General and Inorganic Chemistry

Order, disorder, and geometrical information indices of molecules

S. A. Dembovsky and P. A. Koz'min*

N. S. Kurnakov Institute of General and Inorganic Chemistry, Russian Academy of Sciences, 31 Leninsky prosp., 117907 Moscow, Russian Federation. Fax: +7 (095) 952 2382

Structural features of various molecular systems with symmetry of point groups ranging from C_1 to the icosahedral symmetry are analyzed in the framework of the model suggested previously for the evaluation of order and disorder in the arrangement of atoms in a molecule based on the equation $Q = 1 - P/3n$ (where Q is the index of order, and P is the number of independent coordinates needed to fix an n-atomic molecule in the Cartesian coordinate system). The Q value depends on various structural parameters of the molecule: the number of atoms in it, the symmetry, the dimensionality, and the number of structural degrees of freedom. The disorder index $P/3n = 1 - Q$ correlates with Shannon's entropy of information, and Q correlates with negentropy or excess information; this makes it possible to use $P/3n$ as a new geometrical information molecular index that is obtained by a non-probabilistic method. Analysis of the relationship between order and chaos in molecular systems, as well as of the specific order index $q = Q/n$, makes it possible to identify both general and specific features of molecules.

Key words: molecule, order, chaos, disorder, symmetry, point group, information, information index, information entropy, degree of freedom.

Introduction

The problem of chaos and order and their relationship in particular systems occupies one of prominent places in the structure of scientific knowledge. It is tackled not only in such branches of science as chemistry, physics, and cybernetics,[1–5] but even in medicine, psychology, economics, sociology, *etc.* In these cases, absolutely different objects are considered, and order is generally taken to mean correlations in space and time. In the present paper we analyze the problem of the structural (geometrical) order in molecules.

The concept of structural order/disorder is traditionally related to the presence of so-called long-range order, *i.e.*, translational invariance typical of the crystalline state; in this sense, order corresponds to a crystal with long-range order, and chaos is associated with a gas, a liquid, or a glass. During studies on the structure of matter, it becomes clear that no sharp boundary between chaos and order exists, and then the corresponding assumptions are introduced. For example, it is common knowledge that noncrystalline materials are characterized by short-range order determined by chemical bonds within the limits of 1-2 coordination spheres around an arbitrarily chosen atom (see, for example Ref. 5). Furthermore, the notion of "medium-range order" has been introduced to describe the situation in noncrystalline substances, for example in glasses and

Translated from *Izvestiya Akademii Nauk. Seriya Khimicheskaya*, No. 8, pp. 1912—1927, August, 1996.

1066-5285/96/4508-1810 $15.00

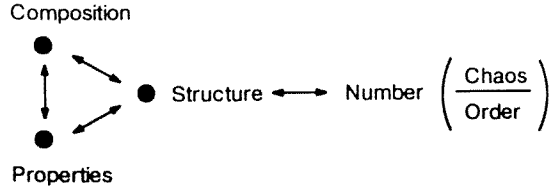
glass-forming liquids. The medium-range order occurs at distances of up to tenths of ångströms.[6] Like the long-range order in crystals, the medium-range order is manifested in the X-ray diffraction pattern of a sample as so-called first sharp diffraction peak, whose nature is being actively debated now.[7,8]

Thus the concepts of disorder (chaos) and order appeal first of all to the kinds of order that can be detected experimentally. A. I. Kitaigorodsky[9] proposed the following classification of states of order: (1) order *viz.*, ideal crystal without defects, (2) chaos, *viz.*, a gas without short-range order; (3) disorder in the order (for example, defects in crystals); (4) order in the disorder (for example, liquid crystals with orienational and short-range order but without long-range order or liquids with short-range order). Real situations can be represented as mixtures of the above states.

All the foregoing refers to order and chaos in macroscopic systems and not to individual molecules, although it is clear that molecules can differ not only in symmetry but also in the degree of order. No quantitative evaluation of order and chaos applied to molecules (or to systems or media consisting of these molecules) existed before the entropy of information was introduced. The example of four-stage description of order and chaos presented above was purely qualitative.

The problem of introducing a universal numerical characteristic of a structure, which would characterize the relationship between the order and chaos in it, is quite timely, since there is a great variety of molecules in existence. The symmetry of molecules varies from the lowest symmetry corresponding to the C_1 (1) point group (the common position multiplicity, *i.e.*, the number of symmetry operations after which the molecule remains invariant, M, is given in parentheses) to icosahedral symmetry I_h (120).

One of the key problems in chemistry and physics as well as in materials technology is the relationship between composition, structure, and properties, which is intimately connected to the problem of order and chaos and to the relationship between them. Regarding molecules, their structures and properties in the gas, liquid, and solid phases may be substantially dissimilar. The condensed solid phase may have a "liquid-like" (for example, glassy) structure or a crystalline structure. Thus the composition—structure—properties triad should be supplemented by an integral numerical (scalar) description of a molecule, which would, on the one hand, reflect its structure and, on the other hand, characterize order and chaos and the relationship between them.



This would permit not only the comprehensive description of separate molecules but also their general systematization and comparison in various series based on modern information theory and the theory of molecular graphs.

The above-noted structural physical and chemical heterogeneity (complexity, diversity, or inhomogeneity) of a molecule reflect the disorder in the arrangement of its atoms. In the statistical thermodynamics, disorder is usually characterized by the Boltzmann entropy $S = -k\ln W$; however, in our case, this value cannot be used if for no other reason than that it refers only to statistically big ensembles. This limitation can be overcome by switching over to the information entropy introduced by Shannon.[10]

$$H(x_i) = -\sum_{i=1}^{N} p_i(x_i)\log_2 p_i(x_i),\ \sum p_i = 1, \qquad (1)$$

where p_i is the probability of the outcome x_i of the experiment, which should be calculated in each particular case. For example, if only the chemical composition of a molecule is taken into account, the calculation by Eq. (1) gives the corresponding information index of the chemical composition $H = I_{cc}$; and with allowance for only the topology of a molecule, which is reflected by its structural graph, this calculation affords the information topological index $H = I_{top}$, *etc.* These indices are expressed in terms of information units (bits) and are numerical measures of the heterogeneity or structural complexity of molecules. At present, they are widely used to establish correlations between structures and properties of molecules, mostly for organic and bioorganic molecules (see, for example, Refs 11—19). In the case of inorganic compounds and solids (crystals and glasses, including those with defects), which cannot be described by a molecular graph, the information topological approach has been much more poorly developed.

The information approach, which uses the Shannon entropy of information, is probabilistic. At present, nonprobabilistic approaches to the determination of information are also being developed. One of these is the Kolmogorov approach based on the idea of the epsilon-entropy (combinatorial approach) and on the algorithmic determination of the quantity of information,[20,21] the approach of Ingarden and Urbanek,[22] *etc.* In this paper, we also describe a nonprobabilistic approach in which the degree of order/disorder is evaluated by a procedure based only on geometrical considerations.[23,24] The application of this method to molecules is considered in detail below (in particular, from the standpoint of information theory); for this purpose, a new notion, *viz.*, geometrical molecular index, is introduced.

We do not deal with the condensed state of matter; however, all the results obtained here can be extended to it by treating, for example, a solid as a set of molecules "fastened" to the equilibrium positions. With certain assumptions, this approach can also be applied to a macroscopic gas or to a liquid, but it is necessary to take

into account the fact that in these cases, the analyzed state is merely an instantaneous picture of the molecules in space. Translational motions of atoms such as diffusion, which are typical of gases and liquids and also take place in real solids, or atomic motions with large amplitudes (characteristic of glasses)[25] and, hence, the corresponding types of order are not taken into account in this approach.

Cyclic molecules

Regular polygons

The index of order is calculated from the formula[23,24]

$$Q = \frac{3n - P}{3n} = 1 - \frac{P}{3n}, \qquad (2)$$

where P is the number of independent coordinates (positional parameters), which fix a group consisting of n points in an arbitrary system of coordinates. We chose the Cartesian coordinates. For example, for $n = 1$ (a separate point), $P = 3$ and $Q = 0$; for a set of n randomly arranged points, $P = 3n$ and $Q = 0$; for a line of a specified length (one stable spatial link is introduced for two points, and thus the P value decreases from $6 = 2 \cdot 3$ for disconnected points to $6 - 1 = 5$ for connected points), $Q = 0.167$; for an isosceles triangle, $Q = 0.111$, *etc.* In this study, points denote atoms, groups of points stand for molecules consisting of n atoms, and the spatial links correspond to chemical bonds in a molecule. Let us begin with the simplest case of regular polygons and molecules represented by them.

Along with the Q values calculated from Eq. (2), Table 1 presents the Schoenflies point groups of symmetry and common position multiplicities (in parentheses) as well as specific indices of order $q = Q/n$. We could not find exact molecular analogs of regular polygons except molecules like S_3 with the D_{3h} symmetry existing in the gas phase. The other examples correspond to a carbon framework or a metallic framework of a cluster without ligands. The order in these molecules is actually much higher than in the framework, since, if ligands are taken into account, the n value increases. Triangles, tetragons, and pentagons are frequently found as faces in polyhedral molecules reflected by regular polyhedra.

The order indices Q and the types of symmetry (common position multiplicities) for regular polygons, the limiting case of which is a circle with $n = \infty$ and $Q = 1$, vary in parallel but not linearly (see Table 1).

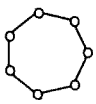
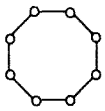
Irregular polygons

Molecules represented by irregular polygons are fairly abundant.[26,27] One should distinguish actually irregular polygons in which all the interatomic distances are different, and semiirregular polygons, which incorporate groups of equal interatomic distances corresponding to identical chemical bonds.

