}_{\text{chr}} = \bar{I}_{\text{orb}}$. The chromatic index can be determined analogously also for the edges. Chromatic information indices have been used in the analysis of the properties of alternant hydrocarbons. The conditions for the existence of the Kekulé structure of aromatic hydrocarbons have been formulated in terms of the chromatic indices.[14]

(d) The matrices of the graph and information indices. The set of the elements of the adjoint matrix A of order n, whose elements are zero or one, can be subdivided into two classes consisting respectively of zeros and ones. If n_1 and n_2 are the numbers of elements in these classes and $n^2 = n_1 + n_2$, then the quantity

$$\bar{I}_A = -\sum_{i=1}^{2} \frac{n_i}{n^2} \log_2 \frac{n_i}{n^2}$$

characterises the inhomogeneity of the composition of the set of the elements of the matrix A. The information indices can be constructed analogously for the edge adjacency matrix, the cycle matrix, the incidence matrix, etc.[14,59] Information indices have been used in a number of studies to analyse the structure of IS consisting of the degrees of the vertices $\{v_i\}$ or the distance degrees of the vertices $\{d_i\}$. Any two elements having equal values were assigned to the same class. The information indices corresponding to these subdivisions were denoted by ${}^{V}I_A^E$ and ${}^{V}I_D^E$ respectively. The information index I_{conn},[59] defined on the bases of the edge adjacency matrix, can be constructed analogously: the non-zero elements of the matrix are subdivided, for example, into the orbits of the group of automorphisms of the matrix. If $n_{c,i}$ is the number of elements in the ith orbit and n_c the total number of non-zero elements in the matrix, then

$$\bar{I}_{conn} = -\sum_{i} \frac{n_{c,i}}{n_c} \log_2 \frac{n_{c,i}}{n_c} .$$

(e) The use of compound topological indices for the construction of TI of the information type. Some of the topological indices considered above are represented by the sum of local or integral TI. Such indices can be conveniently referred to as compound. For example, the Wiener (W), Hosoya (Z), and the Randić (χ) indices as well as the Balaban centric index B are compound. New indices of the information type have been derived in a series of studies on the basis of these TI. In their construction, the set of components whose sum yields the initial TI was subdivided into classes in some way. For example, the Wiener index W, which can be written in two equivalent forms, i.e.

$$W = \sum_i i g_i = \frac{1}{2} \sum_{i,j} d_{ij},$$

has been used to construct the indices I_D^W, $\bar{I}_D^W$, I_D^E, $\bar{I}_D^E$, I_D^M, and $\bar{I}_D^M$, which are defined by the following formulae:[14,16,19]

$$I_D^W = W \log_2 W - \sum_j g_j j \log_2 j; \quad \bar{I}_D^W = -\sum_j g_j \frac{j}{W} \log_2 \frac{j}{W};$$

$$I_D^E = \frac{n(n-1)}{2} \log_2 \frac{n(n-1)}{2} - \sum_i g_i \log_2 i;$$

$$\bar{I}_D^E = -\sum_j \frac{2g_j}{n(n-1)} \log_2 \frac{2g_j}{n(n-1)};$$

$$I_D^M = 2W \log_2 2W - \sum_i d_i \log_2 d_i;$$

$$\bar{I}_D^M = -\sum_i \frac{d_i}{2W} \log_2 \frac{d_i}{2W} .$$

In these formulae, g_j is the number of pairs of MG vertices at a distance j and d_i the distance degree of the ith vertex. Analogous expressions have been constructed also for the indices B, Z, and χ, for example:[16]

$$\bar{I}^Z = -\sum_k \frac{p(G,k)}{Z} \log_2 \frac{p(G,k)}{Z} .$$

(f) Information indices and long range effects. As stated in subsection 3 of this Section, TI which depend on the structure of different coordination spheres of the MG vertices (for example, the indices K_l for $l = 2,3$) are used in theoretical chemistry. In this subsection we shall consider a series of information-theoretic indices of analogous types. The definition of these TI is based on the subdivision of the set of graph vertices into equivalent classes, in which the character of the immediate environment of the vertices, including their degree and the multiplicity of the corresponding edges, is taken into account in detail.

Two vertices i and j belong to the same class if, firstly, they correspond to atoms of the same type and have equal degrees; secondly, each of these vertices has the same number of first neighbours $i_1, \ldots, i_m$ and $j_1, \ldots, j_m$; thirdly, there exists a mutually unambiguous mapping of the set of vertices $i_1, \ldots, i_m$ on the set of vertices $j_1, \ldots, j_m$, $i_\alpha \leftrightarrow j_\beta$, for which the vertices i_α and i_β correspond to atoms of the same type; fourthly, the orders of the vertices i_α and j_β are identical, i.e. $\nu_{i_\alpha} = \nu_{j_\beta}$; fifthly, the multiplicities of the edges (i, i_α) and (j, j_β) are identical. Only TI of information-theoretic type have been compared with this type of sub-division of the vertices of the molecular graph into equivalence classes:[60-72] ${}^1IC$ (Information Content), ${}^1SIC$ (Structural Information Content), ${}^1CIC$ (Complementary Information Content), and ${}^1BIC$ (Bonding Information Content), which are defined by the formulae

$${}^1IC = -\sum_i \frac{n_i}{n} \log_2 \frac{n_i}{n}; \quad {}^1SIC = \frac{{}^1IC}{\log_2 n};$$

$${}^1CIC = \log_2 n - {}^1IC; \quad {}^1BIC = {}^1IC/\log_2 q .$$

In these formulae, n is the number of graph vertices, n_i the number of vertices in the ith equivalence class, and q the number of graph edges. The index ${}^1CIC$ characterises the deviation of the topological index ${}^1IC$ from its maximum value ${}^1IC_{\max} = \log_2 n$. The differences between other coordination

spheres can be taken into account analogously, which leads to the indices ${}^{k}IC$, ${}^{k}CIC$, ${}^{k}SIC$, and ${}^{k}BIC$;[64] the symbol k characterises the order of the coordination sphere.

Together with the information topological indices ${}^{k}IC$, etc. mentioned above, use is also made of "weighted" information indices[73] based on the subdivision of the graph vertices into the equivalence classes described above. However, in this case a weighting factor equal to the van der Waals volume of the corresponding atom is attributed to each vertex and then the weighting factors of the vertices referring to the same class are added together and the index

$$ {}^{k}IC_v = -\sum_{i=1}^{j} \frac{v_{is}}{v_s} \log_2 \frac{v_{is}}{v_s} \text{ and } v_s = \sum_{i=1}^{j} v_{is} ,$$

where v_{is} is the sum of the weighting factors of the vertices of the ith class and j is the number of classes, is calculated. The topological indices ${}^{k}SIC_v$, ${}^{k}BIC_v$, and ${}^{k}CIC_v$ are determined analogously.

We may indicate yet another information index I_{sym}, which is used to characterise the molecular structure but is not topological in the generally accepted sense. The definition is based on allowance for the molecular symmetry and has been proposed in a study[74] where the following criterion of the equivalence of atoms was introduced: the atoms are assumed to be equivalent if they are interconverted by some kind of operation involving the symmetry point group of the molecule.

6. Other Topological Indices

The adjoint matrix of the graph $A = (a_{ij})$ can be used to construct certain other matrices, which are useful in finding new IS and hence also other TI. For example, in the method of potentials, which is used to establish the isomorphism of graphs,[40] the matrices $L = (l_{ij})$ and $H(h_{ij})$, where $l_{ij} = -a_{ij}, (i \neq j)$, $l_{ii} = \nu_i + 1$, and $H = L^{-1}$, are defined. The matrix $D^H = (d_{ij}^H)$ with the elements $d_{ij}^H = 0.5(h_{ii} + h_{jj}) - h_{ij}$ has been introduced.[39] The matrices L, H, and D^H have been used to define new topological indices:

$$\prod_{i \neq j} d_{ij}^H, \det L, \det H.$$

We may mention several further graph invariants, which are discussed in the mathematical literature[11,23] but which have not so far found application in theoretical chemistry as TI: the numbers of the vertex and edge coverings, the vertex and edge independence numbers, thickness, edge connectivity, deformability, and certain others. In conclusion we may note one other TI for simple graphs, which is known as the minimum binary code.[33] For its construction, each adjoint matrix of the graph is formulated and a single row of length n^2 made up of zeros and ones. Each such row can be regarded as the binary formulation of a certain number. The smallest of all these numbers, corresponding to the adjoint matrices, is in fact referred to as the minimum binary code. We may note that the calculation of such a code for large matrices is fairly laborious.

7. Relation between the Topological Indices and Their Comparison

As shown above, hitherto numerous TI of different kinds have been defined. Most of them have found application in the analysis of molecular structure and in the construction of correlations between structure and property. Some of these TI are linked by simple relations. The examples of such functional relations have been discussed above. Undiscovered relations of some kind probably exist also in many other cases. The results of testing the approximate equalities of the type $T_i = a + bT_j$ for certain pairs of topological indices T_i and T_j have been reported.[75] The constants a and b were fitted by the method of least squares. Such analysis has been carried out for TI of three series of hydrocarbons—alkanes, polyalkylbenzenes, and monocyclic structures. It has been shown[75] that approximate linear relations with a fairly high correlation coefficient r ($0.84 \leq r \leq 0.99$) exist between the indices W, Z, χ, N_2, λ_1, and M_1. On the basis of such an investigation, one can postulate that the indices of one group carry approximately the same information about the structure of compounds of the given classes.