Let us consider the case of irregular polygons in relation to the S_9 molecule, which is of interest because it is formed not only in the gas phase but also in some biological objects and, in addition, it exists in the molecular-crystalline state.[28] Four isomers of this molecule are known; three of them have symmetry of the C_s point group (a boat, a cage, and a chair) and the fourth isomer has C_2 symmetry. All the isomers are characterized by equal numbers of symmetry operations ($M = 2$) and identical order indices $Q = 1 - 12/(3 \cdot 9) = 0.567$. Note that the order in the S_9 molecule proves to be much lower than in the corresponding regular polygon with D_{9h} symmetry, for which $Q = 1 - 7/(3 \cdot 9) = 0.741$, which looks quite reasonable.

Numerous examples of homoatomic rings represented by irregular polygons can be given even for $n > 6$ (for example, the Se_8^{2-} ion), whereas for regular polygons, only rings with $n = 3$ or C-frameworks (for example, C_5H_5, C_7H_7, and C_8H_8) are known. Thus, transition to irregular polygons leads not only to a decrease in the symmetry and order but also, as a consequence, to an increase in the probability of the occurrence of polyatomic molecules.

Table 1. Symmetry and order indices in regular polygons

Figure	n	Symmetry	Q	$q = \frac{Q}{n}$	Examples of molecules (ions)
	3	$D_{3h}(12)$	0.222	0.074	S_3
	4	$D_{4h}(16)$	0.417	0.104	C_4H_8 (C-framework)
	5	$D_{5h}(20)$	0.533	0.107	C_5H_5 (C-framework)
	6	$D_{6h}(24)$	0.611	0.102	C_6H_6 (C-framework)
	7	$D_{7h}(28)$	0.667	0.095	C_7H_7 (C-framework)
	8	$D_{8h}(32)$	0.708	0.089	C_8H_8 (C-framework)
	∞	$D_{\infty h}(\infty)$	1.000	0.000	—

Chain molecules

The world of chain molecules is enormous and embraces all branches of chemistry. There exist low-molecular-weight monomeric and oligomeric molecules incorporating relatively small numbers of atoms and macromolecules with $n \approx 10^3$—10^5, as well as giant polymeric macromolecules containing up to 10^6 atoms, *viz.*, $-[CH_2]_n-$ type organic molecules and $-[S(\text{or Se})]_n-$ or $-[PNCl_2]_n-$ type inorganic molecules.

Let us consider chains of the $-[S]_n-$ type. According to experimental data, in a polymer chain in the amorphous state, interatomic distances R are fixed parameters, whereas the bond angles and torsion angles can vary. For comparison, we shall consider the dependence of the order index on the number of atoms in the chain for three model cases (Fig. 1):

(1) constant interatomic distances, bond angles, and torsion angles

$$Q_1 = 1 - \frac{8}{3n}; \qquad (3)$$

(2) constant interatomic distances and bond angles and varying torsion angles

$$Q_2 = 1 - \frac{n+5}{3n} = \frac{2}{3} - \frac{5}{3n}; \qquad (4)$$

(3) constant interatomic distances and varying bond and torsion angles

$$Q_3 = 1 - \frac{2n+2}{3n} = \frac{1}{3} - \frac{2}{3n}. \qquad (5)$$

Each set of the structural degrees of freedom (SDF) corresponding to Eqs. (3)—(5) is described by its own curve and by its own limit of the Q value, the addition of a further SDF decreasing $\lim Q$ by one third

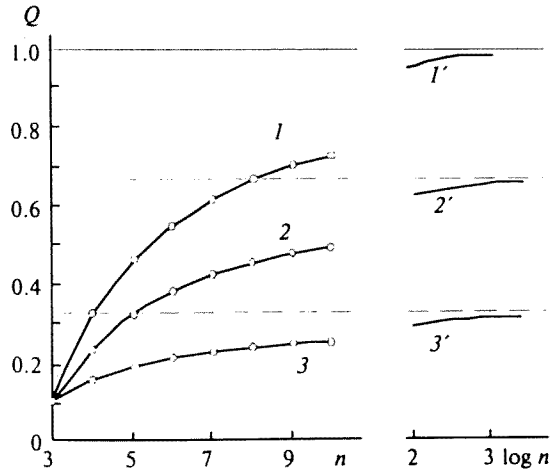


Fig. 1. Dependence of the order index on the number of atoms in chain molecules. The following parameters in the chains are constant: interatomic distances and bond and torsion angles (*1*); interatomic distances and bond angles (*2*); interatomic distances (*3*).

(1→0.667→0.333) (see Fig. 1, curves *1*—*3*). In the hypothetical case where all three SDF are realized, *i.e.*, where all angles and distances are variable, there is no order ($Q = 0$), since the number of independent coordinates (P) in Eq. (2) is $3n$.

In the second and third cases (Q_2 and Q_3), the symmetry of the chain molecule as a whole is C_1 (the molecule is asymmetrical), whereas an elementary unit S_3 (angle) has symmetry of the C_{2v} point group. Thus, the absence of symmetry in polymeric molecules as a whole does not mean that there is no order, since $Q > 0$. In this case, spontaneous ordering occurs, depending on the number of structural degrees of freedom in a polymeric chain of length n. Giant polymeric biomolecules can coil into globules or, conversely, can uncoil into polymeric bundle filaments as occurs in muscles. Deoxyribonucleic acid is a double-stranded helix in which the order is higher than in a single helix, due to the appearance of joints formed *via* hydrogen bonds.

Polyhedral molecules

Regular polyhedra

Five regular polyhedra called Platonic solids are known: tetrahedron, cube (hexahedron), octahedron, dodecahedron, and icosahedron. These polyhedra form the basis for many molecules (examples of these molecules are cited in Table 2). The number of atoms in the molecules can exceed the number of vertices of the corresponding regular polyhedron, due to the presence of a central atom (for example, in the terahedral $SiCl_4$ or octahedral SF_6) or due to replication (for example, B_4Cl_4 is built in such a way that the boron atoms form an inner tetrahedron, while the chlorine atoms constitute the outer tetrahedron).

Table 2. Symmetry and order indices in regular polyhedra and figures based on them

Polyhedron	n	Symmetry	P	Q	$q = \frac{Q}{n}$	Examples of molecules (ions)
Tetrahedron	4	$T_d(24)$	7	0.417	0.104	P_4, As_4
	5	$T_d(24)$	7	0.533	0.107	CH_4, $SiCl_4$
	8	$T_d(24)$	9	0.625	0.078	*closo*-B_4Cl_4
Octahedron	6	$O_h(48)$	7	0.617	0.103	—
	7	$O_h(48)$	7	0.667	0.095	SF_6, PCl_6^-
	12	$O_h(48)$	8	0.778	0.065	*closo*-$[B_6H_6]^{2-}$
Cube	8	$O_h(48)$	7	0.708	0.089	Cubane C_8H_8 (C-framework)
	16	$O_h(48)$	8	0.833	0.052	Cubane C_8H_8
Icosahedron	12	$I_h(120)$	7	0.805	0.067	$B_{12}H_{12}^{2-}$ (B-framework)
Dodecahedron	20	$I_h(120)$	7	0.883	0.044	Dodecahedrane (C-framework)
	40	$I_h(120)$	8	0.933	0.023	Dodecahedrane $C_{20}H_{20}$

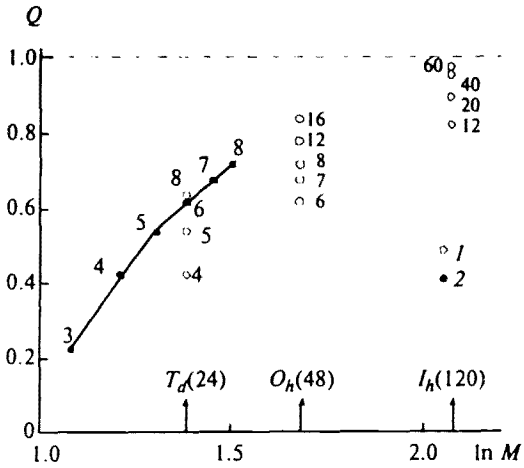


Fig. 2. Dependence of the order index on symmetry (M) and on the number of atoms (numerals near the points) in molecules reflected by regular polyhedra and figures based on them (*1*) and by regular polygons with the D_{nh} symmetry (*2*).

In this case, unlike regular polygons, the order index is a function of both symmetry and the number of atoms in the molecule (see Table 2). For example, in tetrahedral molecules with the T_d symmetry and $M = 24$, the Q value increases from 0.417 to 0.625 as n increases from 4 to 8.

To make this more clear, we present the data of Table 2 as a plot, which is shown in Fig. 2. It can be seen that when $n \geq 4$ and $M \geq 24$, the order index Q is much more sensitive to an increase in the number of atoms in a molecule (especially at small n) than to an increase in symmetry. For comparison, Fig. 2 also presents the $Q = f[M(n)]$ dependence for regular polygons.

Semiregular polyhedral molecules

Faces of semiregular polyhedra are regular polygons of various types. Numerous types of these polyhedra exist, *viz.*, prisms and antiprisms, pyramids and bipyramids (examples are presented in Table 3). One should bear in mind that the coordination polyhedron of a real molecule can deviate somewhat from the prototype polyhedron; for example, faces in the polyhedron formed by the NH_3 molecule are isosceles rather than equilateral triangles.

It can be seen from Table 3 that the main regularities are similar to those observed in the previous case, *i.e.*, $Q = f(M, n)$ and Q increases as M and n increase. A natural distinction is that, at the same n, the symmetry (common position multiplicity) and the order index are lower for molecules corresponding to semiregular polyhedra than for those represented by regular polyhedra (*cf.* Tables 2 and 3).

Irregular polyhedra

Irregular polyhedra are classed as either asymmetrical (possessing no symmetry) polyhedra, for which the symmetry of the C_1 point group corresponds to the common position multiplicity $M = 1$, or dissymmetrical polyhedra, which often posses low symmetry. The asymmetrical figures and many of the dissymmetrical figures are chiral, *i.e.*, their limit figures are a rotating cylinder with $\infty/2$ symmetry and a rotating cone with ∞ symmetry. Numerous groups of the type C_n (n = 1, 2, 3, 4, 5...), D_n (n = 1, 2, 3, 5...), and O are chiral.

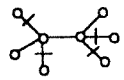
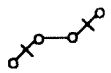
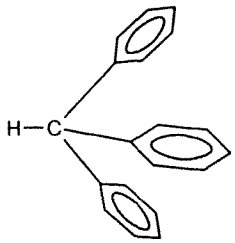
Intuition suggests that asymmetrical figures with C_1 symmetry should have $Q = 0$. However, this statement is true only for simplest figures, formed by four of five atoms arranged arbitrarily (Table 4). Examples are also provided by some triatomic molecules (NOF, O_3) that have $C_1(1)$ symmetry and $Q = 0$, due to lone electron pairs regarded as pseudo-atoms. When the number of atoms increases, identical interatomic distances or identical atoms usually arise; an example is provided by the C_2H_4FCl molecule (see Table 4) in which $Q > 0$ due to the fact that $P < 3n$. This result demonstrates once again that in terms of the model considered, there is no unambiguous relationship between symmetry and order. In the present case, order (Q) is a broader concept incorporating all the features of the spatial arrangement of points (atoms), including those, not reflected by symmetry.

Figures with higher symmetries (C_n and D_n) with $n \geq 2$ are characterized by greater Q values. For example, for the S_2Cl_2 molecule (C_2 symmetry), $Q = 0.250$, whereas for trimethylbenzene (C_3 symmetry), the value $Q =$

Table 3. Order indices in semiregular polyhedral figures

Polyhedron	n	Symmetry	P	Q	$q = \frac{Q}{n}$	Examples of molecules (ions)
Trigonal pyramid	4	$C_{3v}(6)$	8	0.333	0.083	NH_3
	5	$C_{3v}(6)$	9	0.400	0.080	POF_3
Tetragonal pyramid	5	$C_{4v}(8)$	8	0.467	0.093	*closo*-$[B_5H_5]^-$ (B-framework)
	10	$C_{4v}(8)$	11	0.633	0.063	*closo*-$[B_5H_5]^-$
Trigonal bipyramid	5	$D_{3h}(8)$	8	0.467	0.093	—
Trigonal bipyramid with the central atom	6	$D_{3h}(12)$	8	0.445	0.074	PF_5
Pentagonal bipyramid	14	$D_{5h}(20)$	10	0.762	0.054	*closo*-$[B_7H_7]^{2-}$
Square prism	10	$D_{4h}(16)$	9	0.700	0.070	$[Re_2Cl_8]^{2-}$
Prism	18	$D_{3h}(12)$	12	0.778	0.043	$[Re_6Cl_{12}]^-$
	20	$D_{2h}(8)$	15	0.750	0.038	$[Tc_8Cl_{12}]^-$

Table 4. Asymmetrical and dissymmetrical figures

Figure*	Dimensionality	n	Symmetry	P	Q	$q = \frac{Q}{n}$	Molecule
	3D	4	$C_1(1)$	12	0.000	0.000	—
	3D	5	$C_1(1)$	15	0.000	0.000	CHFClBr
	3D	8	$C_1(1)$	21	0.125	0.016	C_2H_4FCl
	2D	4	$C_2(2)$	9	0.250	0.063	S_2Cl_2
	3D	35	$C_3(3)$	14	0.867	0.025	$CHPh_3$

* Equal distances between the points of the figures are marked by dashes.

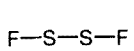
0.867 is commensurable with those observed for regular polyhedral molecules (see Fig. 2).

Thus, asymmetrical figures (C_1) can possess certain degrees of order ($Q > 0$), and in the case of polyatomic chiral dissymmetrical figures, the order index may be fairly high. This nontrivial result provides the possibility of dividing asymmetrical figures (molecules) into disordered asymmetrical figures (C_1, $Q = 0$) and ordered asymmetrical figures (C_1, $Q > 0$), and dissymmetrical molecules can thus be classified not only in terms of their symmetries but also according to the degree of order realized in the arrangement of their points (atoms).

Order in isomers

According to the classification of isomerism (see, for example, Ref. 27), the most pronounced structural changes occur in the case of constitutional isomers, in which not only interatomic distances but also the mutual arrangement of chemical bonds vary. Let us consider, for example, the isomers of S_2F_2.

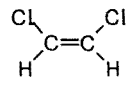
$C_{2v}(4)$, $P = 9$, $Q = 0.250$



$D_{\infty h}(\infty)$, $P = 7$, $Q = 0.417$

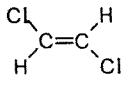
As should be expected, the index of order is greater for the isomer characterized by the higher symmetry.

The structural differences between geometrical isomers, for example, between *cis*- and *trans*-1,2-dichloroethylene are less pronounced.



cis

$C_{2v}(4)$, $P = 11$, $Q = 0.389$



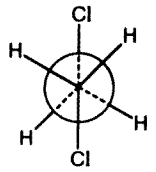
trans

$C_{2h}(4)$, $P = 11$, $Q = 0.389$

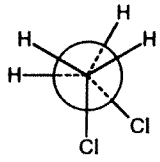
These molecules are characterized by equal order indices and equal common position multiplicities, although the symmetry point groups are dissimilar. The invariability of the Q value means that in the present case, the index of order reflects the number of symmetry operations rather than the symmetry itself.

Geometrical isomers can exist as individual forms (they can be physically separated), since the barrier to the internal rearrangement is relatively high. In the case where this barrier is low, we deal with the conformational isomerism. For example, on passing from dichloroethylene incorporating a C=C double bond to dichloroethane with a C—C single bond, the possibility of rotation around the C—C axis appears. Consider the Newman projections.

The most symmetrical staggered configuration of the molecule with the C_{2h} symmetry, for which $Q = 0.500$, is presented at the left. The lowest C_1 symmetry and $Q = 0.416$ correspond to an arbitrary arrangement of atoms during rotation, *i.e.*, the order substantially de-



$C_{2h}(4)$, $P = 12$, $Q = 0.500$



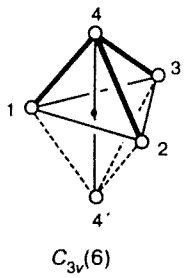
$C_1(1)$, $P = 14$, $Q = 0.41$

creases during rotation. Thus, the appearance of the new (rotational) degree of freedom leads to a decrease in the order, similar to what we have observed previously upon the addition of structural degrees of freedom in chain molecules (see Fig. 1).

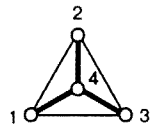
Order in nonrigid molecules

The above example of internal rotation with a low energy barrier, when the isomers cannot be physically separated, indicates that the method that we suggest makes it possible to describe dynamic structures. Below we analyze some types of nonrigid molecules, in which inversions, librations, and permutations can occur along with rotation.

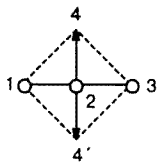
Inversions. Let us consider the simplest tetraatomic molecule with the C_{3v} symmetry like NH_3 as an example. The base of this molecule is a regular triangle consisting of atoms 1, 2, and 3 (**A**).



$C_{3v}(6)$



$D_{3h}(12)$



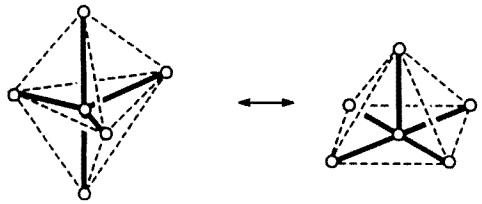
$C_{3v}(6)$

A: inversion **B:** libration

During the inversion, the pyramid formed by atoms 1, 2, 3, and 4 turns inside out like an umbrella; as this takes place, atom 4 occupies position 4′ and, in the general case, it moves along the line connecting points 4 and 4′ and atoms 1, 2, and 3 remain at their positions (are fixed). For the initial position, $P = 8$ and $Q = 0.333$. When atom 4 moves, its momentary fixation requires one independent coordinate; therefore, P and the index of order are retained. However, the above consideration is simplified, since atoms 1, 2, 3 move apart fully symmetrically as atom 4 passes through the base of the pyramid, and the 1—2, 2—3, and 1—3 distances increase, although the regular triangle forming the base of the pyramid retains its D_{3h} symmetry. Thus, during the inversion, coordinates of all of the atoms vary, and the order index should decrease. If the inversion occurs in a gas or in a liquid, the decrease in the Q value is not too pronounced, *i.e.*, the simplified description is also suitable. However, if a nonrigid molecule undergoes an inversion in a solid, the symmetry substantially decreases due to the effect of the environment. As a consequence, the order index decreases considerably, perhaps down to $Q = 0$.

Librations. Let us consider librations (oscillations of atoms accompanied by the deviation of one of them from the plane) in relation to a planar tetraatomic molecule with D_{3h} symmetry (**B**). Atoms 1, 2, and 3 retain their positions, whereas atom 4 executes inversion oscillations. The D_{3h} symmetry is not retained, since a C_{3v} pyramid arises, and one more independent coordinate is now needed for the momentary fixation of the position of atom 4. In this case, $Q = 1 - 8/(3 \cdot 4) = 0.333$, which is lower than the initial $Q = 0.417$.

Permutation isomerism. Let us consider the permutation isomerism using the Berry rearrangement in the PF_5 molecule as an example; the two extreme positions are characterized by D_{3h} symmetry ($P = 8$, $Q = 0.556$) and by C_{4v} symmetry ($P = 8$, $Q = 0.556$). If a polyhedron incorporates a regular polygon in the equatorial plane, which coincides with the plane of the mirror image m, the common position multiplicity increases as the number of vertices increases. Conversely, for all regular polygons, irrespective of the number of vertices, the number of independent parameters P is 7.



$D_{3h}(12)$ $C_{4v}(8)$

The transition from one state to another does not lead to a variation of the order, but the order in the intermediate states (in instantaneous pictures) is lower.