Certain indices, for example, I_D^W and I^Z, can be fitted by second degree polynomials in terms of Hosoya (Z) and Wiener (W) indices as well as other TI.[76]

We shall now proceed to the estimation of the quality of the TI. In coding chemical structures with the aid of TI, it is desirable that different structures should receive different index codes. However, one cannot establish a mutually unambiguous correspondence between the set of square symmetrical matrices, describing the topology of the molecule, and the set of real numbers. It is therefore always possible to select two different structures having the same values of the TI. Consequently, any TI is degenerate with respect to structures of some kind. Topological indices are usually compared in terms of their capacity to "differentiate" compounds of a certain class, i.e. in terms of the degree of their degeneracy. TI are usually employed to differentiate isomeric compounds. For example, the Balaban index $J(G)$[13] is subject to a comparatively small degree of degeneracy. Such degeneracy appears for alkane isomers only in molecules where the number of carbon atoms is greater than 20. In order to obtain TI with a smaller degree of degeneracy, superindices, equal to the sum of simpler TI, or multi-indices, which constitute an ordered sequence of topological indices, are used.[14,47] The question of the degeneracy of different TI has also been discussed in a number of publications.[13,16,77–80]

The effectiveness of 19 TI, used to identify chemical structures, has been investigated.[81] The quantity $e = n/N - t/T$, where n is the number of different values of the index in an assembly of N structures and t and T are the average time of the calculation of the index and the time required to test the isomorphism criterion respectively, has been used as a characteristic of the effectiveness of the TI. Indices defined in the method of potentials[39] and the Randić index[33] have been recognised as the most effective. One TI is regarded as "better" than another (in the estimation of the given physical property) if the correlation equation obtained with its aid is more accurate.

8. Topological Indices as Characteristics of the Shape and Size of Molecules

It has been suggested that certain TI be used to characterise the degree of branching and the shape of molecules.[19] According to intuitive ideas about the degree of branching for a specified number of vertices, the chain and star graphs have the least and most branched structures respectively. In terms of the degree of branching, the remaining structures occupy an intermediate position between them. Certain indices considered above, for example W, χ, B'', λ_1, and ${}^{2}\kappa$ have the property involving their assumption of the smallest (or greatest) value for the chain and of the greatest (or smallest) value for the star.# It has therefore been

#See also L.B.Kier, Quant.Struct.-Act.Relat., 1987, 6, 8.

suggested that they be regarded as measures of the branching of the structure. Another piece of evidence in support of the selection of these branching characteristics is the existence for a number of TI of relations of the type

$$T_j = a + bn + cn_1 + dn_2 \text{ [75]},$$

where n_1 is the number of branching points, n_2 the number of side chains connected to the main chain, and a, b, c, and d are constants.

Topological indices also vary in parallel with molecular size. It has been noted that,[82] for a series of compounds, the van der Waals volume of the molecule varies linearly with molecular refraction (MR) and determines the value of the latter to the extent of 90%. Correlations have been found[82] between MR and TI, i.e. relations of the type $R = a + bT_j$, for 17 TI and it has been concluded that the van der Waals volume can be estimated fairly accurately with the aid of the indices W and M_2 and the two information indices I_D^E and I_D^W.[83]

III. THE USE OF TOPOLOGICAL INDICES FOR THE ESTIMATION OF PHYSICOCHEMICAL PROPERTIES OF COMPOUNDS

The traditional methods for the estimation of physicochemical properties of compounds are based on additive schemes. In the classical theory of chemical structures, it is postulated that certain physicochemical properties of organic compounds (for example, the heat of formation, the boiling point, the polarisability, the dipole moment, etc.) can be calculated as the sum of the contributions ε_i, ε_{ij}, and ε'_{kl}, corresponding to the effective atoms i, the bonds (i, j), and pairs of valence- non-bonded atoms (k, l) of the corresponding molecules.[4] In the simplest case, this approach presupposes a fairly detailed classification of the vertices and edges of the MG and is based on the hypothesis that the contributions ε_x for fragments x of one type are independent of the choice of the molecule of the class under consideration.

This additive scheme for the calculation of molecular properties permits a series of modifications. Thus, one of the versions of additive schemes reduces to the representation of the numerical value y of the given property of a chemical compound by the linear function

$$y = \sum_i a_i k_i,$$

defined on the basis of the corresponding MG. In this formula, the index i characterises a MG fragment of a specific type, which can have a fairly complex structure, k_i is the topological index equal to the number of such fragments, and a_i is the contribution by each fragment of the ith class to the given property; the numbers a_i are independent of the choice of compounds. As an example, we may point out that the heats of formation of alkanes ΔH^0_{298} can be calculated approximately from the formula

$$\Delta H^0_{298}, \text{ (kcal mol}^{-1}) = 16.625 + 4.158N_1 + 1.074N_2 - 0.431N_3,$$

where N_i is the number of paths of lengths i on the MG corresponding to the carbon skeleton of the molecule.[8]

We shall now discuss the results of other studies in which different modifications of the simple additive model have been realised. In most of them, the physicochemical properties of alkanes are discussed. We may note that certain properties of hydrocarbons are satisfactorily correlated with the number of carbon atoms.[84-89] Thus the heats of atomisation and molar refractions of normal alkanes vary linearly and the boiling point and density vary non-linearly with the number of carbon atoms. However, the properties of branched hydrocarbons can no longer be described in terms of the single parameter n, since in this case the degree of branching of the molecule also plays a significant role.[10]§ Wiener has suggested in a number of studies[20,90-93] that properties of alkane isomers such as molar refraction, molar volume, the boiling point, the heat of formation, etc. be described with the aid of equations of the type

$$y - y_0 = a(W - W_0) + b(p - p_0),$$

where y is the numerical value of the given property, W the Wiener index for the MG corresponding to the carbon skeleton of the molecule, p, g_3 the so called polarity number, y_0, w_0, and p_0 are the corresponding characteristics for a normal alkane, and a and b are constants which depend on the property investigated. The value of the index W is interpreted as a quantity characterising the degree of branching of the isomer. For example, the following formula has been obtained[20] for the boiling points of alkanes t:

$$t - t_0, ^\circ\text{C} = \frac{98(W - W_0)}{n^2} + 5.5(p - p_0),$$

where

$$t_0 = 745.42 \log(n + 4.4) - 689.4; \; W_0 = \frac{1}{6}(n^3 - n); \; p_0 = n - 3,.$$

The correlations of the type described above have been analysed,[32] which has led to a conclusion concerning the cases where such indices can be usefully employed. Thus it has been shown that the heat of formation and molar refraction are best described in terms of the invariants f and p (f is the Platt index) and the heat of evaporation, the boiling point, and the molar volume can be most usefully described in terms of the indices p and W. The heats of formation of alkanes are most accurately described by a three-parameter correlation with the indices p, W, and f. Subsequently the Wiener index was used to estimate the properties of other hydrocarbons also. Relations between the Wiener index, on the one hand, and the boiling and melting points, surface tensions, viscosities, densities, and refractions of alkanes, alkynes and alkenes, as well as the boiling points of arenes, on the other have been published.[18,21] The index W has found applications also in other branches of physical chemistry: for the estimation of the width of the forbidden gap in the π-electronic spectra of conjugated micromolecules,[94,95,96] in chromatography,[94,97,98] and in crystal theory.[99,100]

In a number of studies, the Hosoya index has been used to estimate the thermochemical properties of hydrocarbons of this type.[46-48] A linear correlation between the boiling points of alkanes and both the index Z and the index

$$P = \prod_i v_i \text{ [31]}.$$

has been reported.[31] The index Z has also been used to estimate the bond orders and π-electron energies in conjugated hydrocarbons.[101,102]

It has been shown that the octane numbers of acyclic hydrocarbons can be estimated fairly satisfactorily in terms of the centric indices B, B', and B'' (especially the binormalised index B'').[103] The introduction of an additional parameter—the number of carbon atoms or the molecular weight of alkanes—increases the accuracy of the estimate.

Among other TI used to estimate the physicochemical properties of compounds, the connectivity index $^1\chi$ and its varieties, especially $^2\chi$, $^1\chi^v$, and $^2\chi^v$, have enjoyed the greatest popularity. Relations have been obtained[9,10,33,104,105]

---

§ See also J.S.Chickos, R.Annuziata, L.H.Ladon, et al., J.Org.Chem., 1986, 51, 4311.

between these quantities and various physicochemical properties. We shall quote several examples of such equations. It has been shown [33] that the heat of formation of alkanes ΔH_f^0 can be described approximately by the formula

$$\Delta H_f^0 \text{ (kcal mol}^{-1}) = 12\chi - 11n - 9,$$

where n is the number of carbon atoms in the molecule. An equation relating the heats of atomisation of monohydric alcohols ΔH_a to the connectivity index χ and the number of non-hydrogen atoms n has been published:[10]

$$\Delta H_a \text{ (kcal mol}^{-1}) = -21.926 + 22.275\Delta\chi + 559.975\chi_n,$$
$$(r > 0.9999;\ s = 1.01;\ N = 19)$$

where $\Delta\chi = \chi - \chi_n$ and $\chi_n = 0.5(n-3) + 2^{1/2}$ is the connectivity index for a normal alcohol. It has been noted that the separation of the contributions by the group OH and the carbon skeleton to χ, i.e. the use of two independent parameters χ^{OH} and χ^{C} instead of the sum χ, reduces the standard error s by 30–40%. The density d of carboxylic acids can be estimated in terms of the parameters $1/\chi$ and $^3\chi$ by means of the formula[10]

$$d(\text{g cm}^{-3}) = 0.7546 + 0.4358\chi^{-1} + 0.0252^3\chi \quad (r = 0.983;\ s = 0.014;\ N = 20).$$

Table 1. Correlations between physicochemical properties and connectivity indices.[9]

Class of compounds	N	r	Refs.
Solubility in water			
Alcohols	51	0.978	[110]
Ditto	48	0.988	[37]
Ditto	51	0.965	[111]
Alkanes	18	0.958	[110]
Alkylbenzenes	13	0.989	[10]
Esters	38	0.990	[35]
Ethers	22	0.989	[10]
Boiling point			
Alcohols	63	0.970	[110]
Ditto	48	0.993	[37]
Ditto	28	0.999	[10]
Alkanes	51	0.997	[10]
Primary amines	12	0.999	[10]
Secondary amines	13	0.999	[10]
Tertiary amines	4	0.997	[10]
Alkyl halides	24	0.992	[37]
Ethers	11	0.988	[10]
Polyols	9	0.995	[37]
Carboxylic acids	22	0.996	[112]
Surface area of the molecule			
Alkanes and alcohols	69	0.978	[34]
Density			
Liquid aliphatic hydrocarbons	82	0.989	[113]
Hydrocarbons	13	0.910	[10]
Aliphatic alcohols	42	0.931	[113]
Aliphatic acids	20	0.983	[113]
Aliphatic ethers	13	0.910	[10]
Diamagnetic susceptibility			
Hydrocarbons	27	0.999	[10]
Alcohols	15	0.999	[10]
Excluded volume			
Binary mixture of n-alkanes	—	—	[114]
Heat of atomisation			
Alkanes	44	0.999	[10]
Alcohols	19	0.999	[10]
Ethers	12	0.999	[10]
Olefins	37	0.999	[10]
Thiols	14	0.999	[10]
Heat of formation			
Alkanes	67	0.999	[10]
Alcohols	20	0.998	[10]
Heat of evaporation			
Alkanes	44	0.999	[10]
Alcohols	20	0.998	[10]

Table 1 continued.

Class of compounds	N	r	Refs.
Molar refraction			
Alcohols	31	0.998	[10]
Alkanes	46	0.999	[10]
Amines	22	0.988	[10]
Alkylbenzenes	70	0.999	[10]
Ethers	9	0.999	[10]
Mixed group of compounds	65	0.990	[10]
Olefins	39	0.999	[10]
Phenyl-substituted compounds	25	0.990	[37]
Molar volume			
n-Alkanes at 20 °C	14	0.999	[114]
n-Alkanes at 50 °C	13	0.999	[114]
n-Alkanes at 100 °C	15	0.999	[114]
Polarisability			
Mixed group of compounds	36	0.990	[34]
Partition coefficient between water and 1-octanol			
Alcohols	49	0.997	[115]
"	14	0.999	[116]
Amines	12	0.998	[116]
"	28	0.979	[115]
Carboxylic acids	9	0.996	[115]
Hydrocarbons	45	0.975	[115]
Esters	24	0.999	[115]
Ethers	12	0.976	[115]
Ketones	16	0.993	[115]
Phenols	16	0.988	[116]
Mixed group of compounds	66	0.970	[116]
Ditto	138	0.986	[115]
Gas chromatographic retention indices on squalane			
Alcohols	—	0.994	[118]
Esters	—	0.991	[118]
Ethers	—	0.968	[118]
Ketones	—	0.996	[118]
Van der Waals constant b			
Mixed group of compounds	20	0.981	[37]
Positronium annihilation cross-sectional area			
Benzene and its derivatives	17	0.997	[117]
Alkanes	47	0.999	[117]
Halogeno-compounds	31	0.998	[117]
Oxygen-containing compounds	46	0.998	[117]
Unsaturated compounds	33	0.984	[117]
Mixed group of compounds	182	0.996	[117]

Notation. N is the number of compounds and r the correlation coefficient.

The topological index χ proved to be useful in chromatography for the estimation of retention indices.[104,105] However, in this case, together with χ, it is necessary to use also other topological characteristics of the molecule, describing its size, for example, the number of paths of length 3 or the number of six-membered rings in aromatic hydrocarbons.[104]

We shall now proceed to information TI. It has been shown that the Morowitz index I_{MOR} is correlated with the entropies and heats of formation of organic compounds of different classes.[14,106] Equations of the type

$$y = a + bn + cI_D^W + dI_D^W n,$$

where a, b, c, and d are constants, have been used to estimate the molecular volumes and properties of hydrocarbons such as heats of formation, oxidation, and evaporation and the boiling points.[16,94,107] A relation between the index I_D^W for alkylbenzenes and their chromatographic retention parameters has also been reported. Analogous relations have been obtained likewise for the index I_D^E in the case of o-dialkylbenzenes and isoalkanes.[108] The index I_D^M has found an application in the estimation of different thermochemical properties of alkylbenzenes.[109] There is an equation[14] relating the heat of evaporation ΔL_{298} of alcohols and the index I_{orb}:

$$\Delta L_{298} \text{ (kcal mol}^{-1}) = 0.265 I_{orb} + 8.566 \quad (r = 0.989;\ s = 0.414;\ N = 15).$$

A linear relation has been established between the entropies of alkanes, alkenes, alcohols, and organic acids and the information index I_{sym}.[74] Other examples of the use of information topological indices have been discussed in a monograph.[14]

Table 1 summarises data on some of the relations described in the literature between the connectivity indices and various physicochemical properties.

IV. TOPOLOGICAL INDICES AND BIOLOGICAL ACTIVITY

One of the most important applications of TI involves their use to establish quantitative relations between the structure and biological activity of molecules. Biological activity is usually estimated in terms of the molar concentration of the substance necessary to attain a specified level of biological activity. Various parameters are used as characteristics of biological activity, including LD_{50} (or LC_{50}), representing the dose which destroys 50% of the organisms tested, the dose ED_{50} of the antagonist which lowers by 50% the effect of the standard dose of the agonist, the concentration *MIC* of the inhibitor which suppresses the growth of the experimental biological object to a minimal level, the molar concentration I_{50} of the inhibitor which reduces the rate of the enzyme-catalysed reaction by a factor of two, etc.[12] The hydrophobicity parameter lg *P* (*P* is the partition coefficient of the given compound between water and a lipid-simulating phase, usually 1-octanol) as well as the Hammett and Taft constants, reflecting the electronic and steric structures, are most often used as empirical parameters in the study of the relation between structure and activity. An extensive bibliography on the methods used to establish a relation between structure and biological activity has been published[12,40,119] and the principal molecular characteristics used to solve this problem have been described.

Among the TI used to investigate biological activity, the molecular connectivity indices, especially the Randić index χ, have found the most extensive applications. We shall consider certain examples of the use of indices of this type.

A correlation between lg *MBC* and the index χ has been found by Kier et al.[34] from the results of Agin et al.[120] for a series of anaesthetics (aliphatic alcohols, ketones, ethers, amines, phenols, heterocycles, and hydrocarbons):

$$\lg MBC = 3.55 - 0.762\chi \quad (r = 0.983;\ s = 0.390;\ N = 36).$$

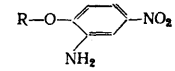
Here *MBC* is the minimum blocking concentration of the anaesthetics for isolated muscle or nerve fibres. We may note that a relation between lg *MBC* and the polarisability α has been discovered for the given group of compounds[120] and a correlation has been found[34] between α and χ. These facts have served as the basis of the above equation relating lg *MBC* and χ.

Certain relations between the index χ and biological activity have been discussed.[115,121,122] The discovery of linear correlations between χ and lg *P* and also between lg *P* and the biological activity of aliphatic alcohols made it possible to establish a linear relation between χ and biological activity.[115] The narcotic effect of alcohols on the larvae of certain crustaceans and the inhibition of the movement of tadpoles have been considered as a manifestation of the biological activity of such substances. For aliphatic alcohols, ethers, esters, ketones, and carbamates, relations have been found between the index χ and the non-specific narcotic activities of the above compounds on tadpoles ($N = 36$).[121] The dependence of the non-specific toxic activity of monofunctional alcohols, ethers, and ketones on fungi of a particular species and the index χ ($N = 45$) was considered in the same study. A linear correlation has also been reported[121] between the index χ for a series of compounds and the inhibition of the enzymes succinate oxidase, thymidine phosphorylase, and butyrylcholinesterase. Thus an equation of the following type has been obtained for a series of saturated and unsaturated aliphatic alcohols, ketones, phenols, aniline, and pyridine, which inhibit succinate oxidase:

$$-\lg I_{50} = 0.916\chi - 1.582 \quad (r = 0.966;\ s = 0.169;\ N = 13).$$

Quadratic correlations between lg I_{50}, lg LC_{50}, and quantities such as lg *P* and χ have also been described. Such correlations have been observed, for example, in estimations of the bactericidal activities of alkylbenzylmethylammonium salts and of the narcotic activities of 5,5-dialkylbarbituric acids on certain animal species.[16,63,122]

Correlations between the connectivity indices and parameter of types of biological activity such as non-specific narcotic activity, enzyme inhibition, bactericidal activity, toxicity, and anaesthesising activity have been examined in a monograph.[10] The possibility of estimating the taste and smell of substances with the aid of topological characteristics of molecular structure has been discussed in the same publication. Thus the relative sweetness *RS* of nitroanilines of the type



is expressed in terms of the parameters $^1\chi$ and $^1\chi^v$ in the following manner:

$$\lg RS = 0.350\,{}^1\chi + 0.694\,{}^1\chi^v - 3.856 \quad (r = 0.953;\ s = 0.222;\ N = 9).$$

According to one of the theories of smell, the smell of a substance is determined by the shape and size of its molecule.[12] It has been suggested[10] that molecular shape be characterised by the parameter χ/n and the hypothesis was put forward that substances having identical smells should have similar values of χ/n. To confirm this hypothesis, a list of structurally different substances having the almond smell and similar values of χ/n is quoted.[10]

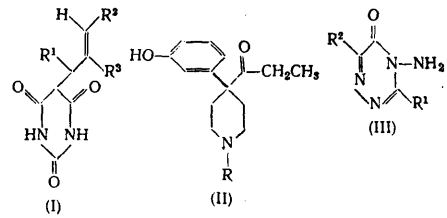
Together with the first-order connectivity index, higher-order connectivity indices have been considered in a number of studies.[10,123] For example, the topological index $^2\chi^v$, calculated for phenol derivatives, has been used to analyse their antifungal activity characterised by the parameter *MIC*.