In fact, if we assume that only the phosphorus atom retains its position, whereas all the five fluorine atoms move, and that three coordinates are required to describe momentary positions of each of them, then $P = 3 \cdot 5 = 15$ and $Q = 1 - 15/(3 \cdot 6) = 0.167$. If the phosphorus atom also changes its position, then $P = 3 \cdot 6 = 3n$ and $Q = 0$.

However, one should take into account that the symmetry is determined by more than just the the common position multiplicity M. For example, in the scheme presented above, the $C_{4v}(8)$ state is more symmetrical from the crystallographic standpoint than the

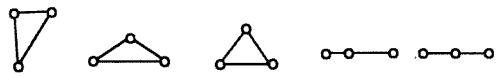
$D_{3h}(12)$ state, since the former possesses a fourfold axis.

Thus, the index of order Q is very sensitive to the static and dynamic isomerism, and correlates well with the symmetry: Q decreases as the symmetry decreases; however, the order remains unchanged if the common position multiplicity does not vary. The appearance of new sorts of motion in nonrigid molecules (librations, inversions, and permutations) is accompanied by a decrease in the symmetry and by a decrease in the order index.

Results and Discussion

1. Order and symmetry. Since, as shown above, in the general case, order varies as a function of both symmetry and the number of atoms in a molecule, it would be appropriate to consider the dependences of Q on each of these parameters separately, with the other parameter remaining constant. Now we discuss the dependence of Q on M at n = const for the simplest tri-, tetra-, and pentaatomic figures.

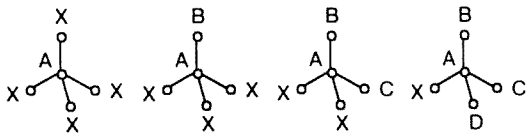
In triatomic molecules, points (atoms) can be arranged either as a triangle (irregular, isosceles, or equilateral) or on a line (with equal of different arms).



Symmetry	$C_s(1)$	$C_{2v}(4)$	$D_{3h}(12)$	$C_{\infty v}(\infty)$	$D_{\infty h}(\infty)$
P	9	8	7	7	6
Q	0	0.111	0.222	0.222	0.333

The limiting value $\lim Q(M\to\infty)$ is equal to 0.333. The data presented above indicate that there exist figures with different types of symmetry [$D_{3h}(12)$ or $C_{\infty v}(\infty)$] but identical order indices Q. Let us call these figures *isoorders*.

Let us consider a tetrahedral molecule with central atom A in which the initially identical ligands are successively replaced as an example of pentaatomic molecules.



Symmetry	$T_d(24)$	$C_{3v}(6)$	$C_{2v}(4)$	$C_1(1)$
P	7	9	11	15
Q	0.533	0.400	0.267	0

The index of order regularly increases in the $C_1\to T_d$ series with an increase in the common position multiplicity, and reaches the limiting value Q = 0.533 in the case of a regular tetrahedron. These results can also serve as an illustration of the effect of chemical composition on the Q and $P/3n$ values.

In the case of tetraatomic molecules, 74 figures can be distinguished. We shall consider only the eleven most typical forms (Table 5, Fig. 3). Among them, molecules with different types of symmetry but identical order indices can be found; for example Q = 0.417 (the limiting value for all dimensionalities) for $D_{\infty h}$, D_{3h}, D_{4h}, and T_d. Thus, the isoorders, which we discovered above as an individual case of triatomic molecules, can also be found in more complex molecules.

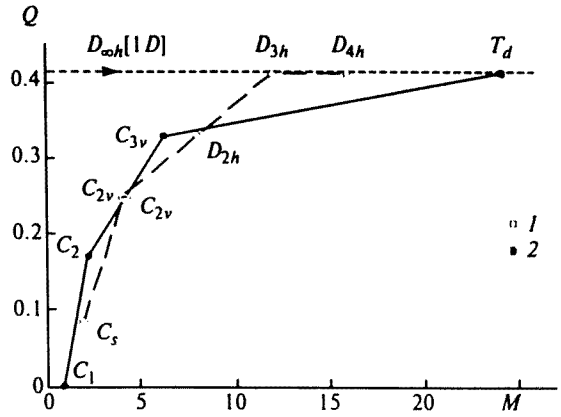


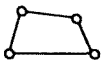
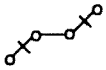
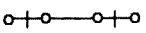
Fig. 3. Dependence of the order index on symmetry for tetraatomic molecules reflected by 2D-polygons (*1*) and 3D-polyhedra (*2*). The horizontal line corresponds to an 1D-structure at $M\to\infty$.

In the case of isoorders, the index of order is insensitive to variations in the molecular symmetry. The opposite case, *i.e.*, the case where figures with identical types of symmetry possess different order indices (we shall call them *isosymmetrics*), is equally interesting. In the example under consideration, this is a polyhedron with $C_2(2)$ symmetry and Q = 0.167 and a polygon with $C_s(2)$ symmetry and Q = 0.083. This implies that in some cases, Q can be more sensitive to the molecular structure than the symmetry (the common position multiplicity).

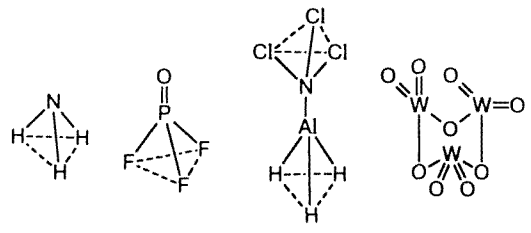
2. Order and the number of atoms in the molecule. The dependences of Q on n have already been discussed in the consideration of chain molecules and molecules having shapes of regular polygons and polyhedra. The symmetry of the chain molecules in the noncrystalline state, *i.e.*, in the presence of structural degrees of freedom, corresponds to the C_1 point group, and the Q value regularly increases as the number of atoms n increases tending to the limiting values $\lim Q$ = 1.000 (no SDF), $\lim Q$ = 0.667 (one SDF), or $\lim Q$ = 0.333 (two SDF) (see Fig. 1).

Figure 4 presents the plot of Q against n at a constant symmetry. It can be seen that for figures (molecules) based on regular polyhedra, the $Q = f(n)$ dependence is essentially the same and varies only slightly upon going from one symmetry point group to another (T_d, O_h, I_h). The situation also does not virtually change when we consider irregular polyhedral molecules, for example, those with the $C_{3v}(6)$ symmetry.

Table 5. Order index and symmetry in planar (2*D*) and three-dimensional (3*D*) tetraatomic molecules

Figure*	Dimensionality	Symmetry	*P*	*Q*	$q = \frac{Q}{n}$
	3*D*	$C_1(1)$	12	0.0	0
	2*D*	$C_s(2)$	11	0.083	0.021
	3*D*	$C_2(2)$	10	0.167	0.042
	2*D*	$C_{2v}(4)$	9	0.250	0.063
	3*D*	$C_{2v}(4)$	9	0.250	0.063
	3*D*	$C_{3v}(6)$	8	0.333	0.083
	2*D*	$D_{2h}(8)$	8	0.333	0.083
	2*D*	$D_{3h}(12)$	7	0.417	0.104
	2*D*	$D_{4h}(16)$	7	0.417	0.104
	3*D*	$T_d(24)$	7	0.417	0.104
	1*D*	$D_{\infty h}(\infty)$	7	0.417	0.104

* Equal distances between the points of the figures are marked by dashes.



n(*P*)	4(8)	5(9)	8(11)	12(12)
Q	0.333	0.400	0.542	0.667
q	0.083	0.080	0.068	0.056

It can be clearly seen from Fig. 4 that the *Q* value increases as the symmetry increases, while the number

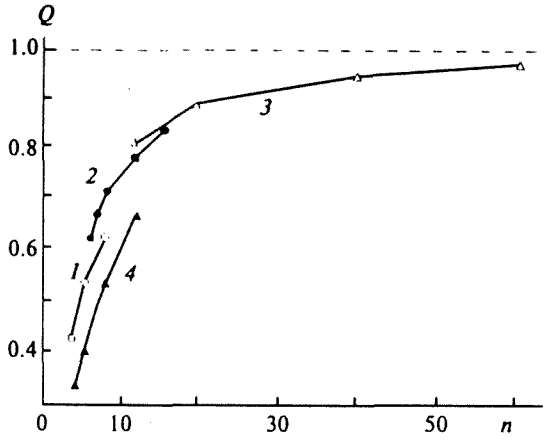


Fig. 4. Dependence of the order index on the number of atoms in a molecule for regular polyhedra and figures based on them with the symmetry T_d (*1*), O_h (*2*), and I_h (*3*), and also for irregular polyhedra with the C_{3v} symmetry (*4*).

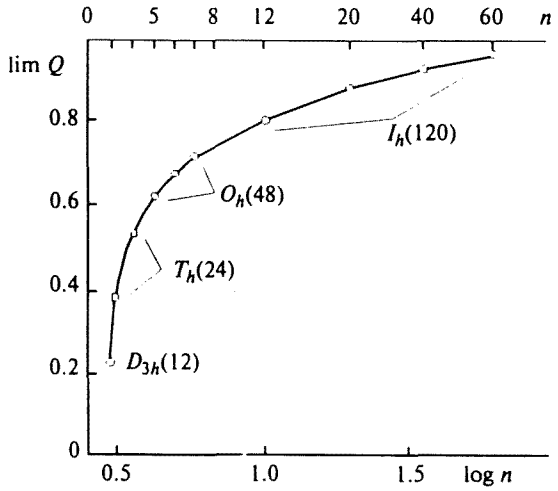
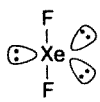


Fig. 5. Limiting order indices and symmetry as a function of the number of atoms in a molecule.

of atoms *n* is constant. The highest symmetry and the limiting index of order lim*Q* corresponding to it can be found for each *n*. For example, the types of symmetry realized for $n = 12$ range from $C_{3v}(6)$ with $Q = 0.667$ to $I_h(120)$ with $Q_{lim} = 0.805$. The results of this analysis carried out for some *n* values located in the range from 3 to 60 are presented in Fig. 5. It can be seen that in polyatomic nonchain molecules with $n > 20$, the order indices are high ($Q \geq 0.9$). As *n* increases the order asymptotically approaches 1.0. The order indices for polyatomic clusters (for example, for fullerenes with $n \geq 60$) are very high, although they are still lower than the values observed in crystals ($Q \approx 1.0$) possessing typical long-range order.