We shall now consider other topological indices. A series of correlations between the indices I_{orb} and I_{sym}, on the one hand, and the inhibition of enzymes, alcohols, and their non-specific narcotic activity and toxicity, on the other, has been discussed in a monograph.[14]

Systematic studies on the "structure–biological activity" problem, based on the information indices *IC*, *SIC*, *BIC*, and *CIC*, have been carried out.[60–73,124] The relation between the structure and toxicity of ethers, characterised by the parameter LD_{50}, has been analysed.[60] The quantities $^1\chi$, $^1\chi^v$, χ^v, *W*, *IC*, *SIC*, *CIC*, I_D^W, and $\bar{I}_D^W$ were used as the TI and it was established that the most suitable structural characteristic among those quoted is the index $^1\chi^v$ in this instance. A correlation equation of the following type has been obtained in terms of the TI:

$$\lg LD_{50} = -0.7207\,{}^1\chi^v - 1.3868 \quad (r = 0.938;\ s = 0.246;\ N = 15).$$

The indices *IC*, *CIC*, $^1\chi$, $^1\chi^v$, I_D^W, $\bar{I}_D^W$, and *W* as well as the parameters lg *P* and the van der Waals volume V_W have been used as structural characteristics to estimate the anaesthesising activities of barbiturates.[61] Four groups of biologically active molecules have been investigated by Basak et al.:[63] alcohols (toxicity); barbiturates (anaesthetic (I); ketobemidones (analgesics) (II), and triazinones (herbicides) (III). There were all characterised with the aid of the indices *W*, *IC*, *SIC*, *CIC*, I_D^W, $\bar{I}_D^W$, $^1\chi$, and $^1\chi^v$.



(I) (R^1,R^2,R^3 = alkyl groups)

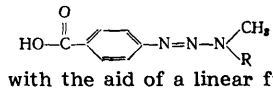
(II) [$R = CH_3(CH_2)_n$, where $n = 0, ..., 9$]

(III) ($R^1 = SCH_3$; R^2 = alkyl groups)

It was established that the most convenient structural characteristics in such cases are the indices *CIC* and $\bar{I}_D^W$ and the most suitable type of functional relation is quadratic.

The indices *IC*, *SIC*, and *CIC* are used to characterise derivatives of piperidine (inhibition of butyrylcholinesterase), barbiturates (hyponotic activity), alkanes (toxicity), aliphatic alcohols (local anaesthesia, narcotic effect on crustacean larvae), and a mixed group of compounds which are tranquillizers.[66-68]

The antitumour activity of triazenes of the type



has been estimated with the aid of a linear function of two variables lg *P* and ^2BIC.[124] The same type of biological activity has been considered in a study[69] where a wider set of parameters was used, namely lg *P*, V_W, *IC*, and *SIC*. The index *SIC* is also useful in the analysis of the biological activity of triazinones (herbicides) and *N*-alkylnorketobemidones (analgesics).[70] The higher-order indices ^kCIC, ^kIC, and ^kSIC, where $k > 1$, have been used[64] to analyse the biological activity and structures of a series of organic compounds. In particular, correlations of the type lg $LD_{50} = ay^2 + by + c$ between the toxicity of barbiturates (LD_{50}) and characteristics such as lg *P*, *W*, I_D^W, $\bar{I}_D^W$, $^1\chi^v$, ^2SIC, ^2CIC, and $^1\chi\lambda$ have been described. The most accurate correlation has been obtained using the indices ^2SIC and ^2CIC.

In some of the studies enumerated above, it is emphasised that the use of TI as structural parameters instead of characteristics such as lg *P* leads to more accurate correlations. The presence of several correlations, obtained in terms of different structural parameters, for the same property makes it possible to identify those which are most significant for the given property. Hypotheses concerning the mechanism of the action of active compounds can be constructed on the basis of such analysis.

The application of the index $^2\kappa$ (see above) in the quantitative estimation of the properties of chemical compounds has been discussed.[142] For example, it has been suggested that the index $^2\kappa$ be used as the characteristic of the molecular shape; it has been shown that, for similar values of $^2\kappa$ (i.e. for similar shapes), the substances have similar smells. A relation has been established between $^2\kappa$ and the bactericidal activity of phenol derivatives with halogenoalkyl substituents:

$$\lg PC = 1.817\,{}^2\kappa - 0.131\,({}^2\kappa)^2 - 2.899 \quad (r = 0.955;\ s = 0.27;\ N = 49),$$

where *PC* is the minimum concentration of the substance inducing specified biological effects in 10 min. The index $^2\kappa$ has also been used[142] to estimate the sweetness of nitro- and cyano-anilines.

Together with the TI mentioned in this Section, a series of topological indices, originating from the study of the electronic structures and chemical reactivities of biologically active compounds with the aid of topological approximations of quantum chemistry, have also been used in estimating biological activities. This field developed most vigorously in the 1960's and it has been described in detail in a monograph.[125]

V. QUANTUM-CHEMICAL METHODS AS A SOURCE OF TOPOLOGICAL INDICES

1. The Hückel Method and Graph Theory

In the examples considered previously, the TI arose as a rule as a result of the formal analysis of the structure of the molecular graph. Such analysis was supplemented in many instances by qualitative considerations concerning the choice of criteria for the estimation of the degree of branching or compactness of the molecule. In this Section, we shall discuss TI which appear naturally within the framework of certain mathematical models used to investigate the electronic structures of molecules. We shall follow in relation to one such model how the solution of the problem of the electronic structures of conjugated molecules reduces to the analysis of the TI of certain MG.

A molecule with a system of conjugated double or triple bonds contains as a rule a fairly extended planar fragment, which makes it possible to sub-divide the electronic system into two parts—the σ- and π-components—and to investigate each system independently.[125-128] Later we shall consider only the π-electron system assuming that it exists in a certain effective field generated by the atomic nuclei, the electron inner shells, and the σ electrons participating in the formation of localised two-centre σ-bonds. In the case of conjugated hydrocarbons, the number of π electrons is usually the same as the number of carbon atoms forming the conjugated part of the molecule. The structure of the conjugated part can be described by a molecular graph in a simplified form. It is obtained from the complete MG by removing the vertices corresponding to univalent substituents.

The properties of a system incorporating *m* π electrons are known[125-128] to be described in terms of one-electron energy levels ε_k and the corresponding molecular orbitals

$$\psi_k(r) = \sum_{j=1}^{m} c_j^{(k)} \varphi_j(r),$$

where $\phi_j(r)$ is an atomic $2p_z$ orbital centred on the *j*th atom of the conjugated part of the molecule. We shall denote by $\mathbf{c}^{(k)}$ the unit vector with the components $c_1^{(k)}, ..., c_m^{(k)}$. Then, in the case of hydrocarbon systems, the vectors $\mathbf{c}^{(k)}$ and one electron energy levels are found by solving the system of equations $H\mathbf{c}^{(k)} = \varepsilon_k \mathbf{c}^{(k)}$, where $H = \alpha E + \beta A$, *E* is the *m*th order unit matrix, *A* the adjoint matrix of the MG describing the conjugated part of the molecule, and α and β are the Coulombic and resonance integrals. For heteroatomic systems, the matrix *A* must be replaced by the adjoint matrix of the weighted MG.

We shall denote by λ_j (where $j = 1, ..., n$) the eigenvalues of the matrix *A* arranged in decreasing order. The one-electron energy levels ε_j can be then represented in the form $\varepsilon_j = \alpha + \lambda_j\beta$, where $\beta < 0$. The eigen-vectors of the matrices *H* and *A* are identical. The energy of the ground state of the system with an even number of *m* of π electrons can be calculated in terms of λ_j by the formula

$$\varepsilon_\pi = 2\sum_{j=1}^{m/2} \varepsilon_j = m\alpha + 2\beta \sum_{j=1}^{m/2} \lambda_j .$$

The density matrix *P*, whose elements are

$$p_{ij} = 2\sum_{k=1}^{m/2} c_i^{(k)} c_j^{(k)},$$

is defined in order to analyse the distribution of electron density. The diagonal element p_{ii} of the matrix *P* is called the π-electron density at the *i*th atom, while the non-diagonal element p_{ij}, corresponding to atoms linked by valence bonds, is called the order of the π-bond (i, j). The elements of the matrix *P* are TI of the local type and in terms of these TI

the formula of the total π-electronic energy of the ground state of a conjugated system containing heteroatoms can be formulated as follows:

$$\varepsilon_\pi = \sum_i \alpha_i p_{ii} + \sum_{i \neq j} p_{ij}\beta_{ij} ,$$

where $\alpha_i = H_{ii}$ and $\beta_{ij} = H_{ij}$ are respectively the Coulombic and resonance intervals. It can be shown[127] that the elements of the matrix P characterise the response of a π-electron system to small perturbations leading to changes in one of the parameters α_i or β_{ij} and can be found as partial derivatives of ε_π with respect to the corresponding parameters α_i and β_{ij}. The use of the topological indices p_{ij} for the interpretation of data on the chemical reactivity of conjugated molecules is based on this property.[127]

In the case of hydrocarbons, the expression for the energy of the ground state ε_π includes as a rule only non-negative eigenvalues λ_j of the corresponding MG. The energy levels ε_j located below the Coulombic integral α correspond to such eigenvalues. The energy levels ε_j satisfying the inequality $\varepsilon_j < \alpha$ are called bonding.[125] The term "antibonding" is used to characterise the energy levels to which $\lambda_j < 0$ correspond. On the other hand, if $\lambda_j = 0$, then $\varepsilon_j = \alpha$; such an energy level is referred to as non-bonding. If the number of bonding orbitals is $m/2$, then it is said that the π-electron system has a closed (in the Hückel sense) electron shell.

The character of the electron shell of conjugated hydrocarbons can be interpreted in terms of the topological index $\omega = n_+(G)/m$, where $n_+(G)$ is the number of positive eigenvalues of the molecular graph G. If $\omega = 1/2$, then the molecule has a closed (in the Hückel sense) electron shell. For $\omega > 1/2$, there is a deficiency of π electrons and, when $\omega < 1/2$, there is an excess of π electrons. The presence of a closed π-electron shell in a conjugated hydrocarbon molecule is attributed to the stability of the π-electron system and the molecule as a whole.[125–127] Qualitative conclusions used[126–139] for *a priori* estimates of the possibility of the existence of chemical compounds are based on this precise hypothesis. The results of studies[135–139] where the problem of the existence of polyhedral C_n carbon clusters is discussed are based to a large extent on the criterion specifying whether or not the electron shell is closed in the Hückel sense.¶

It follows from the data described in this Section that the study of the electronic structure of conjugated systems in terms of the Hückel π-electron approximation reduces to finding a series of invariants of molecular graphs corresponding to conjugated systems. We shall indicate the most important of these quantities: the eigenvalues of the MG $\{\lambda_j\}$, the sum of the positive eigenvalues

$$\lambda_+ = \sum_{\lambda_j > 0} \lambda_j$$

the π-electronic charges on the atom p_{ii}, and the π-electron bond orders p_{ij}. A more complete list of the TI used in the Hückel method can be found in Table 2.

The TI used for the purpose of classification occupy a special place in the theory of conjugated systems. They include, for example, the conjugation energy and the topological resonance energy (TRE). The latter was introduced by Aihara[140] and Gutman et al.[141] and has been used to estimate the relative stabilities and degrees of aromaticity of conjugated molecules. The topological resonance energy is determined in terms of the zero values of the so called acyclic polynomial in the construction of which use is made of the combinatorial properties of the coefficients a_j characterising the polynomial $P(G,x)$ of the molecular graph G. It can be shown[23] that each coefficient a_j represents the sum of the contributions corresponding to special subgraphs of the graph G. Among the subgraphs, there are those which contain cyclic components. We shall denote by a_j' the sum of the contributions introduced into the expression for a_j solely by the subgraphs which do not contain cyclic components. The polynomial $P(G,x)$ with the coefficients a_j' is called acyclic.[141] It has only real roots and is regarded as the characteristic polynomial of the hypothetical π-electron system with the same number of π electrons as in the initial system. The difference between the energies of the ground states of these two systems is called the topological resonance energy. These questions are discussed in greater detail in the review of Graovac et al.[142] Certain new TI used to estimate the degrees of aromaticity[144,145] have been compared.[143]

Table 2. Certain topological indices used in the Hückel method.[126]

Topological index ($m = 2k$)	Name	Application
$\lambda_1, \ldots, \lambda_m$	Eigenvalues of the MG adjacency matrix	Calculation of one-electron energy levels $E_i = \alpha + \beta\lambda_i$; interpretation of photoelectron and UV spectra
$\lambda_+ = \lambda_1 + \ldots + \lambda_{m/2}$	Sum of positive eigenvalues of the MG adjacency matrix	Calculation of the energy of the ground state of the π-electron system ($E_\pi = m\alpha + 2\beta\lambda_+$); estimation of the relative stabilities of the conjugated systems; calculation of the resonance energy
$p_{ij} = 2\sum_{k=1}^{m/2} c_i^{(k)} c_j^{(k)}$	Elements of the density matrix: p_{ii} is the π-electron density at the atom i and p_{ij} at the π-electron bond order	Bond orders—estimation of bond lengths, force constants, and chemical reactivity; p_{ii} elements-estimation of the dipole moments and reactivities
$F_i = N_{\max} - \sum_{j \neq i} p_{ij}$	Free valence index	Used as the reactivity index in radical substitution reactions; $N_{\max}$ is the maximum permissible value of $\sum_{i \neq j} p_{ij}$ for atoms of the given type
$\lambda_{m/2}$, $\lambda_{m/2+1}$; $\{c_i^{(m/2)}\}$, $\{c_i^{(m/2+1)}\}$	The smallest positive and the greatest negative eigenvalues of the MG adjacency matrix and the corresponding eigenvectors	Estimation of the first ionisation potential, electron affinity, and the energy of the first electronic transition; polarography; theory of chemical reactivity

Notation: $m = 2k = N$ is the number of centres in the conjugated systems, where N is the number of π electrons; $\lambda_j \neq 0$ for $j = 1, \ldots, m$.

Various correlations between TI (of the Hückel type) and different physicochemical properties of conjugated molecules have been described fairly completely in monographs.[125–127] The usefulness of the Hückel method for the interpretation of the photoelectron spectra of polycondensed aromatic hydrocarbons has been discussed.[146] It has been shown for more than 250 compounds that the potentials of ionisation from the upper π-energy levels can be fitted with an adequate accuracy (the standard deviation does not exceed 0.1 eV) by analytical

¶ See also D.J.Klein, W.A.Seitz, and T.G.Schmalz, Nature, **1986**, **323**, 703.

expressions whose form is determined by the one-electron energy levels and the bond orders.

Quantities such as the eigenvalues of the MG, the elements of the density matrix, and other characteristics of the molecular systems can be readily calculated with the aid of a computer. However, these data, many of which have been compiled in handbooks (see, for example, Streitwieser and Brauman[147]), are difficult to interpret directly in structural terms. Nevertheless such interpretation is of interest for the understanding of the sources of the appearance of definite functional relations between the structure of the molecule and its properties. The field associated with the search for relations between the results of Hückel calculations and clear characteristics of molecular graphs has therefore developed fairly vigorously in recent years.[142] For example, a complete series of approximate formulae of different degrees of complexity has been obtained for the sum of the positive eigenvalues of the MG, which, as mentioned above, is linearly related to the total π-electronic energy. A catalogue of the most convenient of these together with estimates of their accuracy has been published.[148,149] Analogous formulae have been obtained for the maximum eigenvalue and also for the difference between the smallest positive and greatest negative eigenvalues of the adjoint matrix.[150–152] The latter quantity is of interest for the estimation of the spectroscopic characteristic of conjugated molecules.

The mathematical model of the electronic structures of unsaturated molecules described in this Section can be extended also to the case where it is necessary to take into account all the valence electrons (for example, to saturated systems). One of the possible extensions has been proposed in studies[153,154] where the investigation of the electronic structure of molecules in terms of the valence approximation reduces to the analysis of the spectroscopic characteristics of weighted graphs of a special type. The vertices of these graphs correspond to different hybrid orbitals constructed from atomic functions of the *s* and *p* types.

2. Relations between Certain Topological Indices of the Quantum-Chemical Type

As mentioned above in Section II, a series of TI are linked by simple functional relations. Some of these relations are empirical and have been established by the method of least squares. Similar relations have been obtained in a number of studies[10,95,96,155–158] between certain topological indices defined in the Hückel method and indices which do not have a direct bearing on the method (such as, for example, the connectivity indices, the Wiener numbers, and others). It has been shown[155] that, for conjugated hydrocarbons, the first-order connectivity index $^1\chi^\nu$, determined from the formula

$$^1\chi^\nu = {\sum_{i,j}}' (\delta_i^\nu \delta_j^\nu)^{-1/2},$$

is fairly satisfactorily correlated with the total π-electronic energy. In this formula $\delta_i^\nu = z_i^\nu - h_i$, where z_i^ν is the number of valence electrons of the *i*th atom and h_i the number of univalent substituents of the given atom. Relations of an analogous type have been established[156] between the charges on the atoms and the Mulliken bond orders obtained by the extended Hückel method for aliphatic hydrocarbons, on the one hand, and the atomic connectivity indices χ_i^ν, found from the formula

$$\chi_i^\nu = \frac{1}{2} {\sum_j}' (\delta_i^\nu \delta_j^\nu)^{-1/2},$$

on the other. Such formulae are useful, since they make it possible to imitate fairly simply the results of calculations of the electronic structure of molecules by MO methods, which cannot be carried out without using modern computers.