3. Order and lone electron pairs. The stereochemical role of lone electron pairs is well known. The Gillespie model of repulsion[29] in its current formulation[27] makes it possible to describe a great number of molecular

structures. We shall consider the variation of the structure of molecule AX_5 during the successive substitution of ligands X by lone electron pairs (E) within the framework of these views.

Type	AX_5	AX_4E	AX_3E_2	AX_2E_3
Example				
Symmetry	$D_{3h}(12)$	$C_{2v}(4)$	$C_{2v}(4)$	$D_{\infty h}(\infty)$
P	8	9	8	6
Q	0.556	0.400	0.333	0.333

The molecules are shown schematically without precise geometrical relations between the bond lengths and the angles between bonds. Actually, the following variations of the structure of the molecule occur.

The initial structure AX_5 is an ideal trigonal bipyramid. When one of the F atoms is replaced by E, the axial bonds become longer than equatorial bonds, and the angles deviate from their ideal magnitudes. For example, the length of the axial bond in SF_4 is 1.65 Å, the equatorial bond length is 1.55 Å, and the axial and equatorial angles are 173° and 102°, respectively.[25] The symmetry decreases from $D_{3h}(12)$ to $C_{2v}(4)$.

The further substitution $AX_4E \rightarrow AX_3E_2$ does not lead to a change in the symmetry, but the order index Q decreases. The lone electron pair E affects the Q value in two ways. First, as the number of E increases, the n value formally decreases, and, consequently, all other factors being equal, Q decreases. Second, the appearance of E can cause a decrease in the symmetry (the common position multiplicity) of the atomic group and a further decrease in the Q value or, conversely, an increase in the symmetry and in Q. In the example given, the Q value was affected by the first factor, because the symmetry did not change.

Further substitution $AX_3E_2 \rightarrow AX_2E_3$ leads to a highly symmetrical figure with an infinity-fold axis of symmetry. The increase in the symmetry results in an increase in Q, despite the fact that one more atom was lost. Note that in this series, we once again come across isosymmetrics (AX_4E and AX_3E_2), which arise due to the effect of the lone electron pair on the order in the molecule.

4. Scale of order. Thus, for each molecule in which the spatial arrangement of atoms is known, one can determine not only the symmetry (the point group and the common position multiplicity) but also the index of order Q. We have considered numerous examples in which Q varied from 0 to 1.0, *i.e.*, molecules occupied the whole scale of the order indices.

As shown above, symmetry and order supplement each other. Two special cases should be distinguished: isoorders, *i.e.*, degenerate molecules possessing identical order indices but different types of symmetry (for example, XeF_2 and NH_3 with $Q = 0.333$), and isosymmetrics, *i.e.*, molecules with identical types of symmetry but different order indices (for example, NH_3 and W_3O_9 with the C_{3v} point group).

Molecules in which both the order and the symmetry are identical can be classified as structurally identical or structurally similar molecules. Molecules of the former group incorporate equal numbers of atoms (for example, O_3 and FNO molecules, see Table 4, $Q = 0$, C_1, and $n = 3$), while in the latter case, the numbers of atoms are dissimilar (for example, O_3 and C_2H_4ClF, $n = 3$ and $n = 8$, respectively).

Thus, the introduction of the new structural characteristic, the index of order Q, provides the basis for a new classification of molecules and permits comparison of structural features of various classes of chemical compounds within a unified scale.

5. Geometrical information index. We regarded the Q value in Eq. (2) as an index characterizing the order in the arrangement of atoms in a molecule, while the term $P/3n$ was correspondingly regarded as an index characterizing disorder. In the present section we analyze the geometrical meaning (and, to some extent, the physical meaning) of Eq. (2) and of the parameters incorporated in it and demonstrate that the term $P/3n$ corresponds to the entropy of information, and therefore, it can be regarded as a geometrical information index, and the term Q corresponds to negentropy or excess information.

It was shown above that Q (or $1 - Q$) is the function of several parameters

$$Q = f(M, n, D, \mathrm{SDF}),$$

where M is the symmetry point group of the molecule expressed in terms of the common position multiplicity, SDF is the number of structural degrees of freedom, n is the number of atoms, which reflects the complexity of the molecule (whether it is a simple di- or triatomic molecule or a complex molecule), and D is the dimensionality of the molecule, which serves as a measure of its physical inhomogeneity, since the dimensionality is related to the anisotropy of the molecule. Anisotropy is especially significant for $1D$ (chain) and $2D$ (layered) molecules. The dimensionality plays a very important role in the theory of spatial order and disorder when it is applied to the formation of crystals of polymers; this was reflected in the relevant rules, for example, the rule that the spontaneous crystal order cannot exist in $1D$- and $2D$-systems.[5]

Let us analyze the independent variables of the $Q = f(M, n, D, \mathrm{SDF})$ function from the viewpoint of the information theory. According to the definition given by Ashbey,[30] information is a measure of the diversity in the system, and according to the Glushkov definition,[31] it is a measure of inhomogeneity and nonuniformity of the system. We shall consider the independent variables of the function in terms of these definitions.

In the framework of the approach under consideration, symmetry can be considered to be a measure of structural diversity or structural inhomogeneity, since

symmetry operations (symmetry axes and planes and inversion center) make it possible to divide the set of atoms constituting a molecule into automorphic groups (AMG). We shall make it clear using the CH_2F_2 molecule with the C_{2v} symmetry as an example. This molecule contains three AMG, two of which, [H(1) and H(2)] and [F(1) and F(2)], are connected by a twofold axis, while the third (the C atom) is located on this axis. In this case, the AMG are distributed over symmetry as {2,2,1}. Pay attention to the fact that the atoms in the groups are identical (homoatomic), whereas the groups themselves are heteroatomic. Thus, symmetry reflects the structural nonuniformity and indirectly reflects the chemical diversity in the molecule.

Our analysis indicates that the Q (or $1 - Q$) value characterizes a molecule fairly comprehensively and qualitatively corresponds to the definition of information as a measure of diversity, nonuniformity, and inhomogeneity. In order to verify this tentative inference, we shall consider the relationship between the parameter $P/3n$, which in the present case characterizes disorder, and the information symmetry index I_{sym} introduced by Bonchev (see Refs. 13—16); we chose this index among the other information indices because it is best suited for our geometrical approach. The index I_{sym} is calculated taking into account the distribution of atoms equivalent in symmetry, *i.e.*, atoms that exchange places upon symmetry operations. Each symmetry operation involves a particular number of atoms in a molecule as well as particular groups of atoms, *viz.*, AMG.

Let us consider the foregoing in relation to the $[PtCl_4]^{2-}$ ion with the D_{4h} symmetry. It has a fourfold axis, which makes it possible to distinguish an AMG [Cl(1)Cl(2)Cl(3)Cl(4)] consisting of atoms joined by the fourfold axis, and an atom (Pt), through which this axis passes; thus, the five atoms are characterized by the {4,1} symmetry distribution. The I^a_{sym} value for an averaged atom is calculated from formula (1); for the ion considered, it is equal to

$$I^a_{sym} = -\frac{4}{5}\log_2\frac{4}{5} - \frac{1}{5}\log_2\frac{1}{5} = 0.722 \text{ (bit at.}^{-1}\text{)},$$

and for the whole molecule, $I^m_{sym} = 0.722 \cdot 5 = 3.611$ (bit molec.$^{-1}$). Note that the I_{sym} values characterizing the CH_4 and $SiCl_4$ molecules are the same, although the $T_d(24)$ symmetry of these molecules differs from $D_{4h}(16)$; in other words, in some cases, I_{sym} is insensitive to its main criterion and is degenerate. An analysis carried out previously[13—16] showed that, in addition to the symmetry of the whole molecule, the types of symmetry of separate groups of atoms, for example, of NH_2 groups, *etc.*, should also be involved in the calculation (see below).

To elucidate the relations between $P/3n$ and I^m_{sym}, we shall first consider molecules consisting of $n = 5$ or 8 atoms (Table 6). Since the limiting magnitude of $P/3n$ is 1, in order to ensure better comparison of the values of $P/3n$ and I^m_{sym}, the symmetry index should also be presented in a normalized form. For this purpose, we divide the symmetry index by the factor $n\log_2 n$. The physical meaning of this factor is the maximum information that can be carried by an asymmetrical (completely chaotic) molecule: $I^m_{sym} = 11.610$ for $n = 5$, $I^m_{sym} = 24.00$ for $n = 8$, *etc.* The I^m_{sym} values normalized in this way are listed in Table 6. The correlation between $P/3n$ and I^m_{sym} is shown as a plot in Fig. 6.

Table 6. Information symmetry index and the $P/3n$ value in series of penta- and octaatomic molecules

Molecule	Symmetry	P	AMG	I^m_{sym}[a] /bit · molecule^{-1}	$P/3n$	$\frac{I^m_{sym}}{n\log_2 n}$
			Pentaatomic molecules			
$[PtCl_4]^{2-}$	$D_{4h}(16)$	7	{1,4}	3.611	0.467	0.311
CH_4	$T_d(24)$	7	{1,4}	3.611	0.467	0.311
MeCl	$C_{3v}(6)$	8	{1,3,1}	6.857	0.533	0.591
CH_2F_2	$C_{2v}(4)$	10	{1,2,2}	7.611	0.667	0.656
CH_2FBr	$C_s(2)$	12	{1,1,1,2}	9.611	0.800	0.828
CHFClBr	$C_1(1)$	15	{1,1,1,1,1}	11.611	1.000	1.000
			Octaatomic molecules			
Cube[b]	$O_h(48)$	7	{8}	0	0.292	0
C_2H_6	$D_{3h}(12)$	9	{6,2}	6.488	0.375	0.270
C_2H_6	$D_{3d}(12)$	9	{6,2}	6.488	0.375	0.270
Al_2Br_6	$D_{2h}(8)$	10	{2,4,2}	12.00	0.417	0.500
$C_2H_3Cl_3$	$C_3(3)$	12	{3,3,2}	12.49	0.500	0.520
$(CH_2Cl)_2$	$C_2(2)$	15	{2,2,2,2}	16.00	0.625	0.667
$(CHFCl)_2$	$C_i(2)$	16	{2,2,2,2}	16.00	0.667	0.667
EtCl	$C_1(1)$	20	{1,1,1,1,1,1,1,1}	24.00	0.833	1.000
EtCl	$C_1(1)$	20	{3,1,1,1,1,1}[c]	20.80	0.833	0.870

[a] Taken from the literature.[13] [b] The figure is presented for comparison. [c] With allowance for the local symmetry of the Me group (all the other distributions were obtained neglecting the local symmetry).