A series of TI appear naturally when another quantum-chemical method is used—the valence bond method[159] and some of its modifications.[144,145] Methods of this type are based on the construction of a many-electron wave function in the form of a linear combination of many-electron basis functions of a special type. In the case of the π-electron approximation, where only the π orbital is taken into account for each atom, a certain spin pairing scheme for the π electrons associated with the corresponding carbon atoms corresponds to each basis function. This scheme can be described in terms of the graph theory. Only electrons corresponding to atoms linked by valence bonds are usually combined in pairs. Each such scheme is described by a Kekulé graph. It consists of a set of connected two-vertex subgraphs MG (edges), which cover the vertices of the initial graph and have no common vertices in pairs. The Kekulé graph can be related to a structural formula with single and double bonds. Suppose that an approximate wave function ψ_k corresponds to each such structural formula. Then the total wave function ψ can be formulated as the linear combination

$$\psi = \sum_k a_k \psi_k.$$

The numbers a_k are called the weighting factors of the structures and are treated as variational parameters. The procedure described above is usually employed to investigate the stability of bicyclic benzenoid hydrocarbons and their heteroanalogues. On the basis of qualitative considerations, it can be shown that, among the isomers of benzenoid hydrocarbons, the greatest stability should be shown by the isomer to which corresponds the greatest number of Kekulé structures.[159]

The following procedures are used to estimate the bond strength. The number of all the Kekulé structures is found. Then the edge, for example (i,j), corresponding to a bond C–C in the initial molecule, is selected in the MG. Next this edge together with the edges incident with it are excluded from the MG. We shall denote the number of Kekulé structures in the graph obtained in this way by k_{ij}. The ratio $\tilde{p}_{ij} = k_{ij}/k$ is called the Pauling bond order.[160,161] It has been established that, for alternative hydrocarbons not containing four-membered rings and non-bonding energy levels (which is equivalent to the validity of the inequality $\det A \neq 0$, where A is the adjoint matrix), this quantity can be calculated from the formula

$$\tilde{p}_{ij} = \sum_{\lambda_k > 0} c_i^{(k)} c_j^{(k)} / \lambda_k,$$

where $c_i^{(k)}$ are the components of the *k*th eigenvector of the adjoint matrix A. The quantity on the right-hand side was introduced by Ham and Ruedenberg[161] and is called the Ham and Ruedenberg bond order. A number of other relations between TI, used in various quantum-chemical methods, have also been discussed (see, for example, Hosoya and Hosoi[157]).

The examples presented in Sections II and V show that there are simple relations, determined ultimately by the combinatorial properties of the initial MG, between TI which are at first sight of different nature. The search for various systems of independent TI, in terms of which it would be possible to calculate any other TI, is therefore of interest. The solution of this problem would make it possible to construct various correlations more effectively and with better justification and to predict the properties of compounds which have not yet been synthesised.

---o0o---

Two main aspects of the problem of topological indices can be distinguished. The first is associated with the development of effective methods for the coding of molecules, which is of great interest in connection with the use of information-search systems for different purposes. The second aspect refers directly to the search for sets of structural fragments of molecules which are responsible for the manifestation of particular physicochemical properties of concrete compounds and to the discovery of regularities in the character of their variation.

It appears at first sight surprising that the properties of a substance can be predicted with the aid of simple mathematical models of the kind described in this review. However, experience shows that the properties of many chemical compounds, whose molecules have "like" structural fragments, are indeed similar. Furthermore, one should bear min mind that, in selecting the coefficients a_i in functional relations of the type

$$y = a_0 + \sum_i a_i k_i,$$

between the numerical values of a certain property and the structural characteristics expressed in terms of TI, one takes into account implicitly not only the geometrical parameters of individual molecules (more precisely certain average characteristics) but also the features of the structure of the substance as a whole. The same property can be described in terms of different TI, which leads to unlimited possibilities for the search with the aid of computers for relations between changes in the properties of a substance and the topological characteristics of the molecular structure.† In practice, preference should be given to simpler and more accurate relations which can be used even when it has not been possible to justify reasonably the relations discovered. In his Nobel Lecture, Richard Feynman said "the only true physical description of a phenomenon involves the interpretation of the significance of the quantities in an equation from the standpoint of the experiment or, more precisely, a method of applying equations to experimental results. Once this is admitted, then the best way of creating a new theory is to guess at equations without paying attention to physical models or a physical explanation".[164] From our point of view, the method of topological indices is precisely such an approach, which makes it possible to "guess" equations relating structure to property in those fields of structural chemistry where a sufficiently reliable theory has not yet been developed or the applications of a theory encounter considerable computational difficulties.

REFERENCES

1. A.M.Butlerov, "Izbrannye Raboty po Organicheskoi Khimii" (Selected Works on Organic Chemistry), Izd.Akad.Nauk SSSR, Moscow, 1951.
2. V.V.Markovnikov, "Izbrannye Trudy" (Selected Works), Izd.Akad.Nauk SSSR, Moscow, 1955.
3. V.A.Legasov and A.L.Buchachenko, Uspekhi Khim., 1986, 55(12), 1949 [Russ.Chem.Rev., 1986(12)].
4. V.M.Tatevskii, "Klassicheskaya Teoriya Stroeniya Molekul i Kvantovaya Mekhanika" (Classical Theory of Molecular Structure and Quantum Mechanics), Izd.Khimiya, Moscow, 1973.
5. V.M.Tatevskii, "Khimicheskoe Stroenie Uglevodorodov i Zakonomernosti v Ikh Fiziko-Khimicheskikh Svoistvakh" (The Chemical Structure of Hydrocarbons and Regularities in Their Physicochemical Properties), Izd.Moskov.Gos.Univ., Moscow, 1953.
6. V.M.Tatevskii and Yu.G.Papulov, Zhur.Fiz.Khim., 1960, 34(4), 708 [Russ.J.Phys.Chem., 1960(4)].
7. V.M.Tatevskii, V.A.Benderskii, and S.S.Yarovoi, "Metody Rascheta Fiziko-Khimicheskikh Svoistv Parafinovykh Uglevodorodov" (Methods for the Calculation of the Physicochemical Properties of Paraffinic Hydrocarbons), Gostoptekhizdat, Moscow, 1960.
8. E.A.Smolenskii, Zhur.Fiz.Khim., 1964, 38(5), 1288 [Russ.J.Phys.Chem., 1964(5)].
9. D.V.Jain, S.Singh, and V.Gombar, Proc.Indian Acad.Sci.(Chem.Sci.), 1984, 93, 927.
10. L.B.Kier and L.H.Hall, "Molecular Connectivity in Chemistry and Drug Research", Acad.Press, New York, 1976.
11. F.Harary, "Graph Theory" (Translated into Russian), Izd.Mir, Moscow, 1973.
12. A.J.Stuper, W.E.Brugger, and P.C.Jurs, "Computer Assisted Studies of Chemical Structure and Biological Function" (Translated into Russian), Izd.Mir, Moscow, 1982.
13. A.T.Balaban, Pure and Appl.Chem., 1983, 55, 199.
14. D.Bonchev, "Information Theoretic Indices for Characterisation of Chemical Structures", Research Studies Press, Chichester, 1983.
15. A.T.Balaban, J.Chem.Inf.Comput.Sci., 1985, 25, 334.
16. A.T.Balaban, A.Chiriac, I.Motoc, and Z.Simon, Lecture Notes in Chemistry, 1980(15), 1.
17. A.Sablic and N.Trinajstic, Acta Pharm.Jugoslavia, 1981, 31, 189.
18. D.H.Rouvray, Mathemat.Chem.(MATCH),1975(1), 125.
19. D.H.Rouvray, "Chemical Applications of Topology and Graph Theory" (Edited by R.B.King), Elsevier Science Publ., Amsterdam, 1983, p.159.
20. H.Wiener, J.Amer.Chem.Soc., 1947, 69, 17.
21. D.H.Rouvray and B.C.Crafford, South African J. Sci., 1976, 72, 47.
22. K.Altenburg, Kolloid Z., 1961, 178, 112.
23. D.Tsvetkovich, M.Dub, and H.Sachs, "Spektry Grafov.Teoriya i Primenenie" (Spectra of Graphs. Theory and Application), Izd.Naukova Dumka, Kiev, 1984.
24. A.T.Balaban, Theor.Chim.Acta, 1979, 53, 355.
25. D.Bonchev, A.T.Balaban, and O.Mekenyan, J.Chem. Inf.Comput.Sci., 1980, 20, 106.
26. M.Razinger, J.R.Chretien, and J.E.Dubois. Vestn. Slov.Kem.Drus., 1986, 33, 49.
27. A.A.Dobrynin, N.M.Meirmanova, and V.A.Skorobogatov, in "Tezisy Dokladov VII Vsesoyuznoi Konferentsii 'Ispol'zovanie Vychislitel'nykh Mashin v Khimicheskikh Issledovaniyakh i Spektroskopii Molekul" (Abstracts of Reports at the VIIth All-Union Conference on 'The Use of Computers in Chemical Research and Molecular Spectroscopy'), Riga, 1986, p.203.
28. M.Barysz, G.Jashari, R.S.Lall, V.K.Srivastava, and N.Trinajstic, in "Chemical Applications of Topology and Graph Theory) (Edited by R.B.King), Elsevier Science Publishers, Amsterdam, 1983, p.222.
29. Yu.A.Kruglyak and I.P.Peredunova, in "Tezisy Dokladov VII Vsesoyuznoi Konferentsii 'Ispol'zovanie Vychislitel'nykh Mashin v Khimicheskikh Issledovaniyakh i Spektroskopii Molekul'" (Abstracts of Reports at the VIIth All-Union Conference on 'The Use of Computers in Chemical Research and Molecular Spectroscopy'), Riga, 1986, p.115.

†Information about programs for the calculation of different TI on a computer has been published, for example, by Peredunova et al.[162] and Shtain et al.[163]

30. I.Gutman, B.Ruscic, N.Trinajstic, and C.F.Wilcox, J.Chem.Phys., 1975, **62**, 3399.
31. H.Narumi and M.Katayama, Mem.Fac.Eng.Hokkaido Univ., 1984, **16**(3), 209.
32. J.Platt, J.Phys.Chem., 1952, **56**, 328.
33. M.Randic, J.Amer.Chem.Soc., 1975, **97**, 6609.
34. L.B.Kier, L.H.Hall, W.J.Murray, and M.Randic, J.Pharm.Sci., 1975, **64**, 1971.
35. L.B.Kier and L.H.Hall, Eur.J.Med.Chem.Chim., Therapeut., 1977, **12**, 307.
36. L.B.Kier and L.H.Hall, J.Pharmacol.Sci., 1983, **72**, 1170.
37. L.B.Kier and L.H.Hall, J.Pharmacol.Sci., 1976, **65**, 1806.
38. L.A.Evans, M.F.Lynch, and P.Willet, J.Chem.Inf. Comput.Sci., 1978, **18**, 146.
39. V.V.Drboglav and V.E.Golender, "Tezisy Dokladov VII Vsesoyuznoi Konferentsii 'Ispol'zovanie Vychislitel'nykh Mashin v Khimicheskikh Issledovaniyakh i Spektroskopii Molekul'" (Abstracts of Reports at the VIIth All-Union Conference on 'The Use of Computers in Chemical Research and Molecular Spectroscopy'), Riga, 1986, p.217.
40. A.B.Rozenblit and V.E.Golender, "Logiko-Kombinatornye Metody Konstruirovaniya Lekarstv" (The Logical-Combinatorial Methods for the Design of Medicinal Drugs), Izd.Zinatne, Riga, 1983.
41. M.Gordon and G.R.Scantlebury, Trans.Faraday Soc., 1964, **60**, 604.
42. L.B.Kier, Quant.Struct.-Act.Relat., 1985, **4**, 109.
43. M.Barysz, N.Trinajstic, and J.V.Knop, Int.J.Quant. Chem.Symp., 1983(17), 441.
44. M.Barysz, J.V.Knop, S.Pejakovic, and N.Trinajstic, Polish J.Chem., 1985, **59**, 405.
45. M.Randic, G.A.Kraus, and B.Dznonova-Jerman-Blazic, Chemical Applications of Topology and Graph Theory (Edited by R.B.King), Elsevier Science Publ., Amsterdam, 1983, p.192.
46. H.Hosoya, Bull.Chem.Soc.Japan, 1971, **44**, 2332.
47. K.Kawasaki, K.Mizutani, and H.Hosoya, Nat.Sci., Rep.Ochanomizu Univ., 1971, **22**, 181.
48. K.Mizutani, K.Kawasaki, and H.Hosoya, Nat.Sci., Rep.Ochanomizu Univ., 1971, **22**, 39.
49. L.Lovasz and J.Pelican, Period.Math.Hung., 1973, **3**, 175.
50. M.I.Gverdtsiteli, R.O.Devdariani, and O.G.Devdariani, in "Tezisy Dokladov VII Vsesoyuznoi Konferentsii 'Ispol'zovanie Vychislitel'nykh Mashin v Khimicheskikh Issledovaniyakh i Spektroskopii Molekul'" (Abstracts of Reports at the VIIth All-Union Conference on 'The Use of Computers in Chemical Research and Molecular Spectroscopy'), Riga, 1986, p.212.
51. Yu.A.Zhdanov, "Entropiya Informatsii v Organicheskoi Khimii" (Information Entropy in Organic Chemistry), Izd.Rostov.Gos.Univ., Rostov-on-Don, 1979.
52. S.M.Dancoff and H.Quastler, "Essays on the Use of Information of Theory in Biology" (Edited by H.Quastler), University of Illinois, Urbana, 1953; quoted after Ref.14.
53. H.J.Morowitz, Bull.Math.Biophys., 1955, **17**, 81.
54. S.V.Dozmorov, in "Tezisy Dokladov V Vsoyuznoi Konferentsii 'Ispol'zovanie Vychislitel'nykh Mashin v Molekulyarnoi Spektroskopii i Khimicheskikh Issledovaniyakh" (Abstracts of Reports at the Vth All-Union Conference on 'The Use of Computers in Molecular Spectroscopy and Chemical Research'), Novosibirsk, 1980, p.28.
55. N.Rashevsky, Bull.Math.Biophys., 1955, **17**, 229.
56. E.Trucco, Bull.Math.Biophys., 1956, **18**, 237.
57. E.Trucco, Bull.Math.Biophys., 1956, **18**, 129
58. A.Mowshowitz, Bull.Math.Biophys., 1968, **30**, 533.
59. S.H.Bertz, J.Amer.Chem.Soc., 1981, **103**, 3599.
60. S.C.Basak, D.P.Gieschen, and V.R.Magnuson, Environ.Toxicol.and Chem., 1984, **3**, 191.
61. S.C.Basak, D.P.Gieschen, V.R.Magnuson, and D.K.Harris, IRCS Med.Sci., 1982, **10**, 619.
62. S.C.Basak and V.R.Magnuson, Arzneim.-Forsch., 1983, **33**, 501.
63. S.C.Basak, D.K.Harriss, and V.R.Magnuson, J. Pharm.Sci., 1984, **73**, 429.
64. V.R.Magnuson, D.K.Harriss, and S.C.Basak, in "Chemical Applications of Topology and Graph Theory" (Edited by R.B.King), Elsevier Science Publishers, Amersterdam, 1983, p.178.
65. S.C.Basak, S.K.Ray, C.Raychaudhury, A.B.Roy, and J.J.Ghosh, IRCS Med.Sci., 1982, **10**, 145.
66. S.K.Ray, S.C.Basak, C.Raychaudhury, A.B.Roy, and J.J.Ghosh, Ind.J.Pharmac., 1981, **13**, 301.
67. C.Raychaudhury, S.C.Basak, A.B.Roy, and J.J.Ghosh, Indian Drugs, 1980, **18**(3), 97.
68. S.K.Ray, S.C.Basak, C.Raychaudhury, A.B.Roy, and J.J.Ghosh, Indian J.Chem., 1981, **20B**, 894.
69. S.K.Ray, S.C.Basak, C.Raychaudhury, A.B.Roy, and J.J.Ghosh, Arzneim.-Forsch.Drug.Res., 1983, **33**, 352.
70. S.K.Ray, S.C.Basak, C.Raychaudhury, A.B.Roy, and J.J.Ghosh, Arzneim.-Forsch.Drug.Res., 1982, **32**, 322.
71. R.Sarkar, A.B.Roy, and P.K.Sarkar, Math.Bioscience, 1978, **39**, 299.
72. S.C.Basak, A.B.Roy, and J.J.Ghosh, "Proceedings of II International Conference on Mathematical Modelling in Science and Technology", University of Missouri, Rolla, 1979, p.851.
73. S.K.Ray, S.Gupta, S.C.Basak, et al., Indian J. Chem., 1985, **24B**, 1149.
74. D.Bonchev, D.Kamenski, and V.Kamenska, Bull. Math.Biol., 1976, **38**, 119.
75. I.Motoc, A.Balaban, O.Mekenyan, and D.Bonchev, Mathemat.Chem.(MATCH), 1982, **13**, 369.
76. K.Altenburg, Z.phys.Chem.(Leipzig), 1984, **265**, 257.
77. D.Bonchev, O.Mekenyan, and N.Trinajstic, J. Comput.Chem., 1981, **2**, 127.
78. M.Razinger, Theor.chim.Acta, 1986, **70**, 365.
79. O.E.Polansky, Monatsh.Chem., 1983, **116**, 211.
80. C.Raychaudhury, S.K.Ray, J.J.Ghosh, et al., J.Comput.Chem., 1984, **5**, 581.
81. V.V.Drboglav and V.E.Golender, in "Tezisy Dokladov VII Vsesoyuznoi Konferentsii 'Ispol'zovanie Vychislitel'nykh Mashin v Khimicheskikh Issledovaniyakh i Spektroskopii Molekul'" (Abstracts of Reports at the VIIth All-Union Conference on 'The Use of Computers in Chemical Research and Molecular Spectroscopy'), Riga, 1986, p.219.
82. I.Motoc and A.Balaban, Rev.Roum.Chim., 1981, **26**, 593.
83. D.Bonchev and N.Trinajstic, J.Chem.Phys., 1977, **67**, 4517.
84. N.F.Komshilov, Zhur.Obshch.Khim., 1940, **10**, 945.
85. N.F.Komshilov, Zhur.Obshch.Khim., 1939, **9**, 701.
86. M.L.Huggins, J.Amer.Chem.Soc., 1941, **63**, 116.
87. M.L.Huggins, J.Amer.Chem.Soc., 1954, **76**, 343.
88. K.Li, R.L.Arnett, M.B.Epstein, R.B.Ries, L.P.Bitler J.M.Lynch and F.D.Rossini, J.Phys.Chem., 1956, **60**, 1400.
89. A.Mibashan, Trans.Faraday Soc., 1945, **41**, 374.