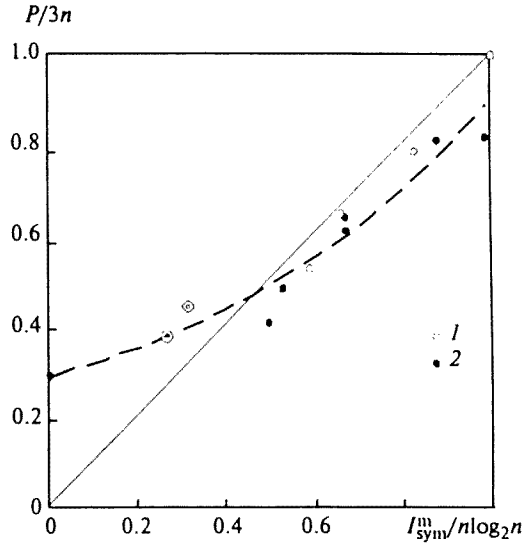
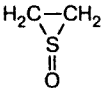
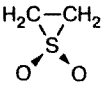
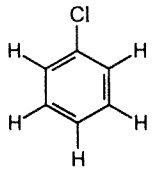


Fig. 6. Correlation of the $P/3n$ parameters and of the normalized symmetry indices in the case of penta- (*1*) and octaatomic molecules (*2*) of various types of symmetry.

Table 7. Information symmetry index and the $P/3n$ value in molecules with the identical C_{2v} symmetry

n	Molecule	Dimensionality	P	$P/3n$	AMG	I^a_{sym} [a] /bit at.$^{-1}$	$\log_2 n$	$\frac{I^a_{sym}}{\log_2 n}$
3		2*D*	8	0.889	{1,2}	0.92	1.59	0.58
4		2*D*	9	0.750	{2,2}	1.00	2.01	0.50
8		2*D*	12	0.50	{2,6}[a]	0.81	3.01	0.27
9		3*D*	12	0.444	{1,2,6}[a]	1.22	3.18	0.38
12		2*D*	10	0.278	{4,8}[a]	0.72	3.60	0.20

[a] With allowance for local symmetry (see the Text).

Now we compare molecules with identical types of symmetry but different numbers of atoms. The relevant data for molecules whose symmetry corresponds to the C_{2v} point group are given in Table 7 and in Fig. 7. In this case, we not only normalized the I^a_{sym} value but also calculated I^a_{sym} taking into account the local symmetries of separate groups of atoms incorporated in the molecule.[13–16] The local symmetry can differ substantially from the symmetry point group of the whole molecule. For example, a molecule with C_1 symmetry can contain CH_3, NH_2, and other groups, whose symmetry is higher than that of the molecule itself. For example, the symmetry of the CH_3 group is C_{3v}. The octaatomic molecule C_2H_5Cl (see Table 6) has C_1 symmetry, and, with allowance for the local C_{3v} symmetry of the CH_3 group, its distribution over symmetry is {3,1,1,1,1,1} instead of {1,1,1,1,1,1,1,1}, which would be the case if the local symmetry were not taken into account. The values calculated taking and not taking into account the local symmetry are markedly different. For example, if we neglect the local symmetry, molecules with C_1 symmetry will be characterized by different I^a_{sym} values; however, the normalized $I^a_{sym}/n\log_2 n$ values will be identical, irrespective of the number n (*cf.* CHFClBr and C_2H_5Cl molecules in Table 6). Thus, the I^a_{sym} values normalized by $\log_2 n$ but not taking into account the local symmetry are insensitive to variations of n. When the local symmetry is involved in the consideration, the degeneracy is removed, and the normalized $I^a_{sym}/\log_2 n$ value becomes a monotonic function of n, if we ignore the 3*D* figure with $n = 9$ (see Table 7 and Fig. 7). Note that the I^a_{sym} index for figures of regular shapes is equal to zero (is degenerate). In these figures, all the vertices (atoms) are identical; hence, $I^a_{sym} = \log(n/n) = 0$. An example of such a figure is provided by a cube (see Table 6, Fig. 6). Unlike I^a_{sym}, the $P/3n$ value is nondegenerate with respect to regular figures.

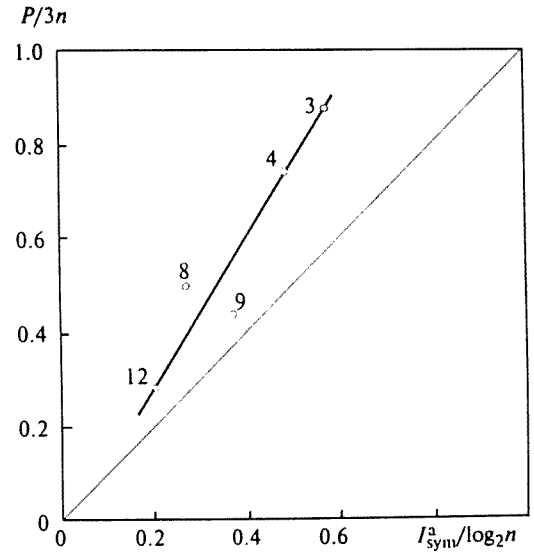


Fig. 7. Correlation of the $P/3n$ parameters and of the normalized symmetry indices (I^a_{sym} with allowance for the local symmetry) for molecules with the C_{2v} symmetry and with various numbers of atoms n (numerals near the points).

Analysis of the data presented in Tables 6 and 7 and in Figs. 6 and 7 indicates that the $P/3n$ and $I^{m}_{sym}/n\log_2 n$ values are not identical (as should be expected) but are correlated. Therefore, $P/3n$ may be regarded as a special information index, namely, geometrical molecular information index

$$I_{geom} = \frac{P}{3n}.$$

The analogy of this value with the symmetry index I_{sym} is due to the fact that both indices take into account the symmetry of the molecule: I_{sym} does it by virtue of the corresponding AMG, and $P/3n$ is based on the geometrical consideration of a particular figure/molecule. They differ not only in the way in which the symmetry is taken into account but also in the method of calculation, which is probabilistic (according to the Shannon formula) in the case of I_{sym} and nonprobabilistic (based on analytical geometry) in the case of $P/3n$. The I_{geom} index supplements the information indices known previously[11–19] and can be used to elucidate the composition—structure—property relationships, for example, to evaluate entropy, enthalpy, and reactivity, to predict the properties of molecules, to accomplish target-directed syntheses of molecules with specified properties, and to carry out systematization for computerization of studies, *etc.*

Note that a numerical correlation involving I^{m}_{sym} can be obtained only if this index is normalized by $n\log_2 n$, whereas the index I_{geom} is actually normalized by $3n$. This means that I_{sym} refers to the information space, whereas I_{geom} refers to the three-dimentional space of Cartesian coordinates. Thus information content can be established in the three-dimensional space by using the set of independent coordinates (P) needed to fix a figure/molecule in this system.

Taking into account the general formula (2), the index of order Q can be written as follows

$$I_{geom} + Q = 1.$$

Let us analyze the information meaning of this expression. Suppose that a degree of freedom associated with the "defrosting" of chains or with isomerism or librations appears in the molecule. As this takes place, $P/3n$ increases, and, consequently, Q decreases; the information content increases as the index of order Q decreases. Then, while I_{geom} can be correlated with I_{sym} (see Figs. 6, 7) and, hence, with the Shannon entropy of information, the order index Q can be correlated with the Brillouin negentropy[32] or with excess information. The limiting case $Q = 1$ is matched by a regular polygon with an infinite number of sides, *i.e.*, a circle (see Table 1) or by an infinite chain with fixed SDF (see Fig. 1). Real molecules always possess the information content $P/3n$, this content being the greater, the closer to the beginning of the scale of order $Q = 0$, they are located.

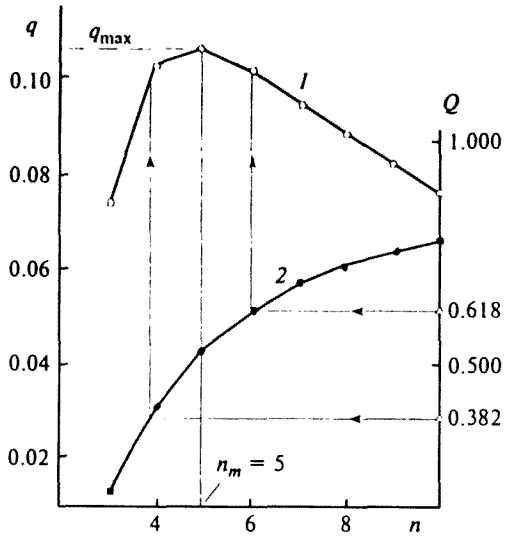


Fig. 8. Specific order index (*1*) and order index (*2*) in regular polygons. Here and in Figs. 9 and 10, the values 0.618 and 0.382 are marked, which correspond to the harmonic ratio of "golden cross-section"; arrows show the transition to the corresponding n and q values.

6. Specific order index. Above we introduced the notion of specific order index $q = Q/n$ and calculated q for various figures (see Tables 1—5). Now we shall consider the variation of the specific order index as a function of the number of atoms (n), symmetry, and dimensionality (1*D*, 2*D*, or 3*D*) of the molecule.