90. H.Wiener, J.Phys.Chem., 1948, 52, 425.
91. H.Wiener, J.Amer.Chem.Soc., 1947, 69, 2636.
92. H.Wiener, J.Chem.Phys., 1947, 15, 766.
93. H.Wiener, J.Phys.Chem., 1948, 52, 1082.
94. D.G.Bonchev, Abstract of Doctoral Thesis in Chemical Sciences, Moscow State University, Moscow, 1984.
95. D.Bonchev and O.Mekenyan, Z.Naturforsch., 1980, 35a, 739.
96. D.Bonchev, O.Mekenyan, and O.E.Polansky, Z. Naturforsch., 1981, 36a, 643.
97. D.Bonchev, O.Mekenyan, G.Protic, and N.Trinajstic, J.Chromatog., 1979, 176, 149.
98. N.Trinajstic, G.Protic, V.Svob, and Dj.Deur-Siftar, Kem.Ind.(Zagreb.), 1979, 28, 527.
99. D.Bonchev, O.Mekenyan, and H.G.Fritsche, Phys. Stat.Sol.(a), 1979, 55, 181.
100. O.Mekenyan, D.Bonchev, and H.Fritsche, Phys. Stat.Sol.(a), 1979, 56, 607.
101. H.Hosoya, K.Hosoi, and I.Gutman, Theoret.Chim. Acta, 1975, 38, 37.
102. H.Hosoya and M.Murakami, Bull.Chem.Soc.Japan, 1975, 48, 3512.
103. A.T.Balaban and I.Motoc, Mathemat.Chem.(MATCH), 1979, 5, 197.
104. S.Hala, in "Sbornik Prazhskogo Khimiko-Tekhnologicheskogo Instituta. Tekhnologiya Topliv" (Collected Reports from the Prague Institute of Chemical Engineering. Technology of Fuels), 1982, Vol.D46, p.229.
105. L.Ekiert, J.Bojarski, and J.L.Mokrosz, Wiadomosci Chemiczne, 1986, 40, 65.
106. M.Valentinuzzi and M.E.Valentinuzzi, Bull.Math. Biophys., 1963, 25, 11.
107. G.Maroulis, M.Sana, and G.Leroy, Int.J.Quant.Chem., 1981, 19, 43.
108. D.Bonchev and N.Trinajstic, Int.J.Quantum.Chem. Symp., 1978(12), 293.
109. O.Mekenyan, D.Bonchev, and N.Trinajstic, Int.J. Quant.Chem., 1980, 18, 369.
110. L.B.Hall, L.H.Kier, and W.J.Murray, J.Pharmacol. Sci., 1975, 64, 1974.
111. A.Cammarata, J.Pharmacol.Sci., 1979, 68, 839.
112. V.Gombar, Pharmacos., 1981, 25, 46.
113. L.B.Kier, W.J.Murray, M.Randic, and L.H.Hall, J.Pharmacol.Sci., 1976, 65, 1226.
114. D.V.S.Jain and V.Gombar, J.Chem.Soc.Faraday Trans.I, 1979, 75, 1132.
115. W.J.Murray, L.N.Hall, and L.B.Kier, J.Pharmacol. Sci., 1975, 64, 1978.
116. J.C.Boyd, J.S.Millership, and A.D.Woolfson, J. Pharm.Pharmacol., 1982, 34, 364.
117. V.K.Gombar, V.K.Kapoor, and D.V.S.Jain, Indian J.Chem., 1980, 19A, 715.
118. L.B.Kier and L.H.Hall, J.Pharmacol.Sci.,1979, 68, 120.
119. S.V.Nizhnii and N.A.Epshtein, Uspekhi Khim., 1978, 47(4), 739 [Russ.Chem.Rev., 1978(4)].
120. D.Agin, L.Hersh, and D.Holtzman, Proc.Nat.Acad. Sci.(USA), 1965, 53, 952.
121. L.B.Kier, W.J.Murray, and L.H.Hall, J.Med.Chem., 1975, 18, 1272.
122. W.J.Murray, L.B.Kier, and L.H.Hall, J.Med.Chem., 1976, 19, 573.
123. A.K.Samanta, S.K.Ray, S.C.Basak, and S.K.Bose, Arzneim.-Forsch., 1982, 32, 1515.
124. S.K.Ray, S.C.Basak, C.Raychaudhury, A.B.Roy, and J.J.Ghosh, IRCS Med.Sci., 1982, 10, 933.
125. B.Pullman and A.Pullman, "Quantum Biochemistry" (Translated into Russian), Izd.Mir, Moscow, 1965.
126. A.Streitwieser, "Molecular Orbital Theory for Organic Chemists" (Translated into Russian), Izd.Mir, Moscow, 1965.
127. K.Higasi, H.Baba, and A.Rembaum, "Quantum Organic Chemistry" (Translated into Russian), Izd.Mir, Moscow, 1967.
128. E.Hückel, Z.Phys., 1931, 70, 204.
129. R.B.King and D.H.Rouvray, J.Amer.Chem.Soc., 1977, 99, 7834.
130. D.M.P.Mingos, J.Chem.Soc.Dalton Trans., 1977(1), 20.
131. D.M.P.Mingos, J.Chem.Soc.Dalton Trans., 1977(1), 26.
132. D.A.Bochvar and I.V.Stankevich, Zhur.Strukt. Khim., 1969, 10, 680.
133. D.A.Bochvar and I.V.Stankevich, Zhur.Strukt. Khim., 1971, 12, 142.
134. D.A.Bochvar and I.V.Stankevich, Zhur.Strukt. Khim., 1972, 13, 1123.
135. D.A.Bochvar and E.G.Gal'pern, Dokl.Akad.Nauk SSSR, 1973, 209, 610.
136. A.D.Haymet, Chem.Phys.Lett., 1985, 122, 421.
137. A.D.Haymet, J.Amer.Chem.Soc., 1986, 108, 319.
138. R.C.Haddon, L.E.Brus, and K.Raghavachari, Chem.Phys.Letters, 1986, 131, 165.
139. T.G.Schmaltz, W.A.Seitz, D.J.Klein, and G.E.Hite, Chem.Phys.Letters, 1986, 130, 203.
140. J.Aihara, J.Amer.Chem.Soc., 1976, 98, 2750.
141. I.Gutman, M.Milun, and N.Trinajstic, J.Amer.Chem. Soc., 1977, 99, 1692.
142. A.Graovac, I.Gutman, and N.Trinajstic, Lecture Notes in Chemistry, 1977(4), 1.
143. L.J.Schaad and B.A.Hess, Pure and Appl.Chem., 1982, 54, 1097.
144. M.Randic, J.Amer.Chem.Soc., 1977, 99, 444.
145. W.C.Herndon, J.Amer.Chem.Soc., 1973, 95, 2404.
146. E.Clar, J.M.Robertson, R.Schlögl, and H.Schmidt, J.Amer.Chem.Soc., 1981, 103, 1320.
147. A.Streitwieser and J.I.Brauman, "Supplemental Tables of Molecular Orbital Calculation", Pergamon Press, New York, 1965.
148. I.Gutman, L.Nedeljcowic, and A.V.Teodorovic, Bull. Soc.Chim.Beograd., 1983, 48, 495.
149. I.Gutman, Z.phys.Chem., 1985, 266, 59.
150. G.G.Hall, Molec.Phys., 1977, 33, 551.
151. I.Gutman and D.Rouvay, Chem.Phys.Letters, 1979, 62, 384.
152. Y.Kiang and E.Chen, Pure and Appl.Chem., 1983, 55, 283.
153. K.Fukui, H.Kato, and T.Yonezawa, Bull.Chem.Soc. Japan, 1960, 33, 1197.
154. K.Fukui, H.Kato, and T.Yonezawa, Bull.Chem.Soc. Japan, 1960, 33, 1201.
155. S.P.Gupta and P.Singh, Bull.Chem.Soc.Japan, 1979, 52, 2745.
156. N.K.Ray and S.Bhargava, Nat.Acad.Sci.Letters (India), 1978, 1, 141.
157. H.Hosoya and K.Hosoi, J.Chem.Phys., 1976, 64, 1065.
158. K.D.Gundermann, C.Lohberger, and M.Zander, Z.Naturforsch., 1981, 36a, 1217.
159. L.Pauling,"The Nature of the Chemical Bond", Cornell University Press, New York, 1960.
160. N.S.Ham, J.Chem.Phys., 1958, 29, 1229.
161. N.S.Ham and K.Ruedenberg, J.Chem.Phys., 1958, 29, 1215.
162. I.V.Peredunova, V.E.Kuz'min, and Yu.P.Konovorotskii, Zhur.Strukt.Khim., 1983, 24, 145.

163. L.M.Shtain, A.S.Kabankin, and M.A.Landau, Zhur.Strukt.Khim., 1985, **26**, 129.
164. R.P.Feynman, in "The Character of Physical Laws" (Translated into Russian), Izd.Mir, Moscow, 1968, p.229.

Lomonosov Moscow State University

Nesmeyanov Institute of Organoelemental Compounds, USSR Academy of Sciences, Moscow