In the simplest case, *i.e.*, in regular polygons (see Table 1), an increase in n is accompanied by successive increase in the symmetry: $D_{nh} = D_{3h}(12)$, $D_{4h}(16)$, *etc.* As this takes place, as Fig. 8 indicates, the order index Q regularly increases, while the dependence of the specific order index $q(n)$ passes through a maximum, $q_{max} = 0.107$ at $n_m = n(q_{max}) = 5$. In the case of regular polyhedra (Fig. 9), no pronounced maximum is observed, but the highest values of q attained in this case are close to the q_{max} value observed for regular polygons and are also located in the $n = 4$—6 region.

In ordered chains, *i.e.*, in chain molecules in which the first interatomic distance and bond and torsion angles are fixed, the extremum q_{max} at $n_m = 5$ and 6 is also observed (Fig. 10). Following the successive defrosting of the structural degrees of freedom, the maximum value decreases and shifts toward lower n ($q_{max} = 0.067$ at $n_m = 5$ and $q_{max} = 0.042$ at $n_m = 4$, Fig. 11, curves *2* and *3*).

In the regular figures considered, the symmetry is not constant: it is D_{nh} for polygons and varies from T_d to I_h for regular polyhedra. In the case of ordered chains with fixed interatomic distances and angles, the symmetry varies from C_{2v} (if the initial triatomic unit is an isosceles triangle) to the alternation of C_{2h} (when the number of atoms in the chain is even) and C_{2v} (when the number of atoms is odd), on the assumption that the

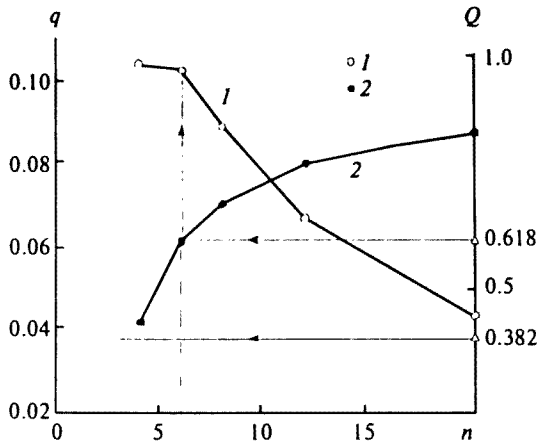


Fig. 9. Specific order index (*1*) and order index (*2*) in regular polyhedra.

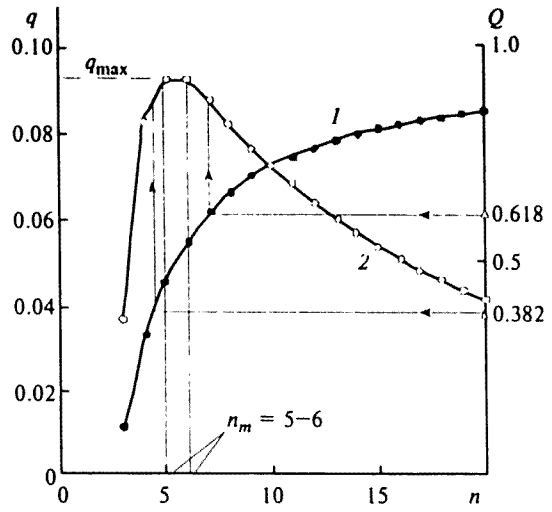


Fig. 10. Specific order index (*1*) and order index (*2*) in chains with fixed SDF, interatomic distances, bond angles, and torsion angles.

chains form 2*D*-structures. The other case, *viz.*, the study of the relationship between q and n in series of figures/molecules with identical types of symmetry, is also of interest. The results of the analysis of such series, together with the above-considered case of regular figures and chains, are presented in Table 8.

It can be seen from Table 8 that the presence of maxima on the $q(n)$ dependences at n_m = 4, 5, 6 is typical of all the groups considered. It is noteworthy that the $Q_m = Q(q_{max})$ values are located near the middle of the 0—1 scale of order indices. More precisely, they are located near or are confined between two numbers, 0.382 and 0.618, corresponding to the harmonic ratio of the "golden cross-section" (see, for example, Figs. 8—10). A similar situation has been observed[18] for the information topological indices of molecular graphs of heteroorganic compounds.

Now let us consider a fixed number of atoms in molecules by using, for example, the data of Table 5 for n = 4. The types of symmetry vary over a wide range, but 1*D*-, 2*D*- and 3*D*-structures can be distinguished. It can be seen from Fig. 12 that for all dimensionalities, the same maximum limiting value q_{max} = 0.104 is attained as the common position multiplicity increases.

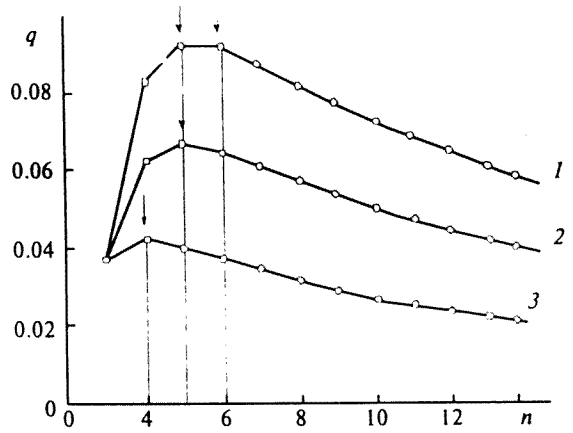


Fig. 11. Variation of the maximum specific order upon variation of the number of SDF.
The following parameters in the chains are constant: interatomic distances and bond and torsion angles (*1*); interatomic distances and bond angles (*2*); interatomic distances (*3*).

Table 8. Maximum values of the specific order index (q_{max}) on the $q = f(n)$ dependences and the corresponding numbers of atoms $n_m = n(q_{max})$ and order indices $Q_m = Q(q_{max})$

Parameter	Regular figures		Chains with frozen SDF	Molecules with identical types of symmetry					
	2*D*	3*D*		$O_h(48)$	$T_d(24)$	$D_{4h}(16)$	$D_{3h}(12)$	$C_{3v}(6)$	$C_{2v}(4)$
q_{max}	0.107	0.104	0.093	0.103	0.107	0.104	0.104	0.080	0.067
n_m	5	4 and 6	5 and 6	6	5	4	4	5	5
Q_m	0.533	0.517	0.512	0.617	0.533	0.417	0.417	0.400	0.267
$1 - Q_m$	0.467	0.483	0.488	0.383	0.467	0.583	0.583	0.600	0.733
$\frac{Q_m}{1-Q_m}$	1.14	1.07	1.05	1.610	1.14	0.72	0.72	0.67	0.36

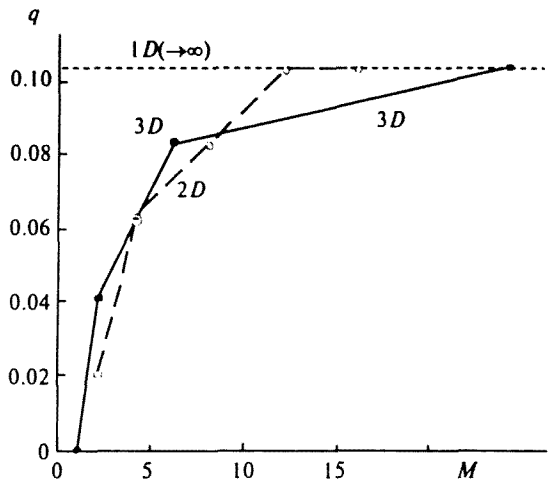


Fig. 12. Specific order index in tetraatomic molecules with various types of symmetry (M is the common position multiplicity).

For $2D$- and $3D$-molecules, the q value increases in a similar way as M increases.

In conclusion, we shall consider the order indices Q_m corresponding to the maximum values q_{max}. Firstly, they all (except for molecules with the lowest C_{2v} symmetry, which provide the only exclusion) fall into the limits of the golden cross-section ratio near the numbers 0.382 and 0.618. Secondly, a relative $Q_m/(1 - Q_m)$ value can be introduced, whose physical meaning is to indicate the chaos/order ratio. Two main classes are distinguished: groups in which order prevails, $Q_m/(1 - Q_m)>1$ (in Table 8, these are regular polygons, regular polyhedra, ordered chains, and molecules with the O_h and T_d types of symmetry) and groups in which disorder prevails (in Table 8, these are all groups with symmetry lower than O_h and T_d).

Let us consider two important results, associated with the specific order index and with the number of atoms $n_m(q_{max})$: firstly, the specific order index as a function of the number of atoms in a molecule normally passes through a maximum at n_m = 4—6, *i.e.*, on the scale of the first coordination sphere; secondly, as the number of varying structural degrees of freedom in polymeric molecules (interatomic distance, bond angle, torsion angle) increases, q_{max} decreases and nm shifts toward lower values.

Let us compare these results with the nonstrict crystal-chemical conclusions made from the strict localization theorem reported by Delone *et al.*[33] for a regular system of Delone.[34] Based on this simplified interpretation of the localization theorem, it may be claimed that long-range order is a consequence of short-range order.[35] If all atoms in a three-dimensional ($3D$) space are regularly surrounded within the limits of 3—4 coordination spheres, long-range order, *i.e.*, an ideal crystal arises.[35] In the case of $2D$ space, the Shtorgin theorem states that a regular environment within two coordination spheres is needed.[35] When the radius of the identical environment decreases, ideal crystal is not formed; instead, there occurs twinning, formation of various polymorphic modifications, *etc.*

Based on the foregoing, it may be concluded that if the radius of the identical environment decreases to the first coordination sphere, either separate molecules are formed rather than condensed polymeric materials (*i.e.*, long-range order does not arise either because the size of the regular environment in the molecule is too small) or long-range order is absent because a condensed noncrystalline state with a short-range order is formed. The latter can be clearly seen in the case of polymeric $1D$ molecules: as $Q_m = Q(q_{max})$ and q_{max} decrease and the number of SDF increases, the $nm(q_{max})$ number of atoms decreases in the series 6→5→4 (see Fig. 11).

Thus, molecules and noncrystalline substances can be defined as systems in which the regular environment of atoms is observed at distances not exceeding the size of the first coordination sphere of short-range order. If the regular environment is disturbed even in the first coordination sphere, only noncrystalline substances can be formed.

* * *

We developed a geometrical approach to the description of the order/disorder in various molecular systems and found certain regularities in the variation of the order index Q as a function of the number of atoms in a molecule, its symmetry, the number of structural degrees of freedom, and dimensionality. It follows from the main equation $Q = 1 - P/3n$ that $P/3n$ can be regarded as a relative index of disorder in the system, while from the viewpoint of the theory of information, the $P/3n$ parameter (P is the number of independent variables needed to fix a molecule consisting of n atoms in the Cartesian system of coordinates) correlates with the Shannon entropy of information and can be interpreted as a geometrical information index. The order index Q correlates with negentropy or excess information. Due to the nonprobabilistic method for the evaluation of information used by us (since the Q values were calculated only in terms of the analytical geometry), the relationship between the Cartesian three-dimensional space (normalization of P by $3n$) and information space (the normalizing factor used for I^m_{sym} is $n\log_2 n$) can be established as a correlation between the geometrical and symmetry information indices. Therefore, our study can be related not only to the structural chemistry of molecules but also to the branch of mathematical chemistry that is based on the theory of information.

The potential of the approach used by us is by no means exhausted within the framework of this study in which the following items were simplified or were not considered at all. First of all, this method for the calculation of the order index in this particular form is applicable only to rigid structures, whose symmetry does not vary. Nevertheless, we managed to use it for the

analysis of nonrigid molecules having degrees of freedom like inversions and librations. The development of this method toward its application to dynamic structures seems fairly interesting.

We also simplified *a fortiori* a practically important problem, *i.e.*, the composition—structure—property correlation. Despite the fact that this problem is intuitively clear and seemingly simple, it is quite complex and, in general, has not been solved. In this connection, it is pertinent to mention, for example, the permanent discussion dealing with the fundamental impossibility of unambiguously predicting structure from data on the chemical composition.

We did not consider the special case of quasi-crystals as an example of icosahedral systems with long-range orientation order in the absence of translational symmetry. Complex systems with a chaotic arrangement of molecules or systems with oriented molecules like liquid crystals also were not analyzed.

It should be noted that the common classification into organic chemistry, in which molecules predominate, and inorganic chemistry, in which complex framework solid substances and compounds predominate, is fairly arbitrary, although it reflects largely the specific character of these fields of chemistry. This division is removed if we proceed to the consideration of mesomolecules and supramolecules (dendrites, cascade systems, molecular tectonics, *etc.*) in organic and inorganic systems in which the covalent bond is no longer significant; in this case, the levels of the arising problems and the ways for tackling them are similar.[36,37] The description of supra- and mesomolecules should be associated with the use of the theory of information.[37] Therefore, we hope that the use of the method that we propose for the evaluation of order/disorder in molecular systems would permit a fresh glance at some problems relating to timely branches of modern science.

The authors wish to sincerely thank Acad. Yu. A. Zolotov, E. M. Dianov, and Yu. D. Tret'yakov for stimulating discussion, Prof. Yu. S. Nekrasov for a series of useful advice, remarks, and discussions. Doctors of Science (Chemistry) A. A. Pasynskii, A. E. Gekhman, and I. L. Eremenko, Doctor of Science (Physics and Mathematics) I. P. Zvyagin, and Ph. D. (Chemistry) A. S. Zyubin for fruitful discussion. The authors are also grateful to Prof. D. G. Bonchev, who supported this work.

This work was carried out with the financial support of the Russian Foundation for Basic Research (Project No. 96-03-33563).

References

1. G. Careri, *Poryadok i besporyadok v strukture materii* [*Order and Disorder in the Structure of Matter*], Mir, Moscow, 1985, 228 pp. (Russ. Transl.) [G. Careri, *Ordine e disordine della materia*, Laterza, Roma, 1982].
2. I. Prigozhin, *Ot sushchestvuyushchego k voznikayushchemu* [*From the Existing to Nascent*], Nauka, Moscow, 1985, 327 pp. (in Russian).
3. I. Prigozhin and I. Stengers, *Vremya, khaos i kvant* [*Time, Chaos, and Quantum*], Progress, Moscow, 1994, 272 pp. (in Russian).
4. Yu. L. Klimontovich, *Turbulentnoe dvizhenie i struktura khaosa: Novyi podkhod k statisticheskoi teorii otkrytykh sistem* [*Turbulent Motion and the Structure of Chaos: A Novel Approach to the Statistical Theory of Open Systems*], Nauka, Moscow, 1990, 317 pp. (in Russian).
5. J. M. Ziman, *Models of Disorder: The Theoretical Physics of Homogeneously Disordered Systems*, Cambridge University Press, London, 1979.
6. S. R. Elliott, *Nature*, 1991, **354**, 445.
7. S. R. Elliott, *J. Phys., Condens. Matter*, 1992, **4**, 7661.
8. E. A. Chechetkina, *J. Phys., Condens. Matter*, 1995, **7**, 3099.
9. A. I. Kitaigorodskii, *Poryadok i besporyadok v mire atomov* [*Order and Disorder in the World of Atoms*], Nauka, Moscow, 1984, 172 pp. (in Russian).
10. C. Shannon and W. Weaver, *Mathematical Theory of Communication*, Univ. Illinois, Urbana, 1949.
11. *Chemical Applications of Topology and Graph Theory*, Ed. R. B. King, Elsevier, Amsterdam, 1983.
12. Yu. A. Zhdanov, *Entropiya informatsii v organicheskoi khimii* [*Entropy of Information in Organic Chemistry*], Izd. RGU, Rostov, 1979, 55 pp. (in Russian).
13. D. G. Bonchev, *Kharakterizatsiya khimicheskikh struktur s pomoshch'yu teorii informatsii i teorii grafov* [*Characterization of Chemical Structures Using Information Theory and Graph Theory*], Sc. D. Thesis (Chemistry), Moscow State University, Moscow, 1984, 409 pp. (in Russian).
14. D. Bonchev, D. Kamensky, and V. Kamenska, *Bull. Math. Biol.*, 1976, **38**, 119.
15. D. Bonchev and G. Lickomanov, *Math. Chem.*, 1977, **3**, 3.
16. D. Bonchev, *Information Theoretic Indices for Characterization of Chemical Structures*, Research Studies Press, Chichester, 1983, 249 pp.
17. M. I. Stankevich, I. V. Stankevich, and N. S. Zefirov, *Usp. Khim.*, 1988, **57**, 337 [*Russ. Chem. Rev.*, 1988, **57** (Engl. Transl.)].
18. Yu. S. Nekrasov, E. E. Tepfer, and Yu. S. Sukharev, in *Prilozhenie teorii grafov v khimii* [*Applications of the Graph Theory to Chemistry*], Izd. IM SO RAN, Novosibirsk, 1994, **151**, 8 (in Russian).
19. V. A. Skorobogatov, E. V. Konstantinova, and Yu. S. Nekrasov, *Math. Chem.*, 1991, **26**, 215.
20. A. N. Kolmogorov, *Vest. Akad. Nauk SSSR* [*Commun. of the USSR Academy of Sciences*], 1965, **5**, 94 (in Russian).
21. A. N. Kolmogorov, *Probl. Peredachi Inform.* [*Problems of Information Transfer*], 1969, **5**, 3 (in Russian).
22. R. S. Ingarden and K. Urbanek, *Colloq. Math.*, 1962, **9**, 131.
23. P. A. Koz'min, *Dokl. Akad. Nauk*, 1995, **343**, 205 [*Dokl. Chem.*, 1995, **343** (Engl. Transl.)].
24. P. A. Koz'min, *Zh. Neorg. Mater.* [*J. Inorg. Materials*], 1995, **31**, 884 (in Russian).
25. S. A. Dembovsky and E. A. Chechetkina, *Stekloobrazovanie* [*Glass Formation*], Nauka, Moscow, 1990, 264 pp. (in Russian).
26. I. Hargittai and M. Hargittai, *Symmetry Through the Eyes of a Chemist*, Plenum, New York, 1995.
27. R. Gillespie and I. Hargittai, *The VSEPR Model for Molecular Geometry*, Allyn and Bacon, Boston—London—Toronto—Sydney—Tokyo—Singapore, 1991.

28. D. Hohl, R. O. Jones, R. Car, and M. Parrinello, *J. Chem. Phys.*, 1989, **89**, 6823.
29. R. Gillespie, *Molecular Geometry*, Van Nostrand Company, London, 1972.
30. W. Ashbey, *An Introduction to Cybernetics*, Wiley, New York, 1956.
31. V. M. Glushkov, in *Kibernetika, myshlenie, zhizn'* [*Cybernetics, Thinking, and Life*], Mysl', Moscow, 1964 (in Russian).
32. L. Brillouin, *Science and Information Theory*, Academic Press, New York, 1966.
33. B. N. Delone, N. P. Dolbilin, M. I. Shtogrin, and R. V. Galiulin, *Dokl. Akad. Nauk SSSR*, 1976, **227**, 19 [*Dokl. Mathem.*, 1976, **227** (Engl. Transl.)].
34. R. V. Galiulin, *Kristallografiya* [*Crystallography*], 1980, **25**, 91 (in Russian).
35. R. V. Galiulin, *Kristallografiya* [*Crystallography*], 1994, **39**, 581 (in Russian).
36. G. D. Mendenhall, A. Greenberg, and J. F. Liebman, *Mesomolecules: From Molecules to Materials*, Chapman and Hall, New York, 1995.
37. J.-M. Lehn, *Supramolecular Chemistry: Concepts and Perspectives*, VCN, Weinheim, 1995.

Received May 13, 1996